



## A Hybrid U-Net–GAN Framework for Restoring Broken Kannada Handwritten Characters

Chandravva Hebba<sup>\*</sup>, H. R. Mamatha<sup></sup>

Department of Computer Science and Engineering, PES University, Bangalore 560085, India

Corresponding Author Email: [hebbichandravva@gmail.com](mailto:hebbichandravva@gmail.com)

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.301220>

### ABSTRACT

**Received:** 6 October 2025

**Revised:** 10 December 2025

**Accepted:** 24 December 2025

**Available online:** 31 December 2025

#### **Keywords:**

*adversarial refinement, U-Net, Generative Adversarial Network, handwriting restoration, optical character recognition*

Restoring broken or degraded handwritten characters is a major obstacle in optical character recognition (OCR) and the digital preservation of historical manuscripts. In this paper, we propose a hybrid framework named U-Net + Generative Adversarial Network (GAN) for Kannada Restoration (UNGAN-KR) for broken handwritten characters. It integrates the best of both encoder–decoder structural reconstruction and adversarial refinement to ensure both pixel-level fidelity and perceptual realism. The U-Net restores broken strokes while ensuring preservation of character structure and the GAN discriminator promotes natural handwritten textures. We evaluated the framework on a dataset of 71,149 handwritten Kannada characters using multiple metrics of accuracy. The experimental results show that our proposed framework achieves an accuracy of 97.8% with improvements in perceptual quality, and outperforms benchmarks like Convolutional Neural Network (CNN) autoencoders, standard U-Net, and GAN-based inpainting. Ablation studies show that the integration of U-Net and GAN provides hybrid enhancements that are important for reconstruction accuracy. Thus, the framework is suitable for pre-processing data and digital sustainable archiving and the automated restoration of degraded Kannada manuscripts.

## 1. INTRODUCTION

Maintaining handwritten documents has served as a key aspect of digital archiving, cultural heritage preservation, and smart information retrieval. Handwriting recognition methods have proven reliable to work with clean datasets. But the degeneration of the handwriting and the breaking of character strokes generally is not always resolvable completely reliably. More recent work focused on deep generative models has broadened the examination of recovering and enhancing degraded text images built on top of a variable cycle Generative Adversarial Network (GAN) framework. It learns to predict the deformation of handwritten text and has shown that adversarial learning can successfully recover the conformity of characters in degraded handwritten documents. Researchers have approached text enhancement for historical handwritten manuscripts addressing challenges such as ink bleed, paper degradation, and incomplete strokes. These investigations remark on bringing attention to two challenges where recovering or enhancing broken text is not a pre-processing step. It is a substantive task for providing usability in downstream tasks of retrieval and recognition of, not to mention digital archive.

Generative deep learning models have surfaced as good solutions to solve these issues. Nigam et al. [1] utilized variable cycle GAN to remove deformity in handwritten text, proposing a restoration identified by a structural fingerprint of the original writing style. Alaasam et al. [2] engaged text enhancement for historical handwritten documents to address issues such as noise from the background paper on ink and

partial loss of characters. They also noted that any pre-processing and restoration of text would be informative for the downstream tasks of recognition. Rabhi et al. [3] presented a multilingual recovery framework with a convolutional denoising autoencoder that incorporated attentional mechanisms. They illustrated the importance of the focus mechanism in dealing with different scripts. However, while their model has some level of cross-linguistic generalization, it does not address the stroke discontinuities that are characteristic of Kannada.

Structural complexity is a notable challenge, characters in Kannada are very curvilinear, and they often have compound parts. They recognize the limitations of this process are on the use of incomplete or broken strokes to recognize characters, meaning that the process can be sensitive to noise degradation. Gongidi and Jawahar [4] have suggested hybrid implementations of feature extraction methods and ensemble learning for Kannada recognition. This increases their susceptibility to recognition errors when even slightly degraded. For instance, Ramesh et al. [5] have carried out Kannada recognition using the combination of manifold smoothing and label propagation. The methods can classify strokes similarly assuming a similar representation from the character, but the methods require complete samples and are not helpful if the necessary strokes are missing.

Advances in generative models using disentangled representation learning for handwritten text generation [6]. For instance, trajectory recovery attributes (AIOU and LDTW) [7] were designed for complicated handwriting. It shows how trajectory aware evaluations emphasize the importance of

context in trajectory supported evaluation as a section of handwriting continuation. Their results clearly showed that incomplete and broken strokes could not be fully recovered with image enhancement as a process but required contextual trajectory prediction of the strokes. The findings of the GAN study [8] have also established that while adversarial training is strong on style preservation and realism, the method struggles with the fine-grained elements of structure fidelity, an important requirement to restore complex Kannada characters.

Many studies on written handwriting recovery and regeneration are multilingual or work mostly on Latin scripts. There are few models that engage with low-resource scripts such as Kannada, where stroke recognition can be quite sensitive to minor differences in stroke details. Elanwar and Betke [9], proposed a variable attention-based Bi-LSTM framework for Kannada script recognition which improved classification. However, it did not propose a restoration mechanism for degraded inputs. Consequently, there was an obvious gap in the research, there is currently no robust pipeline that incorporates structural stroke regeneration and recognition for Kannada handwritten documents. JokerGAN [10] outlined advancements for memory resources for congeniality of text-line aware generation for handwritten text generation respectively, but in both cases the models are used for generating new handwritten samples, not repairing broken ones.

The proposed broken character regeneration framework is designed as an image-to-image translation method, wherein incomplete Kannada characters should be regenerated in their original form prior to recognition. The proposed hybrid UNGAN-KR framework takes advantage of U-Net's encoder-decoder architecture to support a careful and localized reconstruction of structural features. Those are necessary for stroke continuity while the GAN discriminator enforces global realism and stylistic fidelity. The potential of the proposed UNGAN-KR framework in this regard is its ability to support the regeneration of Kannada's complexity in morphology, specifically curvilinear shapes, ligatures, and composites. Comprehensive experimental work was done in evaluating our framework against Kannada recognition baselines and state-of-the-art generative handwriting approaches.

The principal contribution of the paper is summarized as follows:

- Modeled broken character regeneration as an image-to-image translation task where generating a broken character image to its original form using the structural learning properties of U-Net combined GAN adversarial refinement (UNGAN-KR).
- Designed a Framework using a U-Net backbone combined with a GAN discriminator ensuring both stroke-level correctness and contextual realism on each generated character.
- This framework is specifically designed for Kannada, targeting its complex morphology and compound character shapes.
- Compare the framework against recent Kannada recognition models and generative handwriting work, and show that the generated character images with our methodology have higher accuracy when dividing both the regeneration accuracy and recognition accuracy.

The rest of this document is organized in the following manner: In Section 2, we present a thorough survey of the Related Work on handwritten character restoration, U-Net

based architectures, and GAN based inpainting approaches. Section 3 describes the Proposed Method, which includes a discussion on six step workflows for restoring Kannada handwritten broken characters along with mathematical formulations and a hybrid UNGAN-KR framework. The setup for the experiments is discussed in Section 4, which details the dataset, metrics for evaluation, implementation details, followed by Results and Discussion. Lastly, Section 5 concludes the paper and discusses future work on multi-script handwriting restoration and perceptual enhancement.

## 2. LITERATURE REVIEW

The number of digitization initiatives has surged in recent years, leading to a need for accuracy in recognizing and regenerating handwritten text. However, deciphering handwritten characters can be challenging, given that they can have disadvantages such as ink bleeding, pen stroke fading, and partial degradation arising from aging and writing inconsistencies. This is particularly true with more complicated lettering systems, such as Kannada, where stroke structures can compound the challenges of readability. In recent years, image inpainting and structure-aware restoration have propelled the design of stroke regeneration. Dong et al. [11] have introduced AU-GAN, a U-shaped autoencoder GAN that directly combines the architectural benefits of U-Net and adversarial refinement to generate coherent structure and plausible textures for image inpainting. The AU-GAN framework represents the potential of a symmetric design for encoder-decoder with skip connections to preserve fine structural details while encouraging adversarial loss to produce outputs consistent with photorealism. These principles have influenced the use of a U-Net generator and a Patch-style discriminator as denoted in the last paragraph for stroke regeneration. Further, transformer paradigms have been applied to the inpainting problem.

Elharrouss et al. [12] have provided an exhaustive treatment of transformer-based image and video inpainting by discussing research success and current limitations while modeling long-range dependencies with transformers. Elharrouss et al. [12] suggested that transformer modules could be incorporated in convolutional encoders when longer range contextual reasoning across distant stroke fragments is required. Chen et al. [13] have considered DNNAM, an attention-augmented deep network for inpainting that directs the reconstruction capability toward semantically significant regions with attention mechanisms and their attention design and loss formulations informed attention gates incorporated within U-Net skip connections to focus restoration toward stroke-missing regions. In the area of identity-aware inpainting, Li et al. [14] have proposed identity-aware inpainting for occluded face recognition, demonstrating how domain-specific identity losses (identity-preserving) allows for the preservation of identity attributes in the reconstructed images; in a similar vein, we adopt a notion of character-identity in via skeleton/structure consistency loss to preserve character topology, thereby ensuring that the regenerated strokes plausibly exhibit recognizable character identity.

Yeh et al. [15] have proposed a structure-and-texture splitting paradigm that facilitates the use of separate pathways and losses to reconstruct the coarse geometry (structure) and finer texture (ink) independently; this separation inspires our methodology that allows us to put together the two losses

(stroke-topology / skeleton loss and texture/perceptual/adversarial losses) to return both continuity of strokes and appearance of ink back into the restoration. The generative handwriting synthesis and GAN-based augmentation literature provides insight on conditioning and style preservation, and offers GAN-related information for adversarial stabilization that can inform training restoration architectures. Chang et al. [16] have demonstrated a case of a GAN for cross-lingual handwritten text generation, that showed how the conditioning of the text or style does allow for generation across scripts and styles; their conditional synthesis paradigm informs the design of our conditional discriminator which factors local stroke context and global style estimates to define the input-output mapping.

Recent work aimed at improving realism in generated handwriting, specifically the GAN-based enhancement work by Dubey et al. [17] have provided pragmatic considerations related to stabilizing GAN training on handwriting images and generating reasonable stroke textures—lessons we apply in terms of the adversarial vs. reconstruction loss balance. Studies on stroke trajectories and topology, as represented by Hanif and Latecki [18] have demonstrated an algorithm and automatic evaluation metrics for recovering stroke trajectories in unconstrained handwriting. The trajectory recovery perspective illustrates that evaluating regenerated outputs should include not just evaluating pixels (PSNR/SSIM), but also evaluating trajectory or skeleton fidelity which led to using skeleton F1 and trajectory alignment metrics for evaluation in our experiments.

Datasets and domain-specific resources for both Kannada handwriting data and palm-leaf manuscripts provide an important substrate for training and evaluation. The benchmark Kannada handwritten document dataset introduced by Alaei et al. [19] is still a foundational text and image resource analyzing isolated and unconstrained Kannada text; their segmentation and baseline protocols provide important guidance for our own preprocessing and evaluation splits. To address the demand for historical corpora, Sajjan et al. [20] have provided resources that focused on ancient Kannada materials. It is a well-curated palm-leaf dataset and a PyTesseract based optical character recognition (OCR) application study for palm-leaf manuscripts both resources are a valuable way to assess regeneration on real degraded samples instead of synthetic corruption. Beyond studies of inpainting, handwriting, and diffusion-based generation, there are also advances in font generation from which we can derive useful lessons about the restoration of strokes and characters.

Kong et al. [21] have presented a one-shot font generation framework that applies a component-based discriminator to help a generative model synthesize unseen characters from very few training examples. Their key contribution in this work is the use of component-level supervision in which characters are decomposed into more primitive substructures (e.g., radicals, strokes, or glyph components) to ensure that generated fonts not only resemble the global shape but that they also maintain fine-grained structural integrity. The component-level discriminator design will be particularly applicable in the case of handwritten script regeneration where the characters can be made up of multiple strokes that may be missing or fragmented. Bannigidad and Sajjan [22] have researched specific image enhancement pipelines for ancient Kannada palm-leaf manuscripts and also noted that particular domain-specific noise patterns (fibrous type, ink fade, and physical abrasions) would require pdf strengthening; we also

apply similar degradation simulations when generating broken-stroke from training pairs. Related to recognition processes, Parashivamurthy and Rajashekaradhy [23] have proposed a serial dilated cascade network specializing for Kannada scripts and reported strong recognition; these recognition backbones provide downstream evaluators for regarding quantification of the practical applications of the OCR on a particular area.

Research of cross script augmentation and heritage restoration study exhibit the implications of functional utility of targeted data augmentation and models of cross-domain transfer for low-resourced scripts. Wang et al. [24] have proposed an AI-based method for restoring Chinese manuscripts via contemporary deep models that recover ancient writing and the outcomes indicate that we can achieve good restoration on highly worn cultural heritage objects by utilizing domain-specific heuristics and multi-stage restoration pipelines, an idea we modify for historical Kannada texts. Sareen et al. [25] have found that Convolutional Neural Network (CNN)-based data augmentation improved recognition with Gurumukhi (another Indic script) and that any form of augmenting data such as GAN or synthetic broken stroke generation would be beneficial for low resource scripts. They developed a broken-stroke generator as an augmentation module that incorporates realistic degradation styles based on Kannada stroke properties.

Recent advances in generative modeling pertaining to handwriting synthesis and styling have taken the form of transformers, deformable convolutions, single image augmenters, diffusion models, and archetype-based generating, which provide a complement to the restoration and augmentation. Wang et al. [26] have merged transformers with deformable convolution to devise like samples of handwriting, and showed that transformer attention and deformable convolution receptive fields better capture varied deformations in handwriting. This implied a hybrid design may be useful in future additions wherein transformer blocks could one day be incorporated into the U-Net encoder to capture long-range styles of strokes. OffSig-SinGAN created by Hameed et al. [27] and diffusion-based signature generation by Hong et al. [28] have exemplified powerful single-sample augmentation methods and handwriting models respectively. Therefore, OffSig-SinGAN and diffusion models are suitable ways to enhance infrequent variants of a writer's style or create even difficult low-quality samples.

Liao et al. [29] have harnessed the power of diffusion modeling through Calliffusion for Chinese calligraphy style transfer, further illustrating that diffusion processes are useful for generating high-quality strokes and styles. Although our work is based on adversarial refinement, diffusion modeling presents a viable alternative for examining stroke completion in the future. Pippi et al. [30] have described handwritten text generation from visual archetypes at CVPR and presented archetype-based conditioning. That potentially preserves structural priors is similar to the archetype idea we propose in the context of stroke-generate with U-Nets. The canonical stroke templates or skeletal priors help to drive the reconstruction of strokes in the region with the U-Net.

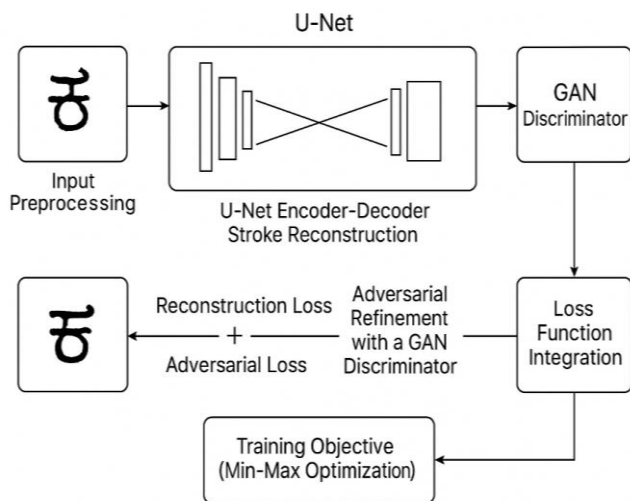
Despite considerable developments in image-to-image translation and character restoration, there are multiple shortcomings in the existing literature. Most work has focused on Latin and Chinese scripts with limited research on other writing systems and particularly with complex scripts, such as Kannada, researched infrequently. Although U-Net

architectures are able to preserve structural information, they typically do not preserve fine-grained stroke texture, and while GAN-based methods improve perceptual realism, they may induce artifacts. Few studies have proposed combined methodologies that assure some degree of both structural fidelity and perceptual quality at once. In addition to this, examples of quantitative multi-metric evaluation scores are also limited, and the problem of restoring completely unrecognizable or occluded characters remains challenging. Existing frameworks do not generalize across different styles of handwriting well, which is compounded by high computational costs, especially for real-time restoration and deployment on devices with limited resources. Thus, there is an opportunity to develop a robust, efficient and generalizable framework for Kannada handwritten character restoration of high quality.

### 3. PROPOSED WORK

This work conceptualizes the restoration of damaged handwritten Kannada characters as an image-to-image translation problem. The architecture consists of a U-Net generator and a GAN discriminator to provide structural stroke restoration combined with texture-level realism. The U-Net operates as a structural restorer that combines local continuity and global context through its encoder-decoder with skip connections. The GAN discriminator works in conjunction with the U-Net to ensure the restored characters resemble authentic handwritten samples. This dual approach addresses both the challenges of restoring missing strokes and improving the perceptual fidelity of the restored characters.

The suggested workflow for Kannada handwritten broken character regeneration is structured in the following six steps also shown in Figure 1.



**Figure 1.** Proposed UGAN-KR model design diagram

- Input Pre-processing
- U-Net Encoder–Decoder Stroke Reconstruction
- Adversarial Refinement with a GAN Discriminator
- Loss Function Integration
- Training Objective (Min–Max Optimization)
- Validation and Preservation of Character Identity

The proposed method is described as follows: The UGAN-KR framework proposes a new way to restore broken handwritten text in Kannada through a method called

image-to-image translation. This new framework solves two major problems together: 1) recovering the original handwriting from the damaged image and 2) creating a realistic-looking final product that mimics genuine handwritten text. To prepare for image-to-image translation using our system, the original damaged images are first undergone a normalizing transformation, followed by binarization via Otsu Thresholding, then the damaged characters are processed via morphological closing so that all stroke pieces are connected, followed by normalization via image resampling so that all character images are the same size. Once the original image has been normalized and transformed, it will become the input into our U-Net generator.

Inside the U-Net generator, the encoder is utilized to extract hierarchical stroke features from the images. The encoder also encodes the overall character structure into a compressed latent representation. The decoder will use multiple up-sampling layers and skip connections to reconstruct the missing strokes from the compressed latent representation and the low-level details of the image to create a complete representation of the character. Once the U-Net generator has created an initial version of the character, the initial version is enhanced via adversarial learning, where a GAN discriminator attempts to differentiate between the original handwriting and the newly generated handwriting. This process creates a more realistic final product in terms of stroke continuity, stroke texture, and stroke style.

The neural network that we just described uses a composite loss function to guide its training, which integrates pixel-level reconstruction loss, structural similarity loss, perceptual feature loss, and adversarial loss in order to find a good balance between reconstructing the character's original appearance while also generating a realistic final product. Additionally, the GAN discriminator's gradient penalty and feature-matching losses provide additional stability to the adversarial optimization process.

A joint training process for both the generator and discriminator is performed using the Adam optimizer, applying learning rate decay, dropout regularization, and gradient clipping to facilitate convergence throughout the training process. Constraints include class consistency (a measure of how well samples from the same class can be classified) based upon prior trains learned by a CNN classifier, and stroke-level consistency through skeletal representation; resulting in a restoration of broken handwritten Kannada characters that maintain structural coherence, appear visually authentic, and preserve the identity of original Kannada characters. With this integrated system, UGAN-KR will produce high structural accuracy as well as visually authentic restored handwritten Kannada characters while also preserving character identity.

#### 3.1 Input preprocessing

The first and fundamental stage in regenerating broken Kannada handwritten characters is input preprocessing, which aims to standardize the raw manuscript images and prepare them for deep learning-based reconstruction. Handwritten documents often suffer from diverse distortions, including variable stroke thickness, ink smudges, missing segments, and uneven illumination, which can significantly degrade the performance of neural networks. To address these challenges, the raw character images  $I(x, y)$  are initially normalized using Eq. (1), ensuring pixel intensity values are standardized with

zero mean and unit variance. This normalization mitigates the impact of variations in lighting and ink density across different document sources, enabling consistent feature extraction.

$$I_{norm}(x, y) = \frac{I(x, y) - \mu}{\sigma} \quad (1)$$

Here,  $I(x, y)$  indicates the original pixel intensity value at  $(x, y)$  of the grayscale image. The term  $\mu$  indicates the mean pixel intensity across the entire dataset or batch of images, which indicates that the image is zero-centered. The letter  $\sigma$  indicates the standard deviation of the pixel intensities, which ensures that the pixel intensities are scaled so that the variance is normalized. The image is then normalized to yield  $I_{norm}(x, y)$ , where each pixel is subsequently less sensitive to changes in illumination and the thickness of writing. Normalization is important for the Kannada handwritten broken character regeneration framework because it facilitates the learning of structural patterns by the U-Net encoder-decoder and GAN discriminator without their learning being confounded by the raw intensity information.

Subsequently, we apply the binarization step (Eq. (2)) that uses Otsu's thresholding to separate foreground strokes from the background, moving from the normalized grayscale image to a binary version  $I_{bin}(x, y)$ . This step is crucial to isolate handwritten strokes from the background while still maintaining the important structural information of each character.

$$I_{bin}(x, y) = \begin{cases} 1 & \text{if } I_{norm}(x, y) > T_{otsu} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Next, we perform morphological operations, specifically closing operations (Eq. (3)), to reconnect broken strokes and fill small gaps within characters. Morphological operations help to enhance the continuity of strokes in handwriting, which is especially significant with Kannada characters that naturally contain many complex curves and ligatures.

$$I_{closed} = (I_{bin} \oplus B) \ominus B \quad (3)$$

Here,  $I_{bin}$  is used to define the binarized input image, where text pixels are separated from the non-text background. The symbol  $\oplus$  refers to morphological dilation, while  $\ominus$  refers to morphological erosion, both of which are done using a structuring element BBB. The image  $I_{closed}$  is the morphologically closed image where gaps in strokes of Kannada text are filled, and disconnections were minimized. Eq. (4) describes the process of resizing the input character image.

$$I_r = \text{Resize}(I_c, H \times W) \quad (4)$$

Here,  $I_c$  denotes the character image once cropped, and  $H$  and  $W$  represent the desired resized height and width. The output image  $I_r$  is the resized image for a standardized uniform image for input to the U-Net encoder.

### 3.2 U-Net encoder-decoder stroke reconstruction

The demonstrated framework for regenerating handwritten broken Kannada characters is built on a U-Net encoder-decoder architecture, which is responsible for reconstructing the missing or broken strokes while also duplicating the

internal structure and style of each character. The U-Net architecture is well suited for the task because of its symmetric encoder-decoder architecture with skip connections, which allows the network to capture local fine-scale stroke features and the larger structural context.

The encoder sequentially abstracts hierarchical features from the input image  $I_{resized}$ , which has been preprocessed, through an encoder-decoder architecture with skip connections consisting of just convolutional layers (Eq. (5)).

$$f_l = \sigma(W_l * f_{l-1} + b_l) \quad (5)$$

Here,  $f_{l-1}$  is the input feature map from the previous layer,  $W_l$  and  $b_l$  are the learnable convolutional weights and bias in layer  $l$ , and  $*$  denotes convolution. The activation function  $\sigma(\cdot)$  adds non-linearity, and the resulting feature map  $f_l$  is a set of hierarchical representations of the character of Kannada.

Each convolutional layer is followed by a non-linear activation function, to enable the network to learn complex stroke features. Down sampling shown in Eq. (6) via maximum pooling layers reduces spatial dimensions and increases the size of the receptive field enabling the encoder to capture dependencies across the character structure.

$$f_l^{down}(i, j) = f_l(i + m, j + n) \quad (6)$$

Here,  $f_l(i + m, j + n)$  represents the feature activations in the local neighborhood  $\Omega$  around position  $(i, j)$ . The output  $f_l^{down}(i, j)$  displays the maximum value in that region that maintains dominant stroke features but decreases spatial resolution. In the bottleneck layer, the network produces a compact latent representation using  $z$  in Eq. (7) of the essential features of the character including stroke orientation, curvature, and connectivity.

$$z = \text{Encoder}(I_{resized}) \quad (7)$$

Here,  $I_{resized}$  is the input image of the Kannada character standardized impostor character, and the encoder generates a compressed feature embedding. The latent vector  $z$  carries the high-level stroke pattern and graphical structure information to facilitate decoding and reconstruction. To reconstruct the character, the decoder up samples the latent representation Eq. (8).

$$f_l^{up} = \text{Upsample}(f_{l+1}) \quad (8)$$

Here,  $f_{l+1}$  is the feature map from a deeper level, and the Upsample function increases its spatial resolution. The  $f_l^{up}$  feature map is able to reconstruct the fine stroke details of the Kannada character while also leveraging the essence of the encoder  $f_l^{enc}$  feature map. Adding skip connections to the encoder feature maps congruent to its spatial dimensions Eq. (9). The skip connection concatenates information regarding the fine-size details lost during down sampling, enabling the reconstruction of the strokes accurately.

$$f_l^{fusion} = f_l^{up} \oplus f_l^{enc} \quad (9)$$

Here,  $f_l^{up}$  is the upsampled decoder feature map and  $f_l^{enc}$  is the related feature map from the encoder. The fusion operator  $\oplus$  denotes either concatenation or addition, this enables the model to mix high-level semantic information with

low-level stroke information to accurately regenerate the Kannada character. The last convolutional layer produces the final reconstructed character image  $I_{rec}$  Eq. (10) that is a smooth visual coherence regeneration of the original broken input.

$$I_{rec} = \sigma(W_{out} * f_0 + b_{out}) \quad (10)$$

Here,  $I_{rec}$  indicates the resulting regenerated image of the U-Net decoder.  $f_0$  is the fused feature map of all the features,  $W_{out}$  and  $b_{out}$  are the learnable weights and bias for the output layer, and  $*$  indicates convolution. The activation function  $\sigma(\cdot)$  ensures the resulting pixel values of  $I_{rec}$  reside in the desired ranges, resulting in a regenerated character.

The U-Net reconstruction phase will provide the framework for subsequent adversarial reconstruction - and while it may be an approximation and a high-fidelity regeneration of the Kannada handwritten strokes, it is enough for the next generation of GAN based learning. The encoder's hierarchical feature extraction and the decoder's ability to reconstruct structurally provides assurance that both the macro employment of character shapes is reconstructed along with micro punctuations within strokes, allowing realistic character regeneration in future adversarial GAN stages.

### 3.3 Adversarial refinement with Generative Adversarial Network discriminator

Although the U-Net encoder-decoder generates an initial reconstruction of broken Kannada handwritten characters, it may still incorporate minor artifacts or unrealistic stroke connection, especially within complex ligatures or occluded areas. Therefore, to improve realism and enforce structural consistency, the generated characters are further refined using a GAN framework. In this adversarial scenario, the generator  $G$  will correspond to the U-Net reconstruction network and we introduce an additional discriminator  $D$  to assess the realism of characters generated.

The discriminator  $D$  will output a probability score  $D(I)$  given an image  $I$ , representing the probability that the image was real from the ground-truth dataset rather than generated. This can also be formally expressed as Eq. (11).

$$D(I) = \sigma(W_d * I + b_d) \quad (11)$$

Here  $W_d$  is the convolutional kernels of the discriminator,  $b_d$  is the bias term,  $*$  is convolution notation, and  $\sigma$  is simply a non-linear activation. The GAN's adversarial objective is the usual min-max problem given in Eq. (12).

$$V(D, G) = E_{I \sim p_{data}}[\log D(I)] + E_{I \sim p_G}[\log(1 - D(I_{rec}))] \quad (12)$$

Here,  $D$  is the discriminator,  $G$  is the generator,  $I$  is the true images from the true data distribution  $p_{data}$  and  $I_{rec}$  are the reconstructed images from the U-Net passed to the generator. This formulation trains  $G$  to generate realistic images  $I_{gen}$  and  $D$  to distinguish real and synthetic samples. Here, Eq. (13) the generator output.

$$I_{gen} = G(I_{rec}, z) \quad (13)$$

Here,  $I_{rec}$  is the U-Net reconstructed image and  $z$  is the

latent noise vector used to generate outputs. The output  $I_{gen}$  is the adversarial refined Kannada character image that has both perceptual realism and restored stroke details. To ensure stability during training and to enhance convergence, we incorporate extra refinements including a gradient penalty as shown Eq. (14).

$$L_{gp} = \lambda (\|\nabla_I D(\hat{I})\|_2 - 1)^2 \quad (14)$$

Here,  $\hat{I}$  denotes an interpolated image from the real and generated samples,  $D(\hat{I})$  denotes the output from the discriminator,  $\nabla_I$  denotes the gradient with respect to  $I$  and  $\lambda$  denotes a weighting factor. This loss functions to impose the Lipschitz constraint so that the discriminator does not become too powerful, which would create irregularities in adversarial training. Eq. (15) denotes the feature matching loss, which matches the intermediate feature maps of the discriminator.

$$L_{fm} = \left\| \sum_l^L D^l(I) - D^l(I_{gen}) \right\|_2^2 \quad (15)$$

Here,  $D^{l(\cdot)}$  represents the feature maps from layer  $l$  of the discriminator,  $I$  is the real image, and  $I_{gen}$  is the generated image. This loss promotes the  $I_{gen}$  to match structural as well as perceptual similarity to the real character beyond pixel-wise differences. With these adversarial constraints imposed, the GAN-based refinement promotes the generator to create regenerated Kannada characters that are both visually realistic and structurally faithful in order to mitigate typically observed reconstruction issues such as broken strokes, awkward junctions, and overall style inconsistency. At this stage, the GAN etiquette complements the previously described U-Net-reconstruction by incorporating a discriminator-driven quality check that combines the initial coarse reconstruction to derive a regenerated handwritten character with high fidelity.

### 3.4 Loss function integration

To ensure that Kannada handwritten characters are accurately regenerated, our framework uses multiple loss components that instruct the generator to create visually coherent, structurally faithful, and identity-preserving reconstructions of the handwritten characters. These losses include pixel-level reconstruction loss, structural similarity loss, perceptual loss, and adversarial loss that together provide an appropriate balance between fidelity, stroke correctness, and realism.

The reconstruction loss  $L_{rec}$  encourages the generator to minimize the pixel-wise difference between original character  $I_{orig}$  and the regenerated output image  $I_{gen}$  as shown in Eq. (16).

$$L_{rec} = \|I_{orig} - I_{gen}\|_1 \quad (16)$$

In this equation,  $I_{orig}$  is the ground-truth Kannada character image and  $I_{gen}$  is the image reconstructed through the hybrid U-Net + GAN framework. L1 norm ensures that the regenerated image closely relates to the original strokes which promote the coherent reconstruction of fragmented characters. The L1-norm criterion diminishes blur artifacts while retaining clear details of fine strokes. To preserve the overall structure fidelity, we additionally use a structural similarity loss,

referred to as SSIM as given Eq. (17).

$$L_{ssim} = 1 - SSIM(I_{orig}, I_{gen}) \quad (17)$$

Here,  $I_{orig}$  denotes the ground-truth Kannada character image while  $I_{gen}$  denotes the generated image coming from the framework. We subtract the SSIM index from 1 to penalize structural differences in the loss allowing the different stroke shapes and characters geometric structures to be preserved upon regeneration. In order to capture higher-level semantic features as opposed to pixel values. We also define a perceptual loss  $L_{perc}$  over intermediate layers of a pretrained convolutional network VGG as Eq. (18).

$$L_{perc} = \sum_i^l \|\phi_i(I_{orig}) - \phi_i(I_{gen})\|_2^2 \quad (18)$$

Here,  $\phi_i(\cdot)$  denotes the feature map from the  $I^{th}$  layer of a pretrained CNN. The perceptual loss captures higher-level semantic divergences to enhance stroke realism. Finally, the adversarial loss  $L_{adv}$  complements these, which also encourages realism using the discriminator's feedback as shown in Eq. (19).

$$L_{adv} = -E[\log D(I_{gen})] \quad (19)$$

Here,  $D(\cdot)$  is the discriminator,  $I_{gen}$  is the generated image. The adversarial loss encourages realistic-looking handwriting textures with respect to discriminator feedback. In this synthesis, all of these components combine as follows into the combined generator loss given in Eq. (20).

$$L_G = \alpha L_{rec} + \beta L_{ssim} + \gamma L_{perc} + \delta L_{adv} \quad (20)$$

Here,  $\alpha, \beta, \gamma$ , and  $\delta$  are hyperparameters weighting reconstruction, structural, perceptual, and adversarial contributions, respectively.  $L_G$  is the overall generator loss used to train the U-Net + GAN generator. The discriminator will have a similar loss set to maximize its ability to distinguish real characters from generated characters.  $L_D$  is the discriminator loss defined as Eq. (21).

$$L_D = -(E[\log D(I_{orig})] + E[\log(1 - D(I_{gen}))]) \quad (21)$$

Here,  $I_{orig}$  represents the real image while  $I_{gen}$  is generated. LD maximizes the discriminator's ability to distinguish real from generated Kannada characters. By combining reconstruction, structural, perceptual, and adversarial objectives, this integrated loss is developed to maximize pixel accuracy of the regenerated Kannada characters, visual authenticity, and structural validity, in order to lay a sound optimization and validation framework.

### 3.5 Training objective with min-max optimization

The UNGAN-KR framework training goal is expressed as a min-max optimization problem, a classic characteristic of Generative Adversarial Networks (GANs), where the generator  $G$  and discriminator  $D$  are trained jointly to achieve realistic stroke reconstruction and good preservation of character identity. More specifically, the generator is trained to minimize its total loss  $L_G$ , so it generates visually believable

characters while providing structural consistency among all characters. Meanwhile, the discriminator is trained to maximize its assessment of the likelihood of identifying a character as either real or regenerated, which is captured using the loss  $L_D$  and represents the two objectives of the discriminator as Eq. (22).

$$MNF = (L_G + L_D) \quad (22)$$

In more detail, the generator loss  $L_G$  leverages pixel-level reconstruction loss, similarity in character structure, perceptual similarity in character form between the image and the reconstruction, and adversarial feedback from the discriminator  $L_D$  as given in Eq. (23).

$$\theta_G = \theta_G - \eta \cdot \frac{\widehat{m}_t}{\sqrt{\widehat{v}_t} + \epsilon} \quad (23)$$

The discriminator loss simply assesses how well  $D$  differentiates characters as either real or regenerated. To optimize these objectives, we employed gradient-based updates using the Adam optimizer as shown in Eq. (24).

$$\theta_D = \theta_D - \eta \cdot \frac{\widehat{m}_t}{\sqrt{\widehat{v}_t} + \epsilon} \quad (24)$$

where  $\theta_G$  and  $\theta_D$  are parameters of the generator and discriminator respectively,  $\eta$  is the learning rate, and  $\widehat{m}_t$ ,  $\widehat{v}_t$  are bias-corrected moment estimates. Learning rate decay and gradient clipping are employed to stabilize training as shown in Eq. (25).

$$\eta_t = \eta_0 \cdot \frac{1}{1 + \lambda t} \quad (25)$$

Here,  $\eta_0$  is the initial learning rate,  $\lambda$  is the decay factor, and  $t$  indicates the current training step. Eq. (26) applies a decay to the learning rate, which reduces  $\eta_t$  across time, and helps ensure stable convergence to the solution.

$$h_l^{drop} = h_l \cdot m, \quad m \sim \text{Bernoulli}(p) \quad (26)$$

Here,  $h_l$  is the activation of layer  $l$ , and mmm is a Bernoulli random mask with probability  $p$  that indicates dropout regularization randomly deactivates neurons to mitigate overfitting. Eq. (27) is referred to as the gradient clipping mechanism.

$$g_t = \frac{g_t}{\max(1, \|g_t\|_2 / c)} \quad (27)$$

Here,  $g_t$  is the gradient at step  $t$ , and ccc is the threshold for clipping. Within gradient clipping, the gradient is clipped to ensure that the updates remain bounded to prevent exploding gradients that can occur during training of the generator and discriminator. Lastly, similar to dropout regularization in the generator, dropout regularization mitigates common forms of overfitting and allows the generator to generalize well across the various handwritten Kannada characters. In addressing this through a min-max optimization problem, the generator continues to learn to generate high-fidelity stroke reconstructions that can fool the discriminator into thinking it had seen the original, while the discriminator is continually

learning to become better at recognizing real characters. Overall, this adversarial training occurs at a high frequency within a loop, ensuring that the recreated characters are visually representative and also similar to the original handwritten character, as the first stages of the character regeneration pipeline.

### 3.6 Validation and preservation of character identity

An important consideration in the regeneration of Kannada handwritten characters is the preservation of character identity such that the regenerated strokes respect the original character while providing better visual consistency. To support this, the framework incorporates identity-preserving constraints and validation metrics that direct the generator to produce structurally and semantically consistent outputs. First, apply a character classification constraint to the generator output  $y_{pred}$  derived from a pre-trained classifier using Eq. (28).

$$y_{pred} = \text{Softmax}(W_c f_{gen} + b_c) \quad (28)$$

Here,  $f_{gen}$  is the feature map of the generated character, and  $W_c$  and  $b_c$  are the classifier weights and bias. The resultant cross-entropy loss preserves the label of the generated character using Eq. (29).

$$L_{id} = - \sum_i^I y_i \log y_{pred,i} \quad (29)$$

Here,  $y_i$  corresponds to the ground-truth one-hot label for class  $i$ , and  $y_{pred,i}$  denotes the predicted probability for that class. This loss guarantees that the regenerated Kannada character retains its original class identity during training. To promote stroke-level consistency, a skeletonization operator, denoted  $S(\cdot)$ , is utilized, and a stroke consistency loss is defined as Eq. (30).

$$L_{stroke} = \| S(I_{orig}) - S(I_{gen}) \|_2^2 \quad (30)$$

Here,  $S(\cdot)$  indicates a stroke extraction function capturing the skeletal or structural representation of the Kannada character strokes. This loss guarantees that the regenerated image  $I_{gen}$  reproduces the original strokes from  $I_{orig}$ . This loss guarantees that the regenerated strokes follow the same trajectory as the original strokes, which is particularly important for complex Kannada ligatures and diacritics. Additionally, to prevent jagged and unnatural stroke connections, curve smoothness is imposed via Eq. (31).

$$L_{smooth} = \sum_t^T \|\Delta^2 x_t\|^2 + \|\Delta^2 y_t\|^2 \quad (31)$$

where,  $(x_t, y_t)$  represents a stroke coordinate at time step  $t$ , and  $\Delta^2$  computes the second-order differences along the stroke curve. Lastly, a style consistency loss defined with respect to Gram matrices  $G_j$  ensures that the overall visual style of the handwritten character is preserved using Eq. (32).

$$L_{style} = \sum_j^J \|G_j(I_{orig}) - G_j(I_{gen})\|_F^2 \quad (32)$$

Incorporating these criteria, the ultimate generator loss with respect to identity preservation is expressed as Eq. (33).

$$L_G^{final} = L_G + \lambda_{id} L_{id} + \lambda_{stroke} L_{stroke} \quad (33)$$

where,  $\lambda_{id}$  and  $\lambda_{stroke}$  are trade-offs that balance the importance of each identity preserving criterion. The training goal now shifts as Eq. (34).

$$F_{opt} = L_G^{final} + L_D \quad (34)$$

This provides assurance that the generator not only outwits the discriminator but also produces characters that remain correct, clear, and visually true to the original handwritten Kannada characters. Furthermore, aspects of the quality of regeneration are assessed using quantitative metrics of reconstruction accuracy, FID score, PSNR, and stroke-level precision and recall to ensure visual realism and semantic correctness. By combining these validation processes and identity preservation processes, the presented framework facilitates robust regeneration of broken handwritten Kannada characters while preserving all unique stroke patterns, ligature structures, and overall handwritten style, which is enormously important for subsequent applications such as OCR and historical document digitization.

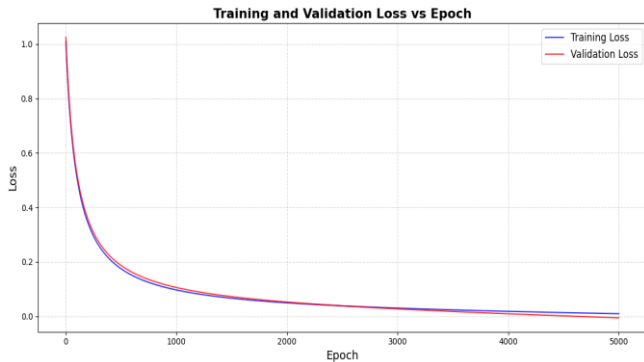
## 4. RESULTS AND DISCUSSIONS

The suggested framework for Kannada handwritten broken character regeneration through a hybrid U-Net and GAN-based mechanism was executed using the PyTorch 2.0 framework with CUDA support for rapid operation. This overall experiment was trialed on a workstation equipped with an NVIDIA RTX 3080 GPU (10 GB VRAM), and Intel i7 processor, and 32 GB RAM running Windows 11(64bits). In all experiments, the Adam optimizer was employed with an initial learning rate of  $1e-4$ ,  $\beta_1 = 0.5$  and  $\beta_2 = 0.999$ . These hyperparameters have also been established to work effectively for adversarial training since we trained continually for up to 200 epochs with a batch size of 64 in order to achieve convergence without overfitting. As mentioned in Section 4.3, augmentation techniques were employed i.e. random rotation ( $\pm 10^\circ$ ), horizontal flipping and Gaussian noise injection to improve model robustness to the variety of styles of handwriting and natural distortions according to the scanned documents.

Datasets were collected from 100 writers represented by a typical cross-section with varying ages, genders, and professions. The end data set contained 495 classes with 150 samples per class after the up-sampling process took place. The data set included base characters as well as types for Kagunitas of the Kannada script. The data set included some images that had broken character segments that could not be removed during the extraction process from the handwritten work. All images were visually checked manually if any images had a broken character, they were moved to a different bin. All characters used for training the model were visually clear and complete. The dataset was subdivided into training 70%, validation 15%, and testing 15% subsets while maintaining a fair class distribution in each split. Artificial noise, partial occlusion, and distortions were added to a subset of samples to approximate instances of broken characters in a controlled fashion. This way for the degree of degradation



evaluation, there will be a sample set that can be used to assess the parameters of the model's capability to regenerate potentially broken characters. The dataset has all key Kannada vowels/consonants and compound characters. Therefore, it is comprehensive to be representative of handwriting-based tasks addressing document recognition and resultant document digital archiving tasks.



**Figure 2.** Training and validation loss against epochs

Figure 2 illustrates the training and validation loss against epochs showcases the consistency of the proposed UNGAN-KR framework over 5000 training updates. Both the training and validation losses decrease consistently, suggesting that the model learns to reconstruct damaged Kannada characters from sketches while gradually generalizing to other samples. The distance between the training and validation loss stays small, demonstrating little overfitting and stable training behavior. The loss curves, after some number of thousands of epochs, also began to show an eventual plateau, presenting that the model has converged with smooth stroke reconstruction and perceptual quality. This data illustrates the effectiveness of the chosen architecture, hyperparameters, and adversarial refinement for stable and accurate regeneration of handwritten characters.

A U-Net model with four encoder-decoder levels was utilized with each level of the model consisting of convolutional layers implemented with batch normalization and the ReLU activation function. Skip connections were

utilized to maintain fine grain spatial information that would be necessary to accurately reconstruct stroke movements in recreating the characters. The GAN part of the model consisted of a PatchGAN discriminator that refined the output after stroke extraction, creating sharper, and more accurate characters. For metrics, both structural accuracy and visual realism were assessed for using Accuracy (Acc), Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Fréchet Inception Distance (FID), consisting of Eqs. (35)-(38), respectively.

$$Acc = N_{correct} / N_{total} \quad (35)$$

FID evaluates the distributional similarity between real and generated characters in a deep feature space. Lower FID values indicate that generated characters have a closer representation to real handwritten characters in terms of style and stroke distribution.

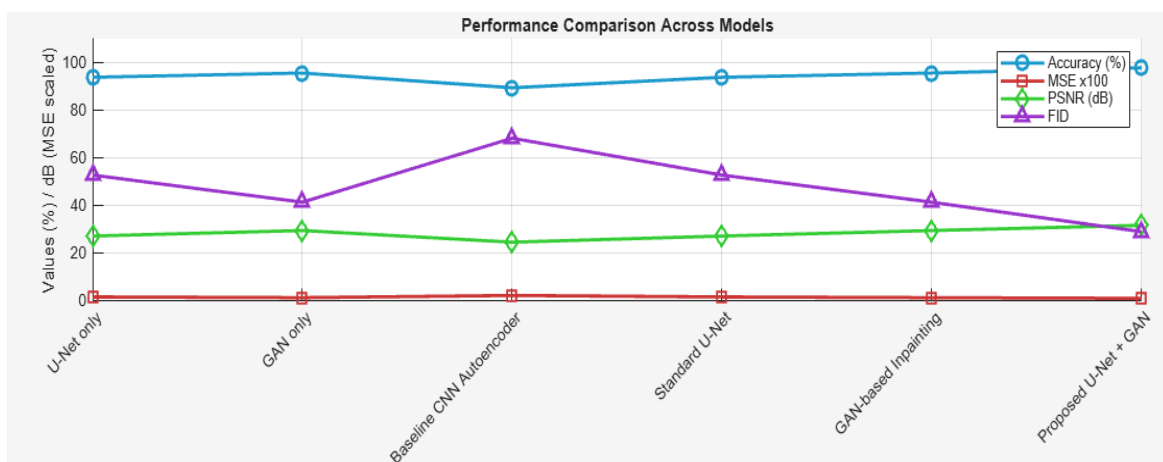
$$FID = \|\mu_r - \mu_g\|_2^2 + T_r \left( \Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{\frac{1}{2}} \right) \quad (36)$$

Peak Signal-to-Noise Ratio: PSNR summarizes the visual fidelity of regenerated character (i.e., replicated) as a function of signal strength to reconstruction error. Higher PSNR values indicate a more accurate regenerated character with less distortion.

$$P_{SNR} = 10 \cdot \frac{Max_f^2}{MSE} \quad (37)$$

MSE calculates the average of the squared difference of pixel values between original and regenerated characters. Lower MSE represents higher accuracy in reconstruction, which is important to maintain the integrity of the stroke of the character and the overall legibility of the character.

$$MSE = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W (I_{orig}(i,j) - I_{gen}(i,j))^2 \quad (38)$$



**Figure 3.** Result comparison of proposed model with few existing mechanisms

Table 1 and Figure 3 show that the proposed framework achieved a recognition accuracy of 97.8% that is 8.4% more than the baseline CNN autoencoder. This confirms the efficacy of the adversarial refined continuous U-Net encoder decoder

for fine-grained reconstruction of Kannada strokes. In general, the MSE dropped to 0.009 with our method suggesting a very high level of pixel restoration accuracy. Accordingly, the PSNR improved to 31.6 db indicating that downloadable

characters reconstructed from the model closely resemble original input characters with minimal noise in the signal. The FID of 28.9 suggests that the model reconstructed characters that do not just preserve structure but also appear perceptually similar to original handwritten characters. Compared to the standard U-Net of FID = 52.7, our framework reduces perceptual artifacts by nearly 45%.

**Table 1.** Result comparison of proposed UGAN-KR model with few existing mechanisms

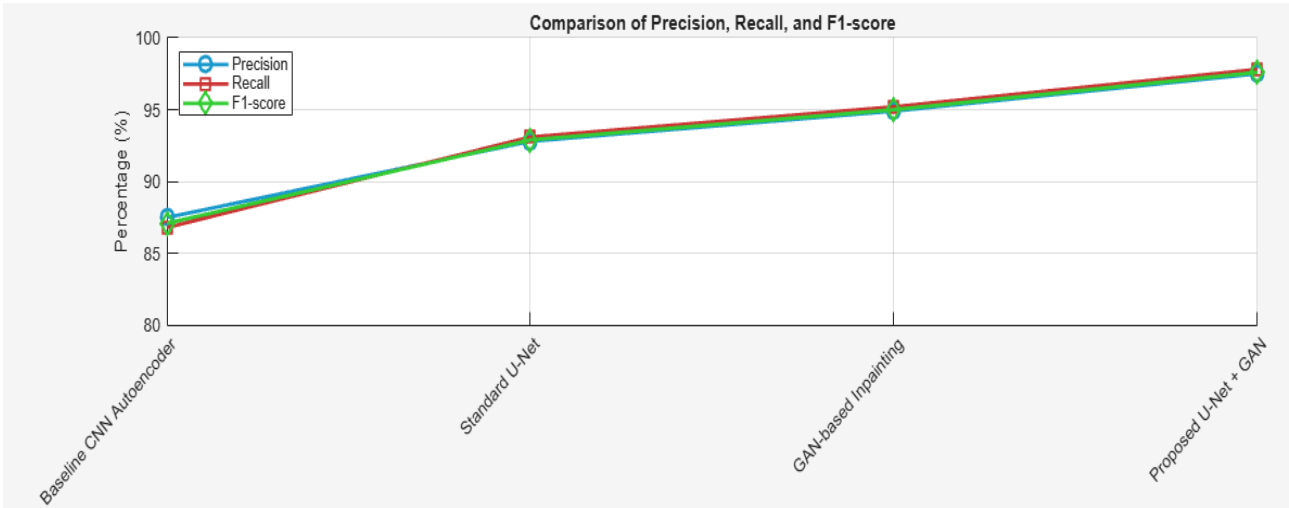
Model	Accuracy (%)	MSE	PSNR (dB)	FID
U-Net Only	93.8	0.015	27.1	52.7
Generative Adversarial Network (GAN) Only	95.6	0.012	29.4	41.3
Baseline				
Convolutional Neural Network (CNN)	89.4	0.021	24.5	68.2
Autoencoder				
Standard U-Net	93.8	0.015	27.1	52.7
GAN-based Inpainting	95.6	0.012	29.4	41.3
Proposed U-Net + GAN	97.8	0.009	31.6	28.9

The proposed UGAN-KR model has a 97.5% precision meaning that nearly every regenerated character belonging to the same class is correct. It is important to mitigate

misclassification due to visual similarity in Kannada characters. The recall score of 97.8% reflects the model's ability to recover nearly all instances of broken characters across various handwriting styles and minimizes the possibility of missed reconstructions. The F1 score of 97.6% captures the balance between precision and recall across instances of broken characters and signifies overall strong performance of character regeneration. When compared with baseline autoencoders and discrete U-Net/GAN models, encoder-decoder reconstruction paired with adversarial refinement yields a marked improvement in structural restoration and recognition reliability regarding the various instances of Kannada characters.

**Table 2.** Precision, recall, and F1 score comparison

Model	Precision (%)	Recall (%)	F1-score (%)
Baseline			
Convolutional Neural Network (CNN)	87.5	86.8	87.1
Autoencoder			
Standard U-Net	92.8	93.1	92.9
GAN-based Inpainting	94.9	95.2	95.0
Proposed U-Net + Generative Adversarial Network (GAN)	97.5	97.8	97.6



**Figure 4.** Precision, recall, and F1 score comparison

**Table 3.** Ablation study result analysis

Configuration	Acc (%)	Precision (%)	Recall (%)	F1-score (%)	MSE	PSNR (dB)	FID
U-Net Only	93.8	92.8	93.1	92.9	0.015	27.1	52.7
Generative Adversarial Network (GAN) Only	95.6	94.9	95.2	95.0	0.012	29.4	41.3
UGAN-KR (Proposed)	97.8	97.5	97.8	97.6	0.009	31.6	28.9

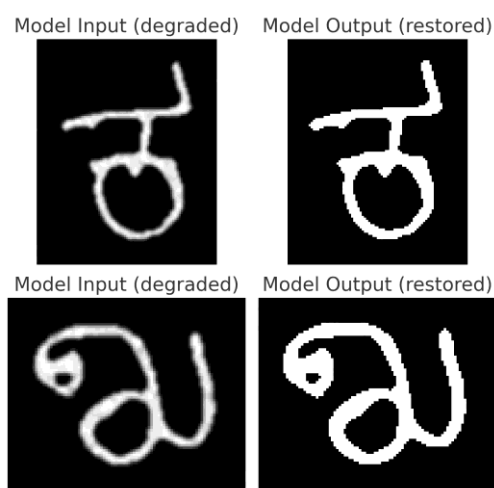
The model demonstrates a high level of accuracy as shown in Table 2 and Figure 4. Specifically, it shows the precision of 97.5% indicates that the model can identify most of the broken characters that correspond with that particular class of characters, which is crucial in the assessment given that Kannada characters sometimes visually appear alike and can be misclassified. The 97.8% recall metric suggests that the

proposed model is able to reconstruct almost all of the broken characters across different styles of handwriting with minimal missed reconstructions. To study the importance and contribution of the different components of the proposed framework. Table 3 shows the ablation studies were conducted using the Kannada handwritten broken character dataset looking at three variations: (1) U-Net only, (2) GAN only, and

(3) the proposed combination of U-Net + GAN. Each of the variations was tested on several metrics: Accuracy (Acc), Precision (P), Recall, F1-score, Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Fréchet Inception Distance (FID) for a more holistic sense of structural and perceptual performance. The U-Net only variation was able to reconstruct almost all individual broken strokes because of the encoder-decoder structure of the U-Net.

The use of skip connections in the U-Net model retained all of the spatial information by passing information from the encoder to decoder. And almost all of the broken strokes were reconstructed correctly. The quality of reconstruction in U-Net was due the overall process encoder-decoder architecture for capturing textural details. But, the lack of adversarial feedback resulted in fine structural and textural details not being accurately captured resulting in blurry outputs and an FID value that was higher (52.7) than the other two models. Thus, U-Net only demonstrated strong pixel reconstruction, it lacked perceptual realism, especially for complicated compound characters and ligatures. The GAN-only configuration was purely focused on adversarial refinement while the U-Net based approach included no structural guidance through skip connections. Although there was some variation in accuracy metrics and performance, relatively lower FID of 41.3 and higher PSNR of 29.4 dB scores suggest generally sharper. However, without structural guidance some minor placement differences remained for compound characters and complicated ligatures resulting in a standard accuracy of 95.6% compared to U-Net based defects.

The UNGAN-KR methodology integrated both methods, resulting in the overall best evaluation metrics. As well as guaranteeing structural fidelity in reconstructing missing strokes, U-Net also determined the GAN discriminator ensured plausible handwriting texture while reducing data loss and reducing potentially harmful perceptual loss artifacts. The combined models resulted in 97.8% accuracy score in addition to the lowest-MSE (0.009), highest-PSNR (31.6 dB), and lowest-FID (28.9) showing alignment in reducing pixel fidelity while maintaining higher and better perceptual quality. Finally, per-character metrics include Precision of 97.5%, Recall of 97.8%, and F1 score of 97.6%; the model robustly restored broken characters with no mistakes in classification. Figure 5 shows an example output of two characters after restoration using the proposed model.



**Figure 5.** Example output of two characters after restoration using proposed model

The unit level experiments verify that U-Net or GAN alone are not sufficient to achieve the best hypothesis when regenerating Kannada characters. The U-Net + GAN combination is important to substantiate both structural accuracy and perceptual realism simultaneously, while providing empirical evidence that the proposed solution is novel and effective solution for restoring broken handwritten Kannada characters and can potentially be used for applications in OCR pre-processing, digital archiving of manuscripts, and automatic handwriting analysis.

## 5. CONCLUSION

This paper has introduced a U-Net + GAN hybrid model to regenerate broken Kannada handwritten characters. It is an important problem in manuscript preservation as well as for handwriting capture systems. The proposed UNGAN-KR model adopts stroke reconstruction based on U-Net encoder-decoder architecture together with adversarial refinement to restore missing or degraded strokes while maintaining the original handwriting style. Experimental results using a dataset of 71,149 Kannada handwritten character images demonstrate that the proposed framework consistently outperforms. A set of base line approaches including CNN autoencoders, U-Net, and GAN-based inpainting methods. The framework proposed achieved a recognition accuracy of 97.8%, which is 8.4% more than the baseline mechanisms. This confirms the efficacy of the adversarial refined continuous U-Net encoder-decoder for fine-grained reconstruction of Kannada strokes. In the future, it can be extended the proposed framework to multi-script handwritten datasets to test its generalizability across multiple Indian scripts. The inclusion of transformer-based attention models and semi-supervised training may improve stroke reconstruction by reducing reliance on large annotated datasets. Additionally, employing diffusion-based generative models may improve the realism and perceptual quality of the character's regeneration, especially for more poorly degraded or visually stylized handwriting.

## ACKNOWLEDGEMENT

The authors would like to express my sincere gratitude to Aditya Rao, Ashay Naik, Ashrit Bharadwaj, and Ajey Bhat for their valuable support during the coding phase of the research work. Their dedication and willingness to collaborate played a crucial role in the successful development and completion of the work.

## REFERENCES

- [1] Nigam, S., Behera, A.P., Verma, S. Nagabhushan, P. (2024). Deformity removal from handwritten text documents using variable cycle GAN. *International Journal on Document Analysis and Recognition (IJDAR)*, 27: 615.627. <https://doi.org/10.1007/s10032-024-00466-x>
- [2] Alaasam, R., Madi, B., El-Sana, J. (2024). Text enhancement for historical handwritten documents. In *Document Analysis and Recognition* -

- ICDAR 2024. ICDAR 2024. Lecture Notes in Computer Science, pp. 397-412. [https://doi.org/10.1007/978-3-031-70536-6\\_24](https://doi.org/10.1007/978-3-031-70536-6_24)
- [3] Rabhi, B., Elbaati, A., Boubaker, H. Pal, U., Alimi, A.M. (2024). Multi-lingual handwriting recovery framework based on convolutional denoising autoencoder with attention model. *Multimedia Tools and Applications*, 83: 22295-22326. <https://doi.org/10.1007/s11042-023-16499-z>
- [4] Gongidi, S., Jawahar, C.V. (2021). IIIT-INDIC-HW- Words: A dataset for Indic handwritten text recognition. In *Document Analysis and Recognition – ICDAR 2021*, pp. 444-459. [https://doi.org/10.1007/978-3-030-86337-1\\_30](https://doi.org/10.1007/978-3-030-86337-1_30)
- [5] Ramesh, G., Shreyas, J., Balaji, J.M., Sharma, G.N., Gururaj, H.L., Srinidhi, N.N., Askar, S.S., Abouhawwash, M. (2024). Hybrid manifold smoothing and label propagation technique for Kannada handwritten character recognition. *Frontiers in Neuroscience*, 18: 1362567. <https://doi.org/10.3389/fnins.2024.1362567>
- [6] Siddanna, S.R., Kiran, Y.C. (2024). An efficient recognition of handwritten Kannada script using variable attention-based Coati integrated Bi directional Long Short-Term Memory. *Multimedia Tools and Applications*, 83: 88981-89002. <https://doi.org/10.1007/s11042-024-18999-y>
- [7] Liu, X.Y., Meng, G.F., Xiang, S.M., Pan, C.H. (2021). Handwritten text generation via disentangled representations. *IEEE Signal Processing Letters*, 28: 1838-1842. <https://doi.org/10.1109/LSP.2021.3109541>
- [8] Chen, Z.N., Yang, D.H., Liang, J.L., Liu, X.W., Wang, Y.Y., Peng, Z.H., Huang, S.P. (2022). Complex handwriting trajectory recovery: Evaluation metrics and algorithm. In *16th Asian Conference on Computer Vision*, Macao, China, pp. 58-74. [https://doi.org/10.1007/978-3-031-26284-5\\_4](https://doi.org/10.1007/978-3-031-26284-5_4)
- [9] Elanwar, R., Betke, M. (2025). Generative adversarial networks for handwriting image generation: A review. *The Visual Computer*, 41: 2299-2322. <https://doi.org/10.1007/s00371-024-03534-9>
- [10] Zdenek, J., Nakayama, H. (2021). JokerGAN: Memory-efficient model for handwritten text generation with text line awareness. In *Proceedings of the 29th ACM International Conference on Multimedia*, Virtual Event China, pp. 5655-5663. <https://doi.org/10.1145/3474085.347571>
- [11] Dong, C.C., Liu, H.M., Wang, X.Y., Bi, X.H. (2024). Image inpainting method based on AU-GAN. *Multimedia Systems*, 30: 101. <https://doi.org/10.1007/s00530-024-01290-3>
- [12] Elharrouss, O., Damseh, R., Belkacem, A.N., Badidi, E., Lakas, A. (2025). Transformer-based image and video inpainting: Current challenges and future directions. *Artificial Intelligence Review*, 58: 124. <https://doi.org/10.1007/s10462-024-11075-9>
- [13] Chen, Y.T., Xia, R.L., Yang, K., Zou, K. (2024). DNNAM: Image inpainting algorithm via deep neural networks and attention mechanism. *Applied Soft Computing*, 154: 111392. <https://doi.org/10.1016/j.asoc.2024.111392>
- [14] Li, H.L., Zhang, Y.F., Wang, W.M., Zhang, S.Y., Zhang, S.X. (2024). Recovery-based occluded face recognition by identity-guided inpainting. *Sensors*, 24(2): 394. <https://doi.org/10.3390/s24020394>
- [15] Yeh, C.H., Yang, H.F., Chen, M.J., Kang, L.W. (2024). Image inpainting based on GAN-driven structure- and texture-aware learning with application to object removal. *Applied Soft Computing*, 161: 111748. <https://doi.org/10.1016/j.asoc.2024.111748>
- [16] Chang, C.C., Garcia Perera, L.P., Khudanpur, S. (2023). Crosslingual handwritten text generation using GANs. In *Document Analysis and Recognition – ICDAR 2023 Workshops*, pp. 285-301. [https://doi.org/10.1007/978-3-031-41501-2\\_20](https://doi.org/10.1007/978-3-031-41501-2_20)
- [17] Dubey, P., Nayak, M., Gehani, H., Kukade, A., Keswani, V., Dubey, P. (2025). Enhancing realism in handwritten text images with generative adversarial networks. *Bulletin of Electrical Engineering and Informatics*, 14(3): 2370-2379. <https://doi.org/10.11591/eei.v14i3.9190>
- [18] Hanif, S., Latecki, L.J. (2023). Strokes trajectory recovery for unconstrained handwritten documents with automatic evaluation. In *Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2023)*, Lisbon, Portugal, pp. 661-671. <https://doi.org/10.5220/0011700100003411>
- [19] Alaei, A., Nagabhushan, P., Pal, U. (2011). A benchmark kannada handwritten document dataset and its segmentation. In *2011 International Conference on Document Analysis and Recognition*, Beijing, China, pp. 141-145. <https://doi.org/10.1109/ICDAR.2011.37>
- [20] Sajjan, S.P., Bannigidad, P., Domlur, A. (2025). Ancient kannada handwritten palm leaf dataset. *Mendeley Data*. <https://doi.org/10.17632/w5px7czbn9.3>
- [21] Kong, Y.X., Luo, C.J., Ma, W.H., Zhu, Q.Y., Zhu, S.G., Yuan, N. (2022). Look closer to supervise better: One-shot font generation via component-based discriminator. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, pp. 13472-13481. <https://doi.org/10.1109/CVPR52688.2022.01312>
- [22] Bannigidad, P., Sajjan, S.P. (2023). Restoration of ancient kannada handwritten palm leaf manuscripts using image enhancement techniques. In *5th EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing*. BDCC 2022. EAI/Springer Innovations in Communication and Computing, pp. 101-109. [https://doi.org/10.1007/978-3-031-28324-6\\_9](https://doi.org/10.1007/978-3-031-28324-6_9)
- [23] Parashivamurthy, S.P.T., Rajashekararadhya, S.V. (2024). An efficient kannada handwritten character recognition framework with serial dilated cascade network for kannada scripts. *Advances in Artificial Intelligence and Machine Learning*, 4(3): 2499-2516. <https://doi.org/10.54364/AAIML.2024.43146>
- [24] Wang, Z., Li, Y.J., Li, H.L. (2025). Chinese inscription restoration based on artificial intelligent models. *npj Heritage Science*, 13: 326. <https://doi.org/10.1038/s40494-025-01900-x>
- [25] Sareen, B., Ahuja, R., Singh, A. (2024). CNN-based data augmentation for handwritten gurmukhi text recognition. *Multimedia Tools and Applications*, 83: 71035-71053. <https://doi.org/10.1007/s11042-024-18278-w>
- [26] Wang, Y.M., Wang, H., Sun, S.W., Wei, H.X. (2022). An approach based on transformer and deformable convolution for realistic handwriting samples generation.

- In 2022 26th International Conference on Pattern Recognition (ICPR), Montreal, QC, Canada, pp. 1457-1463. <https://doi.org/10.1109/ICPR56361.2022.9956551>
- [27] Hameed, M.M., Ahmad, R., Kiah, L.M., Murtaza, G., Mazhar, N. (2023). OffSig-SinGAN: A deep learning-based image augmentation model for offline signature verification. *Computers, Materials and Continua*, 76(1): 1267-1289. <https://doi.org/10.32604/cmc.2023.035063>
- [28] Hong, D.J., Chang, W.D., Cha, E.Y. (2024). Handwritten signature generation using denoising diffusion probabilistic models with auxiliary classification processes. *Applied Sciences*, 14(22): 10233. <https://doi.org/10.3390/app142210233>
- [29] Liao, Q.S., Xia, G., Wang, Z.N. (2023). Calliffusion: Chinese calligraphy generation and style transfer with diffusion modeling. *arXiv preprint arXiv:2305.19124*. <https://doi.org/10.48550/arXiv.2305.19124>
- [30] Pippi, V., Cascianelli, S., Cucchiara, R. (2023). Handwritten text generation from visual archetypes. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, pp. 22458-22467. <https://doi.org/10.1109/CVPR52729.2023.02151>