GANs for Cultural Legacy: Colorizing Indonesia's Historical Images

Zinadine Zidan Alsyahana[#1], Jeffrey Nobel Martin[#2],

Affa Ndaru Rabbany Wijaya[#3], Raden Fauzan Wisnuwardhana[#4]

[#]*Robotics and AI Engineering,*
Faculty of Advanced Technology and Multidiscipline,
Airlangga University
Jl. Dr. Ir. H. Soekarno, Mulyorejo, Kec. Mulyorejo, Kota SBY, Jawa Timur 60115

[1] `zinadine.zidan.alsyahana-2022@ftmm.unair.ac.id`

[2] `jeffrey.nobel.martin-2022@ftmm.unair.ac.id`

[3] `affa.ndaru.rabbany-2022@ftmm.unair.ac.id`

[3] `raden.fauzan.wisnuwardhana-2022@ftmm.unair.ac.id`

Abstract — This study applies Generative Adversarial Networks (GANs) to colorize historical Indonesian photographs. The results demonstrate the model's ability to produce realistic colorizations, accurately replicating details such as skin tones, skies, and environmental elements, while eliminating the need for manual adjustments. Despite minor inconsistencies in certain areas, the approach offers significant potential for cultural preservation, supporting government programs for archiving and restoration, with possible extensions to video reconstruction. Future work aims to improve accuracy by incorporating cultural and contextual nuances and expanding the dataset to ensure greater diversity and fidelity.

***Keywords* — Generative Adversarial Networks (GANs); Image Colorization; Historical Photographs; Cultural Preservation; Automated Restoration; Digital Heritage; Convolutional Neural Networks (CNN).**

I. Introduction

The foundation of a nation is inseparable from the long history of the nation itself. Starting from the political, economic, social and cultural aspects. History itself comes from the Greek ἱστορία, historia, which means "investigation, knowledge gained through investigation". History itself is a general term that includes the events of the past as well as the memory, discovery, collection, organization, presentation, and interpretation of these events [1]. While according to Isidore, the bishop of seville answered the question "what is history". According to him, history is "a narrative of events through which we learn about what happened in the past" [2].

Saffo once wrote in one of his articles that "To see the future, we must look back twice as far", this reinforces the importance of making history a reference in preparing for a better future. Understanding history is not just remembering a series of past events, but also analyzing how various events are interconnected and shape the current reality. This understanding allows us to recognize patterns that may repeat themselves, but in a different context in the future. History also teaches us that change does not happen in a simple linear fashion, but through a complex web of cause and effect. By studying history, we can develop a more critical and holistic way of thinking in facing the challenges of the present and preparing for the future. Moreover, it helps us avoid rigid determinism and encourages more open-minded thinking about the possibilities of the future [3].

A nation shapes its social identity through shared memories that are passed down from generation to generation. These memories are stored in various forms of cultural media, including museums, monuments, textbooks, movies, and especially photographs [4]. During the Indonesian Revolution, black and white photographs played a vital role in shaping public opinion. Newspapers such as Kedaulatan Rakjat used these photos to build the spirit of nationalism by featuring the nation's leaders, documenting the people's struggle, and exposing the atrocities of the colonizers [5]. More than just documentation, these black and white photographs are silent witnesses to the journey of the Indonesian nation that reflect a unique blend of global influences and local wisdom, while illustrating the diversity that characterizes Indonesian society.

With the development of technology in this day and age, the coloring of historical images is no longer something that is difficult to do. Researchers have developed a variety of methods that could only be done manually, using art tools and digital software, such as Adobe Photoshop, to modern artificial intelligence-based approaches. The integration of artificial intelligence and deep learning algorithms has revolutionized the field, offering time-saving and highly accurate automated coloring solutions [6]. While Convolutional Neural Networks (CNNs) have been the conventional choice for image processing tasks, Generative Adversarial Networks (GANs) are emerging as a superior alternative [7]. GANs have shown remarkable capabilities in enhancing the quality and authenticity of color images, surpassing the performance of other neural network models.

The coloring of historical images has a huge impact not only for historians, but also has the potential to assist archivists and anthropologists in deciphering new meanings from archival collections. By coloring these artifacts, historians and preservationists can make them more accessible and appealing to a wider audience, fostering a deeper connection to our collective history [8]. Colorization technology using GANs can bridge the generation gap in understanding history, where younger generations who are familiar with colored visual content can better engage with the nation's historical narrative [9].

II. Literature Review

A. *Generative Adversarial Networks (GANs)*

First introduced by Goodfellow et al. [10], Generative Adversarial Networks (GANs) are made up of two neural networks—a discriminator $D(x)$ and a generator $G(z)$—that are trained concurrently using adversarial learning. While the discriminator learns to differentiate between generated and real samples, the generator learns to map a latent space vector z to the data distribution. GANs mathematically resolve a minimax game that is expressed as follows:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z)))] \quad (1)$$

In this framework, the discriminator D maximizes its capacity to discriminate between generated and real images, while the generator G attempts to minimize the negative log-likelihood of the discriminator successfully differentiating between real images x and generated images $G(z)$.

1. Conditional GANs (cGANs)

Conditional GANs (cGANs) [11] are an extension of the conventional GAN framework for image colorization. They do this by conditioning the discriminator and generator on extra information, like grayscale images. For cGANs, the objective function is written as follows:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[\log D(x|y)] + \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z|y)))] \quad (2)$$

The grayscale image is denoted by y in this case, and the generator's job is to figure out how to map the grayscale image to the colorized image, or $G: y \rightarrow x_{color}$.

B. Image Colorization Techniques

1. Traditional Methods

Image colorization mainly relied on heuristic-based algorithms or manual techniques prior to the development of deep learning [12]. Earlier methods used transfer learning techniques, like coloring an image using pre-established color models and known correspondences between grayscale intensity. Mapping functions could be used to mathematically express these techniques. $f: I_{gray} \rightarrow I_{color}$, frequently based on the grayscale image's gradient or texture information.

2. CNN-Based Approaches

Because convolutional neural networks (CNNs) can capture hierarchical spatial features, they have shown promise in image colorization tasks [13]. Grayscale images are frequently fed into CNN-based colorization models, which forecast a color distribution for every pixel [14]. For instance, a CNN is trained to predict a color image I_{color} given a grayscale image I_{gray} in such a way that:

$$I_{color} = \text{CNN}(I_{gray}) \quad (3)$$

CNNs frequently predict the color space in the CIE Lab color space, which separates luminance (L) from chrominance (a, b) channels, in order to model this task probabilistically. This makes colorization possible without changing the input grayscale intensities [15].

3. GAN-Based Colorization

By incorporating adversarial training to enhance the output, GAN-based colorization expands upon CNN methods. By employing a discriminator to assess the authenticity of the generated images, GANs generate more realistic colors while maintaining texture, structure, and context throughout the colorization process. While the discriminator D learns to differentiate between generated and real color images, the generator G learns a mapping from a grayscale input y to its corresponding color image x_{color}:

$$\mathcal{L}_{GAN}(G,D) = \mathbb{E}_{x_{color}}[\log D(x_{color})] + \mathbb{E}_y[\log(1-D(G(y)))] \quad (4)$$

To improve stability, an additional $L1$ or $L2$ loss is often incorporated:

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x_{color},y}[||x_{color} - G(y)||_1] \quad (5)$$

The combined loss function is expressed as:

$$\mathcal{L}_{total}(G,D) = \mathcal{L}_{GAN}(G,D) + \lambda\mathcal{L}_{L1}(G) \quad (6)$$

where λ is a weight balancing the adversarial loss and the reconstruction loss [16].

C. Advanced GAN Architectures for Image Colorization

1. Pix2Pix GAN

One well-known conditional GAN model for image-to-image translation tasks is the Pix2Pix GAN, developed by Isola et al. [17]. When it comes to image colorization, Pix2Pix allows direct supervision for mapping grayscale images to their colorized counterparts by conditioning the discriminator and generator on the grayscale input image. The standard GAN loss plus a reconstruction term derived from the $L1$ loss make up the loss function for Pix2Pix:

$$\mathcal{L}_{Pix2Pix} = \mathcal{L}_{GAN}(G,D) + \lambda\mathbb{E}_{x,y}[||x - G(y)||_1] \quad (7)$$

When working with culturally significant imagery, where artifact preservation is vital, this enables the model to generate high-quality images while preserving fine details in historical images.

2. CycleGAN

Another development in GANs, CycleGAN [18] is made especially for unpaired image-to-image translation tasks. CycleGAN can be modified for colorization tasks where paired data is limited, as may be the case with historical datasets, even though this is more frequently utilized in style transfer. The architecture uses two discriminators, D_X and D_Y, and two generators, G and F, to ensure cycle consistency through the loss:

$$\mathcal{L}_{cycle}(G, F) = \mathbb{E}_x[||F(G(x)) - x||_1] + \mathbb{E}_y[||G(F(y)) - y||_1] \tag{8}$$

In order to maintain historical accuracy, cycle-consistent loss makes sure that images converted back from the colorized domain to grayscale stay consistent with the original grayscale images.

D. Related Work

Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) are two examples of contemporary deep learning techniques that have advanced image colorization, an active research area. Transforming grayscale images into realistic color representations is the aim of colorization, and a number of methods have been developed to meet this challenge. In this section, we review important contributions to the image colorization problem, with a focus on GAN-based methods [19].

1. Traditional Colorization Approaches

Within the past, picture colorization was ordinarily done physically or semi-automatically, some time recently profound learning. Levin et al. [20] prescribed a novel approach based on user-placed color writes. These jots were dispersed to adjacent pixels employing a dissemination preparation based on luminance likeness. This approach performed well for little pictures, but since the client had to physically give color clues, it battled to handle expansive picture datasets or complex surfaces.

Cheng et al. [21] displayed a calculation that minimized a smoothness imperative within the color space whereas consequently anticipating the color of grayscale pictures employed a learned mapping between grayscale force and color disseminations. Their approach, be that as it may, was confined to characteristic settings and did not work well with pictures from history or culture that had unmistakable color plans.

2. GAN and Deep Learning-Based Approaches

The development of Convolutional Neural Networks (CNNs) in computer vision revolutionized image colorization by automating the process through the use of large datasets [22]. Zhang et al. [22] presented a CNN-based method for predicting the ab channels in the Lab color space from grayscale images (where L stands for lightness and a and b for color information). Due to the probabilistic nature of the predicted colors, each pixel may have multiple colors. To introduce a classification framework, they assigned each pixel to one of multiple bins representing a color distribution:

$$\hat{p}(x, y) = \arg\max_c P(c|I_{gray}(x, y)) \tag{9}$$

This probabilistic method increased color diversity while requiring less human intervention. CNN-based models, however, had trouble achieving photorealistic outcomes, leading to colorizations that were inconsistent or desaturated.

Colorization was one of the image-to-image translation tasks that Isola et al. [18] used conditional GANs (cGANs) for in their Pix2Pix framework. Pix2Pix combines a CNN encoder-decoder structure with skip connections with a U-Net generator to improve the preservation of fine details in generated images. In order to ensure that the generator learned to produce more realistic and coherent colorizations, the adversarial training process included a discriminator that assessed the generated color images' realism. The loss function encouraged realistic yet accurate colorization by fusing adversarial loss with L1 reconstruction loss.

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x_{color}}[\log D(x_{color})] + \mathbb{E}_y[\log(1 - D(G(y)))] + \lambda \mathbb{E}_{x,y}[||x - G(y)||_1] \tag{10}$$

Particularly on datasets containing paired grayscale and color images, Pix2Pix produced state-of-the-art outcomes in supervised image colorization.

Pix2Pix relied on paired training data, while CycleGAN handled unpaired image-to-image translation [18]. CycleGAN was effective when historical grayscale images lacked corresponding color images. Cycle-consistent losses were used to ensure that images created by converting from grayscale to color and back again were consistent.

$$\mathcal{L}_{cycle}(G, F) = \mathbb{E}_x[||F(G(x)) - x||_1] + \mathbb{E}_y[||G(F(y)) - y||_1] \tag{11}$$

CycleGAN is a helpful tool for recovering old images because it offers flexible image colorization without requiring large paired datasets. In contrast to supervised models like Pix2Pix, the results of this unsupervised approach were less reliable.

Radford et al. [24] proposed the Deep Convolutional GAN (DCGAN) architecture, which incorporates convolutional layers into both the discriminator and the generator to improve training reliability and efficiency. Several projects have used DCGAN for image colorization. A two-stream architecture was employed by Iizuka et al. [25] to guarantee coherence in large structures. One stream used convolutional layers to capture local image features, while the other stream captured global image context. Especially in scenes with intricate textures or objects, this dual-stream architecture enhanced the consistency of color prediction.

GAN-based colorization models have also been enhanced with attention mechanisms, which enable the network to concentrate on particular regions of the image during the colorization process. Tang et al. [26] proposed an attention-guided GAN that uses spatial attention to colorize important regions of historical images more accurately and consistently, such as faces or culturally significant elements. Particular architectural details or traditional attire in old photos required special attention, so this was very helpful for applications like cultural heritage preservation.

$$\mathcal{L}_{attn} = \sum_{i=1}^{N} \alpha_i \mathcal{L}_{GAN}^{i} \tag{12}$$

where N represents each image region's attention weights and the total number of image regions.

3. Applications to Historical Image Restoration

Due to the increasing interest in digital cultural heritage preservation, the application of GAN-based colorization for historical images has garnered a lot of attention despite being relatively new. A recent study on ancient Chinese painting restoration introduces a GAN-based approach called the UGAN model, which utilizes a U-shaped network (UNet) for the generator and incorporates a dilated convolution-gated residual block (DCGR-Block). This configuration enables the model to effectively capture both shallow and deep features, significantly enhancing the restoration quality compared to other mainstream inpainting methods [27].

The Doc-Attentive-GAN [28] model is designed to improve historical document denoising by focusing on regions with high levels of degradation. By utilizing an attention map, the model's generator learns to prioritize and correct noise-affected areas, enhancing the quality of restored text. The model's architecture includes an attentive-recurrent network and a contextual autoencoder, which work together to achieve sharp and clean outputs. Through perceptual loss, expressed as

$$L_P(O,T) = \mathrm{L}_{\mathrm{MSE}}(\mathrm{VGG}(O), \mathrm{VGG}(T)) \tag{13}$$

where O is the output image and T is the ground-truth clean image. The network evaluates discrepancies between generated images and ground-truth clean images, ensuring visual fidelity. Doc-Attentive-GAN has shown notable improvements across various historical datasets, contributing significantly to the field of document image restoration.

These studies conclude by demonstrating the effectiveness of GAN-based image colorization techniques, particularly in the restoration of historical and cultural heritage images. However, models that can incorporate cultural and contextual information into the colorization process are still required. This is especially important when working with historical Indonesian photographs, where colorization should enhance aesthetics while maintaining cultural authenticity. Our work aims to build on these developments and preserve Indonesia's rich legacy by developing a GAN-based model that not only colorizes old photos but also takes cultural differences into account.

III. Methods

A. Data Collection

For this study, data was collected through web scraping using specific keywords, including "Soekarno," "Soeharto," and "G30S-PKI." This approach allowed for the retrieval of a diverse set of historical images relevant to the subject matter. After collecting the initial dataset, We applied data augmentation techniques to enhance the quantity and variability of the images. These techniques included rotations, flips, and adjust color brightness, which helped create a more robust dataset for training the Generative Adversarial Networks (GANs) model. This enriched dataset ultimately contributed to improving the model's performance in generating realistic colorized images.

B. Loading and Processing Grayscale Image

The grayscale input image I_{gray}, representing a historic photograph, must first be preprocessed for optimal input to the GANs. This step involves standardizing the image dimensions (e.g., resizing to 256×256) and normalizing pixel values to a range suitable for the model (e.g., [0,1]). Mathematically, the preprocessing function Preprocess}(I_{gray}) is defined as:

$$I_{gray}^{norm} = \frac{I_{gray}}{255} \tag{14}$$

where I_{gray}^{norm} is the normalized grayscale image used as input for the generator.

C. Defining the Generator and Discriminator Models

In a GANs framework, the generator and discriminator models work together in an adversarial manner to produce realistic color images from grayscale input.

1. Generator Model

The generator G is designed to map the grayscale input I_{gray} to a colorized output I_{color}. The architecture typically consists of a sequence of convolutional and transposed convolutional (deconvolutional) layers, which refine spatial details and color information progressively. The forward pass through the generator can be formalized as:

$$I_{color} = G(I_{gray}^{norm}; \theta_G) \tag{15}$$

where θ_G represents the trainable weights of the generator. The architecture is optimized to minimize the perceptual and adversarial loss.

2. Discrimination Model

The discriminator D is a binary classifier tasked with distinguishing real color images I_{real} from the generator's synthetic outputs I_{color}. Structurally, D is a convolutional network designed to assess whether an image is authentic or generated. Its function can be expressed as:

$$y = D(I; \theta_D) \tag{16}$$

where $y \in [0,1]$ is the discriminator's prediction (1 for real, 0 for fake), I is either I_{real} or I_{color}, and θ_D are the discriminator's parameters.

D. Compiling and Training the Models

1. Compile the Discriminator

The discriminator D is compiled with a binary cross-entropy loss function, defined as:

$$\mathcal{L}_D = -\mathbb{E}_{I_{real} \sim p_{data}}[\log D(I_{real})] \tag{17}$$
$$-\mathbb{E}_{I_{color} \sim G(I_{gray}^{norm})}[\log(1 - D(I_{color}))]$$

where p_{data} denotes the distribution of real images. The Adam optimizer, often with learning rates such as $\alpha = 0.0002$ and momentum parameters $beta_1 = 0.5$, is applied for stable training.

2. Construct the Model for GAN

The GAN model combines G and D, where D is set to be non-trainable when updating G. The objective is to minimize the generator's loss, encouraging it to produce outputs that are classified as real by D. This adversarial objective for G is:

$$\mathcal{L}_G = -\mathbb{E}_{I_{color} \sim G(I_{gray}^{norm})}[\log D(I_{color})] \tag{18}$$

Additionally, perceptual or pixel-wise loss $\mathcal{L}_{perceptual}$ may be included to ensure better fidelity.

3. Train the GAN Model

The GAN training involves iteratively updating G and D:

- **Input Batch Selection**: Select a batch of grayscale images I_{gray}^{batch}.
- **Image Generation**: Generate color images $I_{color}^{gen} = G(I_{gray}^{batch})$.
- **Label Assignment**: Label real images I_{real} as 1 and generated images I_{color}^{gen} as 0.
- **Discriminator Training**:
  - Train D on I_{real} with label 1.
  - Train D on I_{color}^{gen} with label 0.
- **Generator Training**:
  - Freeze D and update G to maximize $D(I_{color}^{gen})$.

The training objective is formalized as:

$$\min_G \max_D \mathcal{L}_{GAN} = \mathbb{E}_{I_{real} \sim p_{data}}[\log D(I_{real})] \tag{19}$$
$$+\mathbb{E}_{I_{color} \sim G(I_{gray}^{norm})}[\log(1 - D(I_{color}))]$$

E. Testing with New Grayscale Images

After training, the generator is tested on new grayscale images I_{gray}^{test}, not seen during training. Given I_{gray}^{test}, the generator produces a colorized output:

$$I_{color}^{test} = G(I_{gray}^{test}) \tag{20}$$

This process evaluates the generator's performance on previously unseen data, assessing its ability to generalize and produce realistic colorizations.

F. Generate and Display Colorized Images

The generator's output, I_{color}^{test}, is visually inspected and compared with real color images (if available). As training progresses, generated images are expected to exhibit improved realism and visual coherence, validating the generator's colorization performance.

G. End of Process

The GAN colorization process concludes after the generated images are displayed and evaluated. If the results do not meet expected quality thresholds, hyperparameters (e.g., learning rate, batch size, or epoch count) may be tuned, or additional training data may be incorporated to enhance the model's performance.

Algorithm 1: GAN-Based Image Colorization

Input: Dataset of grayscale images, pre-trained or initialized generator G and discriminator D
Output: Colorized images for new grayscale inputs
1: **Load and Preprocess Grayscale Images:**
2: For each grayscale image I_{gray} in dataset:
3: Resize I_{gray} to fixed dimensions (e.g., 256x256)
4: Normalize pixel values of I_{gray} to range $[0,1]$
5: Store preprocessed images in `preprocessed_data`

6: **Define Generator Model (G):**
7: Initialize G with a series of convolutional and deconvolutional layers
8: Configure activation functions (e.g., ReLU, Tanh) in layers
9: Output colorized image $I_{color} = G(I_{gray})$

10: **Define Discriminator Model (D):**
11: Initialize D as a binary classifier with convolutional layers
12: Set final layer activation as sigmoid for binary output

Output probability $y = D(I)$ where I is either real or generated

Compile and Train Discriminator:
Set loss function for D as binary cross-entropy
Use Adam optimizer with learning rate lr and momentum parameters
Set $D.trainable = \mathrm{True}$

Compile GAN Model:
Combine G and D into a single GAN model
Freeze D during G training by setting $D.trainable = \mathrm{False}$
Set GAN loss to binary cross-entropy, assigning "real" label to generated images
Use Adam optimizer for GAN with chosen hyperparameters

Train GAN:
For each epoch in `total_epochs`:
  For each batch in `preprocessed_data`:
    Generate Fake Images:
      Select batch of grayscale images I_{gray}^{batch}
      Generate fake color images $I_{color}^{fake} = G(I_{gray}^{batch})$

    Train Discriminator (D):
      Define `labels_real` as array of ones for real images
      Define `labels_fake` as array of zeros for fake images
      Train D with real images $(I_{color}^{real}, labels\_real)$
      Train D with fake images $(I_{color}^{fake}, labels\_fake)$

    Train Generator (G) to Fool Discriminator:
      Freeze D by setting $D.trainable = \mathrm{False}$
      Define `labels_trick` as array of ones to "trick" D
      Train GAN with I_{gray}^{batch} and `labels_trick`, updating G

    Display Progress:
      Every N epochs, display samples of generated images and record loss values

Testing with New Grayscale Images:
For each new grayscale image I_{gray}^{test}:
  Preprocess I_{gray}^{test}
  Generate colorized output $I_{color}^{test} = G(I_{gray}^{test})$
  Display or save I_{color}^{test}

End of Process:
Evaluate model performance; if unsatisfactory, adjust hyperparameters or expand training data

Using a structured deep learning approach, this GAN-based colorization algorithm turns grayscale images—like old photos—into colorized versions. Images are first preprocessed, resized, and normalized for consistent input. While the discriminator model distinguishes between real and generated images, the generator model, which is made up of convolutional and deconvolutional layers, is trained to map grayscale pixels to realistic color values. When combined, they create a GAN, in which the generator gradually gains the ability to "trick" the discriminator by creating colorized images that are more lifelike.

The generator processes grayscale images during training, and the discriminator assesses the colorized images that are produced. The generator learns to generate high-quality colorizations by adjusting to reduce the discriminator's

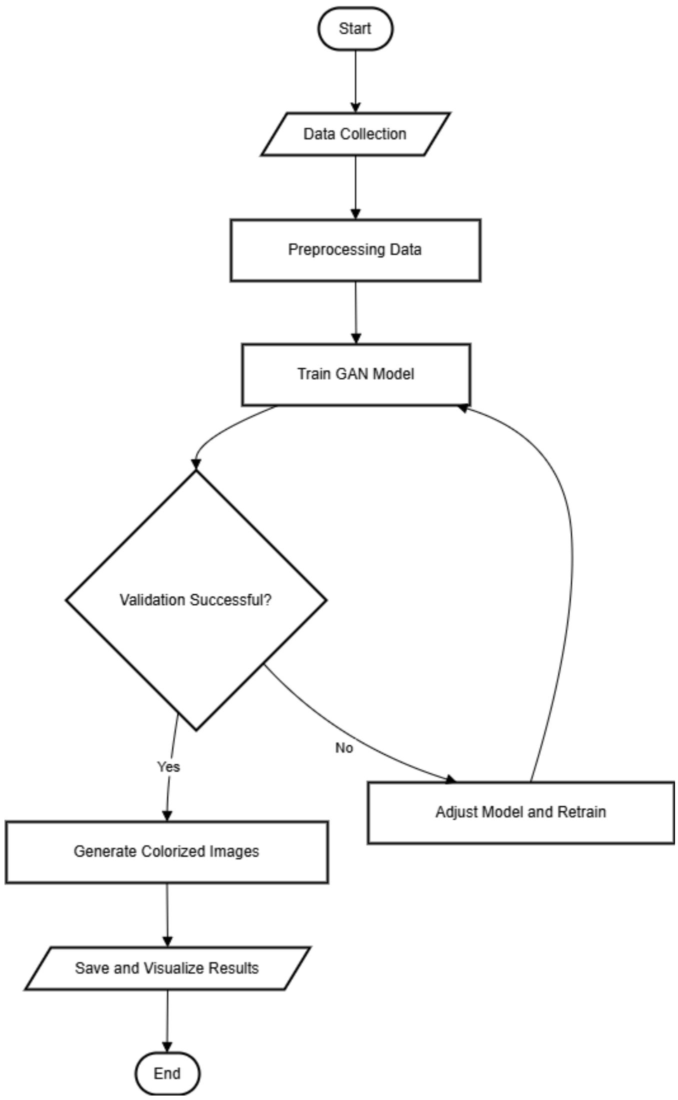

Fig. 1. Flowchart of Colorizing Grayscale Image using GANs.

success. The model can produce colored images for fresh grayscale inputs after training. By adding color to old photographs, this technique enhances their historical and aesthetic value while preserving cultural heritage.

The workflow for the suggested GAN-based image colorization method is shown in Fig. 1. The grayscale image is loaded first, and then the discriminator and generator models are defined. The training process begins once the discriminator has been compiled and the GAN model has been built. While the discriminator assesses the images' authenticity, the generator converts grayscale inputs into colorized images. Both models are iteratively improved, and the generator is trained to increase its output by tricking the discriminator. The model generates and displays the final colorized outputs after being trained on fresh grayscale images.

IV. Result

The following images are the results of image colorization using GANs.

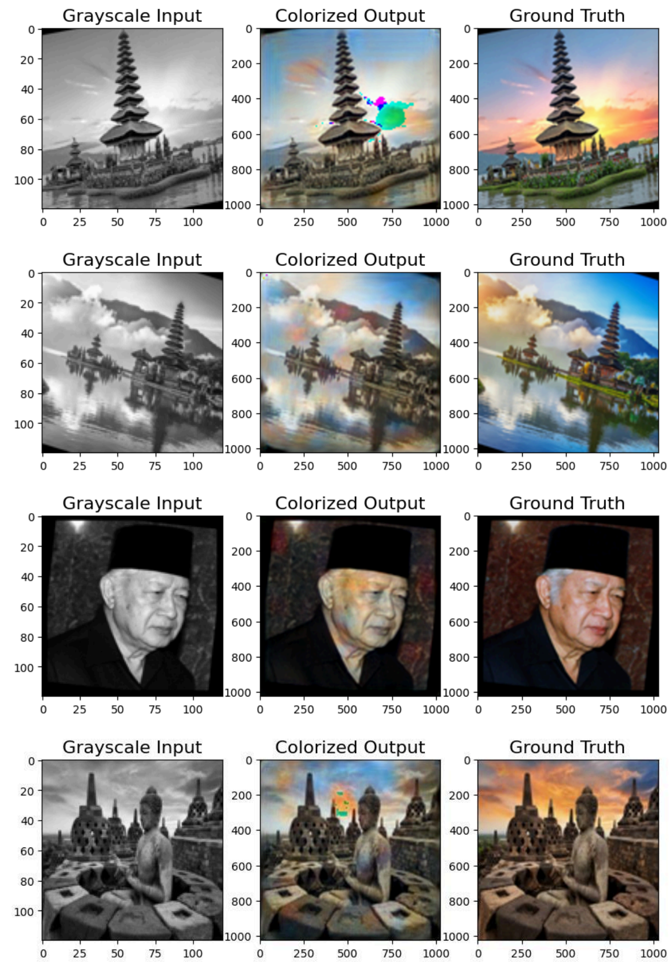

Fig. 2. The Result of Image Colorization using GANs.

The image colorization process using Generative Adversarial Networks (GANs) [10] has shown successful results when applied to grayscale images, as demonstrated by the outputs generated in this experiment. The GANs model was trained to transform grayscale images into realistic, colorful representations.

Each output image was processed by the generator model, which took a grayscale image as input and produced a colorized image. The model was trained over 150 epochs, using a dataset of resized RGB and grayscale images. The generator network, designed with several convolutional layers, effectively learned how to map grayscale images to colorized outputs, while the discriminator network ensured that the generated images closely resembled real-world images.

The outputs demonstrate realistic color details, such as natural skin tones and environmental hues like blue skies. The model automatically balanced color distributions without requiring manual intervention, producing visually accurate colorizations. Each output was compared against ground-truth images to assess quality.

To evaluate the quality of the colorized images, Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Map (SSIM) metrics were employed. These metrics provide quantitative assessments of the reconstructed images' fidelity and similarity to the original ground-truth images. Higher PSNR values indicate better quality, while SSIM measures structural consistency and perceptual similarity. Here are the mathematical formulations for Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Map (SSIM) [29]:

1. Peak Signal-to-Noise Ratio (PSNR):

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{L^2}{\mathrm{MSE}}\right) \tag{21}$$

Where:

- L: Maximum possible pixel value of the image (e.g., 255 for 8-bit images).
- MSE: Mean Squared Error between the original and reconstructed images, calculated as:

$$\mathrm{MSE} = \frac{1}{M \cdot N}\sum_{i=1}^{M}\sum_{j=1}^{N}[I(i,j) - K(i,j)]^2 \tag{22}$$

Here:

- $I(i,j)$: Pixel value at position (i,j) in the original image.
- $K(i,j)$: Pixel value at position (i,j) in the reconstructed image.
- M, N: Dimensions of the image.

2. Structural Similarity Index Map (SSIM):

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{23}$$

Where:

- μ_x: Mean of the original image (x).
- μ_y: Mean of the reconstructed image (y).

- σ_x^2: Variance of the original image.
- σ_y^2: Variance of the reconstructed image.
- σ_{xy}: Covariance between the original and reconstructed images.
- C_1, C_2: Small constants to avoid division by zero, typically $C_1 = (k_1 L)^2$ and $C_2 = (k_2 L)^2$, where L is the dynamic range of pixel values, and k_1, k_2 are constants (e.g., $1k_1 = 0.01$, $k_2 = 0.03$).

These equations were used to evaluate the fidelity and perceptual similarity of the colorized images compared to their ground truth.

The performance of the image colorization model was quantitatively evaluated using two metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Map (SSIM). The model achieved an average PSNR of 29.548, indicating a high level of similarity between the colorized images and the ground truth in terms of pixel-wise accuracy. Additionally, the SSIM averaged 0.866, reflecting a strong structural similarity and perceptual alignment with the original images. These metrics demonstrate the effectiveness of the GANs-based approach in generating realistic and visually accurate colorized images from grayscale inputs.

While the results are high-quality, some artifacts are noticeable in regions with intricate details. These artifacts can likely be mitigated with further training or model fine-tuning. Overall, the results highlight the versatility and effectiveness of GANs in image colorization tasks.

V. Conclusion

Generative Adversarial Networks (GANs) [10] were successfully used in this study to colorize old Indonesian photos. The outcomes showed how well the GANs model worked, generating pictures with realistic color details like skin tones, skies, and other environmental features. The model did not, however, attain perfect accuracy; small mistakes were noted in some places where the colorization was not entirely consistent. Notwithstanding these drawbacks, the model exhibits enormous promise, especially considering that these outcomes were obtained without the need for manual corrections or preset thresholds.

The study's findings have the potential to greatly advance government initiatives aimed at conserving historical archives. Furthermore, this method could be expanded to reconstruct historical videos with additional development. When compared to conventional colorization techniques, the GAN-based approach provides a notable time advantage. The generator and discriminator's cooperative architecture, which combines to create genuine colorization results, is the source of this efficiency.

Further studies could concentrate on enhancing the model's accuracy through the colorization process's incorporation of explicit cultural and contextual information. By doing this, the results would be more faithful to the original shades and subtleties of Indonesian culture. Furthermore, obtaining greater accuracy requires growing the dataset, particularly when dealing with a wide variety of old photos.

References

[1] A. Valeiro, "A History of Portugal Through Images: The Role of Musical Iconography in the Conception of Symbology's of Narration, Meaning and Power," *Springer series in design and innovation*, pp. 643–654, 2022, doi: https://doi.org/10.1007/978-3-031-09659-4_46.

[2] A. Büttner and C. Mauntel, "Does Klio Count, Too? Measurement and Understanding in Historical Science," *Measurement and Understanding in Science and Humanities: Interdisciplinary Approaches*, pp. 39–49, 2023, doi: https://doi.org/10.1007/978-3-658-36974-3_4.

[3] R. Sandford, "Thinking with heritage: Past and present in lived futures," *Futures*, vol. 111, pp. 71–80, 2019, doi: https://doi.org/10.1016/j.futures.2019.06.004.

[4] J. K. Yamashiro and E. Pérez-Amparán, "Memory media design shapes perceived temporal distance of depicted historical events: Color versus Black and White photographs.," *Journal of Applied Research in Memory and Cognition*, 2023, doi: https://doi.org/10.1037/mac0000144.

[5] M. Y. Zara, "Photographic Representations of the Dutch–Indonesian Conflict in the Yogyakarta-Based Kedaulatan Rakjat Newspaper, 1945–1946," *Bijdragen tot de taal-, land- en volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia*, vol. 180, no. 2–3, pp. 212–241, 2024, doi: https://doi.org/10.1163/22134379-bja10058.

[6] P. Sharma, M. A. Ahmad, K. Gupta, and Pashupati Baniya, "Image Colorization Enhancement Techniques using Deep Convolutional GANs," *2023 IEEE 15th International Conference on Computational Intelligence and Communication Networks (CICN)*, pp. 576–583, 2023, doi: https://doi.org/10.1109/cicn59264.2023.10402312.

[7] R. Rahul and D. Walecha, "Reviving Black and White Images: Enhancing Colorization with Generative Adversarial Networks (GANs)," *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2023, doi: https://doi.org/10.1109/icccnt56998.2023.10306534.

[8] E.-C. Gross, "Colours of the Past: Considerations on Photographic Colourisation of Archival Photographs," *Baltic Journal of Art History*, vol. 23, 2022, doi: https://doi.org/10.12697/bjah.2022.23.06.

[9] B. Figueiredo et al., "The Power Of Colors To Maximize Attention And Readability In Visual Communication: Insights From An Eye-Tracking Behavioural Study," *World Congress on Electrical Engineering and Computer Systems and Science*, 2024, doi: https://doi.org/10.11159/mhci24.101.

[10] I. J. Goodfellow *et al.*, "Generative Adversarial Networks," *arXiv preprint*, 2014, doi: https://doi.org/10.48550/arxiv.1406.2661.

[11] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," *arXiv preprint*, 2014, doi: https://doi.org/10.48550/arXiv.1411.1784.

[12] N. Wang, G.-D. Chen, and Y. Tian, "Image Colorization Algorithm Based on Deep Learning," *Symmetry*, vol. 14, no. 11, p. 2295, 2022, doi: https://doi.org/10.3390/sym14112295.

[13] S. Srinivas, R. K. Sarvadevabhatla, K. R. Mopuri, N. Prabhu, S. S. S. Kruthiventi, and R. V. Babu, "A Taxonomy of Deep Convolutional Neural Nets for Computer Vision," *Frontiers in Robotics and AI*, vol. 2, 2016, doi: https://doi.org/10.3389/frobt.2015.00036.

[14] K. Kiani, R. Hematpour, and R. Rastgoo, "Automatic Grayscale Image Colorization using a Deep Hybrid Model," *Journal of AI and Data Mining*, vol. 9, no. 3, 2021, doi: https://doi.org/10.22044/jadm.2021.9957.2131.

[15] H. Carrillo, "Guiding neural networks for image colorization through user interactions," *Doctoral dissertation, Université de Bordeaux*, 2024, doi: https://theses.hal.science/tel-04446168.

[16] H. Safiq and B. Lee, "Image Colorization Using Color-Features and Adversarial Learning," *IEEE access*, vol. 11, pp. 132811–132821, 2023, doi: https://doi.org/10.1109/access.2023.3335225.

[17] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," *arXiv preprint*, 2016, doi: https://doi.org/10.48550/arxiv.1611.07004.

[18] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," *arXiv preprint*, 2017, doi: https://doi.org/10.48550/arxiv.1703.10593.

[19] M. R. Joshi, L. Nkenyereye, G. P. Joshi, S. M. R. Islam, M. Abdullah-Al-Wadud, and S. Shrestha, "Auto-Colorization of Historical Images Using Deep Convolutional Neural Networks," *Mathematics*, vol. 8, no. 12, p. 2258, 2020, doi: https://doi.org/10.3390/math8122258.

[20] A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 689–694, 2004, doi: https://doi.org/10.1145/1015706.1015780.

[21] Z. Cheng, Q. Yang, and B. Sheng, "Deep Colorization," *arXiv preprint*, 2016, doi: https://doi.org/10.48550/arxiv.1605.00075.

[22] R. Shiva. Shankar, G. Mahesh, K. V. S. S. Murthy, and D. Ravibabu, "A Novel approach for Gray Scale Image Colorization using Convolutional Neural Networks," *2020 international conference on system, computation, automation and networking (ICSCAN)*, pp. 1–8, 2020, doi: https://doi.org/10.1109/ICSCAN49426.2020.9262377.

[23] R. Zhang, P. Isola, and A. A. Efros, "Colorful Image Colorization," *arXiv preprint*, 2016, doi: https://doi.org/10.48550/arxiv.1603.08511.

[24] A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," *arXiv preprint*, 2016, doi: https://doi.org/10.48550/arXiv.1511.06434.

[25] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Globally and locally consistent image completion," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–14, 2017, doi: https://doi.org/10.1145/3072959.3073659.

[26] H. Tang, D. Xu, N. Sebe, and Y. Yan, "Attention-Guided Generative Adversarial Networks for Unsupervised Image-to-Image Translation," *arXiv preprint*, 2019, doi: https://doi.org/10.48550/arxiv.1903.12296.

[27] R. Luo, R. Luo, L. Guo, and H. Yu, "An Ancient Chinese Painting Restoration Method Based on Improved Generative Adversarial Network," *Journal of Physics: Conference Series*, vol. 2400, no. 1, p. 12005, 2022, doi: https://doi.org/10.1088/1742-6596/2400/1/012005.

[28] H. Neji *et al.*, "Doc-Attentive-GAN: attentive GAN for historical document denoising," *Multimedia Tools and Applications*, vol. 83, no. 18, pp. 55509–55525, 2024, doi: https://doi.org/10.1007/s11042-023-17476-2.

[29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004, doi: https://doi.org/10.1109/tip.2003.819861.

Appendix

Appendix 1: Source Code of Image Colorization using GAN

The source code is available at the following github link: https://github.com/aimldlnlp/RKK303-Computer-Vision-Final-Project/blob/main/Image_Colorization_With_GANs_Final.ipynb