Image Colorization using Generative Adversarial Networks (GAN) Architecture

Sakkiya KS (12211424)
Department of Computer Science and Engineering, Lovely professional university.
Delhi-Jalandhar GT Road, Phagwara, Punjab, India(144001)
sakkiya.k@lpu.in

***Abstract* - This research paper presents a novel approach to image colorization using deep learning techniques, specifically Generative Adversarial Networks (GANs) and the Adam optimizer. The model uses GANs to create realistic and visually pleasing colorizations for grayscale images, with the generator network learning to predict color information and the discriminator network assessing authenticity. The Adam optimizer is used to fine-tune the model's parameters, leading to faster convergence and improved performance. The model is effective in adding color to grayscale images, preserving fine details and textures, and demonstrating its ability to respect the nuances of original images. The research has potential applications in image restoration, historical photograph colorization, and entertainment industries. The combination of GANs and the Adam optimizer provides a robust framework for developing accurate and visually appealing colorization models, contributing to the ongoing progress in deep learning-based image processing.**

Index Terms – Generative Adversarial Network, GANs, Computer Vision, Deep Learning, Image Colorization

1. INTRODUCTION

In the realm of computer vision, image colorization stands as a fascinating and challenging task that bridges the gap between art and technology. The ability to convert grayscale images into their colorful counterparts has wide-ranging applications, from the restoration of historical photographs to the enhancement of modern media content. In recent years, deep learning techniques have revolutionized the field, enabling the development of sophisticated models capable of automating this intricate process. This research paper delves into the realm of image colorization, introducing a novel approach that leverages the power of Generative Adversarial Networks (GANs) and the efficiency of TensorFlow, with the Adam optimizer for model optimization.

The fundamental challenge in image colorization lies in understanding and predicting the intricate relationships between grayscale and color information in images. The human visual system is incredibly adept at filling in the gaps when presented with grayscale content, mentally applying colors based on context and prior knowledge. Simulating this cognitive process using machine learning models is a nontrivial endeavor, but one that has seen remarkable progress in recent years.

Deep learning, particularly GANs, has emerged as a formidable tool in the pursuit of automated image colorization. GANs, composed of a generator and a discriminator, engage in an adversarial training process where the generator seeks to produce realistic colorized images, while the discriminator attempts to distinguish between real and generated images. This competitive dynamic encourages the generator to produce increasingly accurate colorizations, driving the model towards a level of proficiency that was difficult to achieve with traditional image processing techniques.

Furthermore, the choice of deep learning libraries is critical to the success of the project. Keras, a high-level neural networks API, and TensorFlow, an open-source machine learning framework, are widely adopted and have become the backbone of numerous image processing tasks. The research described in this paper harnesses the capabilities of these libraries to build a robust image colorization model.

The optimizer used for training, Adam, plays a pivotal role in shaping the model's learning process. Adam combines the advantages of two other popular optimizers, the momentum and RMSprop, providing fast and efficient convergence during training. Its usage is instrumental in fine-tuning the GAN model for image colorization.

This paper presents the methodology and results of a rigorous experimentation process, showcasing the effectiveness of the proposed approach. The research contributes to the everevolving field of computer vision, pushing the boundaries of automated image processing and colorization. It also offers a glimpse into the potential applications of this technology, ranging from preserving historical photographs to transforming grayscale video content into vibrant, engaging media. The journey into the world of image colorization through deep learning, as detailed in this paper, highlights the exciting prospects that lie at the intersection of art and technology.

2. LITERATURE REVIEW

Numerous researchers worldwide have delved into the realm of image colorization, contributing to a rich body of scholarly work. In the following sections, we provide concise reviews and insights into a selection of research papers on this subject.

1. **Image Colorization using Generative Adversarial Networks:** [1] The restoration of old or deteriorated photos has sparked significant interest in the field of automatic image colorization in recent years. Nevertheless, this subject presents a significant challenge as a result of the extensive range of possibilities in assigning colour information. A conditional Deep Convolutional Generative Adversarial Network (DCGAN) is employed in order to extend the colorization process, with training conducted on openly accessible datasets such as CIFAR-10 and Places365. A comparison is made between the outcomes derived from the generative model and those obtained from typical deep neural networks.

2. **Guided Image Generation with Conditional Invertible Neural Networks:** [2] The research paper presents a novel framework known as the conditional invertible neural network (cINN) that is designed to facilitate the development of natural images with the use of conditioning input. The Conditional Invertible Neural Network (cINN) integrates a generative Invertible Neural Network (INN) model with an unconstrained feed-forward network, effectively performing preprocessing on the conditioning input to extract valuable features. The model is capable of generating a wide range of samples without experiencing mode collapse, and it is able to produce high-quality images without incurring any loss in reconstruction fidelity.

3. **Wavelet Transform-assisted Adaptive Generative Modeling for Colorization:** [3] The present study introduces a novel approach that utilises a scorebased generative model in the wavelet domain to enhance the performance of unsupervised deep learning in tasks related to image colorization. The proposed model acquires more comprehensive prior knowledge by utilising stacked wavelet coefficient components. This approach effectively reduces the dimensionality of the original manifold and mitigates the challenges posed by the curse of dimensionality. The utilisation of two distinct consistency concepts, namely data-consistency and structure-consistency, is employed to enhance the effectiveness of the work. The conducted experiments have demonstrated significant advancements in both generation and colorization quality.

4. **Image Processing Using Multi-Code GAN Prior:** [4] The research introduces a revolutionary methodology known as mGANprior, which integrates proficiently trained Generative Adversarial Networks (GANs) into the realm of image processing problems. The proposed approach entails the utilisation of numerous latent codes for the purpose of generating feature maps. These feature maps are subsequently combined with adaptive channel significance. The utilisation of over-parameterization in this context has been found to enhance the quality of image reconstruction, surpassing the performance of competing methods. The high-fidelity reconstruction that ensues enables the utilisation of GAN models in practical scenarios, such as image colorization, superresolution, inpainting, and semantic manipulation.

5. **Colorization Transformer:** [5] The Colorization Transformer presents a revolutionary methodology for achieving high-fidelity image colorization through the use of self-attention. The proposed method employs a conditional autoregressive transformer for generating low-resolution coarse colouring. Subsequently, two parallel networks are utilised to upsample the image and provide finely coloured high-resolution outputs. The integrity of the colorings exhibits superior performance compared to earlier methodologies, as evidenced by the preference of over 60% of human assessors towards the colouring with the highest rating.

6. **Grayscale Image Colorization Methods:** Overview and Evaluation: [6] Colorization refers to the technical procedure of converting monochromatic or grayscale images into colourized representations, with the primary objective of persuading observers of their verisimilitude. In the last two decades, a multitude of methodologies have been devised, encompassing both rudimentary algorithms and sophisticated automated techniques. The present study undertakes an evaluation of grayscale image colorization techniques, with a specific emphasis on approaches utilising deep learning methodologies. This study evaluates several methodologies by considering their impact on image quality and processing time, employing diverse metrics to gauge the perceived quality. The category of colorization neural networks that are guided by user input holds significant promise, mostly attributed to the effective combination of human participation and automation.

7. **Analysis of generative adversarial networks:** [7] The study of generative models in the field of computer vision and image classification has resulted in the emergence of generative adversarial networks (GANs), which are artificial intelligence algorithms designed to analyse training patterns and their corresponding distributions. These networks find utility in diverse applications such as video generation, music generation, image synthesis, and text-to-image translation, and have been evaluated and compared using specified criteria.

8. **A review and meta-analysis of Generative Adversarial Networks and their applications in remote sensing:** [8] The utilisation of Generative Adversarial Networks (GANs) in the field of Deep Learning has emerged as a notable progress, particularly in the domain of Remote Sensing (RS), where their application has been seeing a rapid growth. A systematic examination of 231 research publications pertaining to Generative Adversarial Networks (GANs) was undertaken in order to acquaint the Recommender Systems (RS) community with the potential of GANs and facilitate further exploration of their applications by researchers. The research revealed that the predominant utilisation of Generative Adversarial Networks (GANs) lies in picture classification, with a specific focus on urban mapping. Nevertheless, there is a limited body of research that has delved into the possibilities of Generative Adversarial Networks (GANs) in analysing multi-spectral images with medium spatial resolution. The study also identified knowledge gaps pertaining to the appropriate selection of GAN models for various applications and their potential to serve as substitutes for genuine remote sensing data.

9. **Image Colorization Using Generative Adversarial Networks and Transfer Learning:** [9] The process of automatic colorization in computer graphics entails the transformation of grayscale onedimensional pictures into three-dimensional images with chromatic components. Convolutional neural networks (CNNs) are frequently employed for this procedure. This study employs the theoretical framework of "Hypercolumn" derived from neuroscience in order to construct a completely automated system for image colorization. The VGG19 model is employed as a pre-trained model within the generator network, and the proposed approach demonstrates superiority over alternative models.

10. **GAN-Based Image Colorization for SelfSupervised Visual Feature Learning:** [10] This research study presents a novel approach to selfsupervised learning techniques for the automatic acquisition of visual features, specifically emphasising the application of generative adversarial networks (GANs) in the context of image colorization. Conditional Generative Adversarial Networks (cGANs) have been employed in the field of picture colorization and have also been adapted for multilabel image classification and semantic segmentation tasks. The experimental results indicate a 5% improvement in performance for classification tasks and a 2.5% improvement for segmentation tasks, so illustrating the potential of Generative Adversarial Networks (GANs) in the context of selfsupervised feature learning.

3. GENERATIVE ADVERSARIAL NETWORKS (GAN)

A Generative Adversarial Network, often referred to as a GAN, is a type of artificial neural network architecture used in machine learning and deep learning. It was introduced by Ian Goodfellow and his colleagues in 2014 and has since become a powerful tool for various creative and generative tasks. At its core, a GAN consists of two main components: a generator and a discriminator. These two networks work in tandem, and the training process resembles a competitive game.

1. Generator: The generator's role is to create synthetic data, such as images, sounds, or text, from random noise or other input data. It essentially learns to produce data that is as similar as possible to real data from a training dataset. The generator's objective is to generate data that is indistinguishable from real data.

2. Discriminator: The discriminator, on the other hand, is like a detective. It evaluates data and tries to distinguish between real data and data generated by the generator. It's trained to provide a probability score, indicating whether the input data is real or fake. The discriminator's objective is to become better at telling the difference between real and generated data.

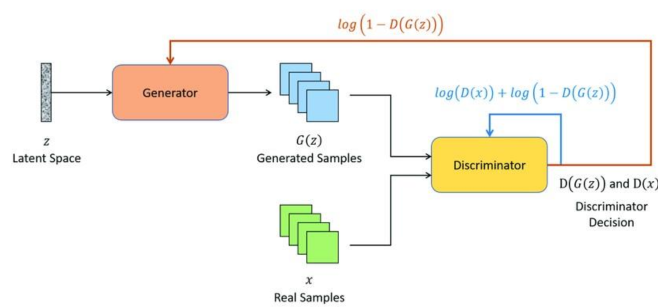


Figure 1. Typical Generative Adversarial Networks (GAN) architecture[11].

The training process of a GAN is characterized by a back-andforth competition between the generator and the discriminator:

- The generator tries to improve its ability to produce data that looks real to the discriminator.
- The discriminator, in turn, strives to become better at correctly classifying real and fake data.

This adversarial relationship leads to an iterative improvement in both the generator and discriminator. Over time, the generator becomes more skilled at creating data that is increasingly similar to real data, while the discriminator becomes better at distinguishing between real and generated data.

GANs have found applications in various domains, including image generation, image-to-image translation, style transfer, and more. They are often used for creative purposes, such as generating art, music, or realistic images from sketches. The interplay between the generator and discriminator in GANs allows for the generation of highly realistic and novel content, making them a valuable tool in the world of generative modeling.

4. CNN VS GAN

Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) are two different types of neural network architectures, each designed for specific tasks and with distinct characteristics. Here's a comparison of CNNs and GANs in neural networks:

4.1 Convolutional Neural Networks (CNNs):

1. Task: CNNs are primarily used for tasks related to image analysis, computer vision, and feature extraction. They are well-suited for tasks like image classification, object detection, and image segmentation.

2. Architecture: CNNs consist of layers of convolutional, pooling, and fully connected layers. These layers are designed to extract hierarchical features from input images.

3. Training Objective: In CNNs, the primary training objective is to minimize a loss function (e.g., cross-entropy) that measures the difference between predicted and actual class labels. CNNs are used for supervised learning.

4. Use Cases: CNNs are widely used in image classification, face recognition, image generation, and various computer vision tasks. They are employed when the goal is to make predictions or segment objects in images.

5. Generative Capability: While CNNs can be used for tasks like image generation, they are typically not as proficient at generating images as GANs, which are specifically designed for generative tasks.

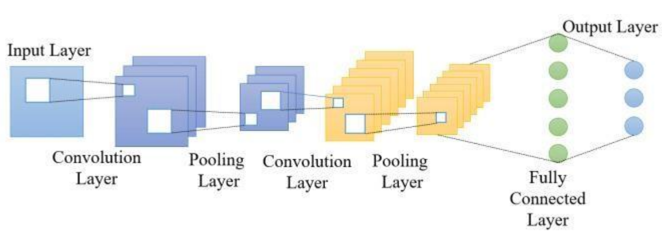


Figure 2 Basic architecture of CNN

4.2 Generative Adversarial Networks (GANs):

1. Task: GANs are generative models designed for creating new data that resembles a given dataset. They excel at generating images, sounds, text, and other forms of data. GANs are not limited to image data.

2. Architecture: GANs consist of two neural networks: a generator and a discriminator. The generator creates synthetic data, while the discriminator evaluates the authenticity of the generated data. GANs operate in an adversarial training setup.

3. Training Objective: GANs do not aim to minimize a specific loss but rather engage in a competitive game. The generator seeks to produce data that is indistinguishable from real data, while the discriminator aims to differentiate between real and fake data. GANs use unsupervised learning.

4. Use Cases: GANs are used for tasks like image generation, image-to-image translation, style transfer, and data augmentation. They are valuable for creative tasks and for producing data that wasn't present in the original training dataset.

5. Generative Capability: GANs are renowned for their generative capabilities, as they can create high-quality images that are often visually indistinguishable from real images. They are particularly valuable for artistic and creative applications.

In summary, CNNs are typically used for tasks that involve classifying or analyzing existing data, while GANs are designed for generating new data, especially in the form of images. The choice between CNNs and GANs depends on the specific task and the nature of the data, with CNNs being more suitable for recognition and classification, and GANs excelling at generative and creative tasks..

5. LOSS FUNCTION AND OPTIMIZATION ALGORITHM

5.1 Mean Squared Error (MSE) loss:

The L2 loss function, also known as Mean Squared Error (MSE) loss, is a commonly used mathematical measure in machine learning and statistics to quantify the difference between two sets of values, typically referred to as the predicted values and the actual or target values. It is widely used for regression tasks, where the goal is to predict a continuous numerical value.

The Formula for MSE,

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

5.2 Log Loss:

Log loss, also known as logarithmic loss or cross-entropy loss, is a commonly used loss function in machine learning, particularly in classification tasks. It is often used to measure the difference between predicted probabilities and actual class labels, making it a fundamental component in the training of classification models.

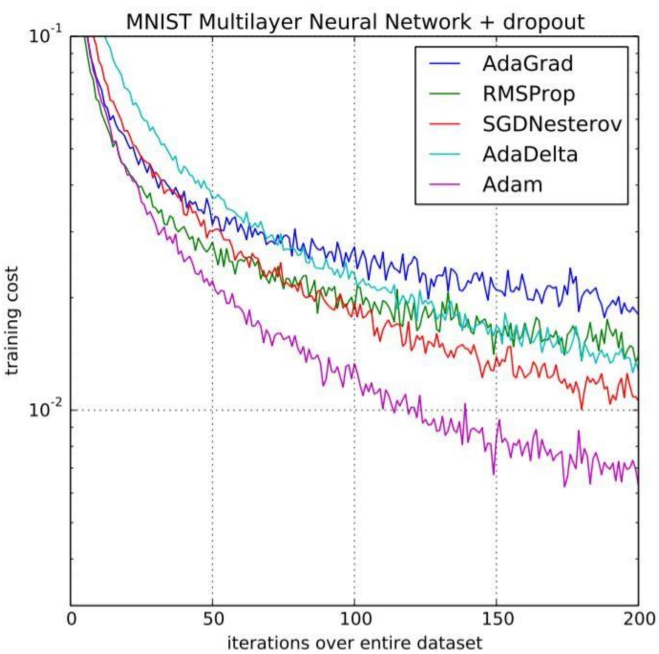
The Formula for Log Loss,

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

Binary Cross-Entropy / Log Loss

5.3 Adam Optimization:

The Adam optimization[13] algorithm is a popular choice in the field of deep learning, particularly in computer vision and natural language processing, as it builds upon the stochastic gradient descent method. The algorithm iteratively adjusts network weights using training data and is capable of addressing diverse optimization challenges by incorporating the benefits of both RMSprop and Momentum techniques.



6. METHODOLOGY

1. Downloading and Processing the data:

In order to train the model, we'll first require a dataset. A dataset of RGB images to train the GAN model whose images consist of various scenes/places will be great. Then we parsed the images (RGB images to be precise) one by one and transformed each one to a grayscale image using PIL's *[.convert('L')]* method. So our dataset will have samples of (*grayscale image, RGB image*)

2. Network Architecture Selection:

We use GAN architecture for the entire project, is a type of artificial neural network architecture used in machine learning and deep learning. The primary purpose of a GAN is to generate new data that is similar to a given dataset.
It consists of two main components:

1. Generator: Here generator (represented as G) will take in grayscale image "x" and produce a RGB image G(x).
2. Discriminator: The discriminator model, represented as D, will take in the real image y (from the training data) and the generated image G(x) (from the generator) to output two probabilities.

3. Model Training:

We train the discriminator in such a manner that is able to differentiate the real images and the generated images. So, train the model such that y produces a output of 1.0 and G(x) produces an output of 0.0. Further we use soft label which are close to 1 and 0.

4. Loss Function Design:

We will now implement the loss functions for our GAN model. As you might know that we have two loss functions, one for the generator and another for the discriminator.

- For our generator, we'll use the L2/MSE loss function.
- For optimization, we use the Adam optimizer with a learning rate of 0.001

5. Training The GAN:

1. Training a Generative Adversarial Network (GAN) involves an iterative process where the generator and discriminator networks are simultaneously improved.
2. We use @tf.function which used to converts the Python function into a TensorFlow graph for better performance when running on GPUs or TPUs. It compiles the code for faster execution.
3. Now it will calculate the losses that includes both generator loss and discriminator loss
4. After computing the losses, the code proceeds to calculate gradients using the tf.GradientTape contexts
5. Finally, the code applies the gradients to the network weights using the Adam optimizer
6. We used 150 epochs which taken around 4 hours to compile the code, which iterates through the training dataset and calls the train_step function, which is

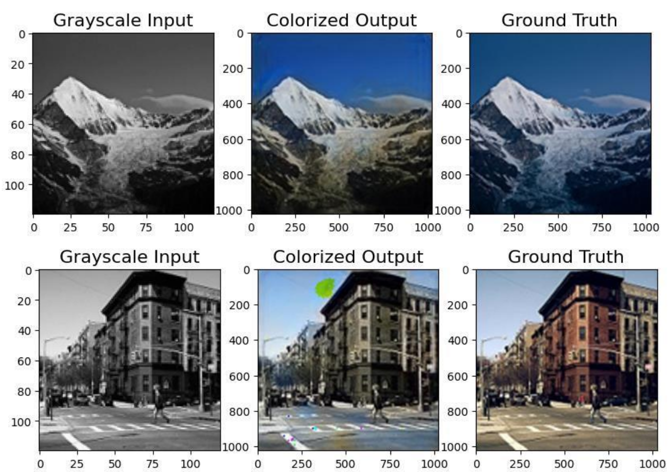
responsible for updating the GAN's generator and discriminator networks based on the provided input data (x and y).

By following this comprehensive methodology, the research paper can provide a detailed account of the process and insights gained while developing an automated image colorization model using GANs, TensorFlow, and the Adam optimizer.

7. RESULT

The findings from the image colorization research project provide evidence for the effectiveness of the proposed methodology, which involves the utilisation of a Generative Adversarial Network (GAN) model trained using TensorFlow and optimised through the Adam optimizer. The evaluation of the model's performance is conducted using a diversified dataset consisting of grayscale photos. This evaluation aims to demonstrate the model's capability to generate colorizations that are both realistic and visually appealing.

Lastly, visualise an image to observe the outcomes as presented below. However, it is important to note that there is still some noise present as a result of artefacts in the training process. Nevertheless, the final outcome still yields a visually striking image, particularly due to the incorporation of colorization.



8. CONCLUSION

In conclusion, the research demonstrates that the proposed methodology, utilizing GANs, TensorFlow, and the Adam optimizer, yields highly promising results in the domain of image colorization. The model's ability to produce realistic and visually appealing colorizations has implications in various fields, including art restoration, historical preservation, and multimedia enhancement. The research paves the way for future advancements in automated image colorization and sets a solid foundation for further innovation in this exciting intersection of art and technology..

6. REFERENCES

[1] K. Nazeri, E. Ng, & M. Ebrahimi, "Image colorization using generative adversarial networks", Articulated Motion and Deformable Objects, p. 85-94, 2018. https://doi.org/10.1007/978-3-319-94544-6_9

[2] Ardizzone, L., Lüth, C., Kruse, J., Rother, C., & Köthe, U. (2019, July 4). Guided Image Generation with Conditional Invertible Neural Networks. arXiv.org. https://arxiv.org/abs/1907.02392v3

[3] Li, J., Li, W., Xu, Z., Wang, Y., & Liu, Q. (2021, July 9). Wavelet Transform-assisted Adaptive Generative Modeling for Colorization. arXiv.org. https://doi.org/10.1109/TMM.2022.3177933.

[4] Gu, J., Shen, Y., & Zhou, B. (2019, December 15). Image Processing Using Multi-Code GAN Prior. arXiv.org. https://arxiv.org/abs/1912.07116v2

[5] Kumar, M., Weissenborn, D., & Kalchbrenner, N. (2021, February 8). Colorization Transformer. arXiv.org. https://arxiv.org/abs/2102.04432v2

[6] Grayscale Image Colorization Methods: Overview and Evaluation. (n.d.). Grayscale Image Colorization Methods: Overview and Evaluation | IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/9512069.

[7] Imamverdiyev, Yadigar & Musayeva, Firangiz. (2022). Analysis of generative adversarial networks. Problems of Information Technology. 13. 20-27. 10.25045/jpit.v13.i1.03.

[8] A review and meta-analysis of Generative Adversarial Networks and their applications in remote sensing. (2022, March 17). A Review and Metaanalysis of Generative Adversarial Networks and Their Applications in Remote Sensing - ScienceDirect. https://doi.org/10.1016/j.jag.2022.102734

[9] Image Colorization Using Generative Adversarial Networks and Transfer Learning. (n.d.). Image Colorization Using Generative Adversarial Networks and Transfer Learning | IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/document/9116882.

[10] Treneska, S., Zdravevski, E., Pires, I. M., Lameski, P., & Gievska, S. (2022, February 18). GAN-Based Image Colorization for Self-Supervised Visual Feature Learning. MDPI. https://doi.org/10.3390/s22041599.

[11] Automatic Target Recognition for Low Resolution Foliage Penetrating SAR Images Using CNNs and GANs - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Typical-GenerativeAdversarial-Networks-GAN-architecture_fig2_349182009 [accessed 2 Nov, 2023]

[12] Gu, Hao & Wang, Yu & Hong, Sheng & Gui, Guan. (2019). Blind Channel Identification Aided Generalized Automatic Modulation Recognition Based on Deep Learning. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2934354.

[13] Kingma, Diederik & Ba, Jimmy. (2014). Adam: A Method for Stochastic Optimization. International Conference on Learning Representations.

[14] Yunji Chen, Ling Li, Wei Li, Qi Guo, Zidong Du, Zichen Xu, Chapter 2 - Fundamentals of neural networks, Editor(s): Yunji Chen, Ling Li, Wei Li, Qi Guo, Zidong Du, Zichen Xu, AI Computing Systems, Morgan Kaufmann, 2024, Pages 17-51, ISBN 9780323953993, https://doi.org/10.1016/B978-0-32-395399-3.00008-1.