Grenze International Journal of Engineering and Technology, Jan Issue

Satellite Image to Map Translation using GANs

Sheetal Phatangare[1], Mohammed Mansoor Khalifa[2], Sejal Kharche[3], Amaan Khatib[4] and Atharva Kshirsagar[5]
[1-5]Vishwakarma Institute of Technology /Computer Engineering Department, Pune, India
Email: sheetal.phatangare@vit.edu, {mohammedmansoor.khalifa20, sejal.kharche20, amaan.khatib20, Atharva.kshirsagar20}@vit.edu

***Abstract*— Generative adversarial networks (GANs) are a promising method for learning deep representations without spending a lot of time on data for annotated training. They accomplish this by obtaining signals that are back propagated through a competitive mechanism in-transforming two networks. The possible representations of a variety of applications may employ the knowledge gathered by GANs, image synthesis, semantic image editing, fashion categorization, transmission, and image super-resolution. The goal of this research article is to give an overview of GANs. the signal processing community, utilizing well-known whenever feasible, use concepts and analogies. Additionally, to recognize various training techniques and building GANs, we also highlight lingering issues with their theory and utilization.**

***Index Terms*— GANs, deep learning, image to image translation, neural network, generator and discrimator model, cGAN, pix2pixGAN.**

I. INTRODUCTION

Network of Generative Adversaries, i.e. Deep learning techniques include GANs. GANs are the methods used to produce human faces that are photorealistic as well as stunning image translations like photo colorization, face de-aging, super-resolution, and more. The creation of new photographs with a realistic appearance is the ideal application of GAN.

Examples: the creation of fresh anime characters, creating fresh logos, creating fresh Pokémon, and creating new apparel, among other things. The neural network architecture used for generative modelling is known as a Generative Adversarial Network or GAN.

Generative modelling is the process of using a model to generate new examples that are consistent with a given distribution of samples. It can be used for a variety of applications, including creating new images that share certain characteristics with an existing dataset while being significantly different from it. A pair of neural network models known as a Generative Adversarial Network (GAN) are frequently used in this approach. The "generator" or "generative network," the initial model, is taught to generate credible samples. The second model, also known as the "discriminator" or "discriminative network," develops the ability to tell created samples from genuine ones.

A GAN exposes the discriminator to both real and created instances while the generator model attempts to trick the discriminator model. These models are created to compete against one another in a game or contest that follows the rules of game theory.

During the initial stages of training, the generator often generates samples that are easily identifiable as fake by the discriminator. The discriminator, in turn, rapidly learns to differentiate between the real and generated data

As the training progresses and the generator improves, an interesting phenomenon occurs: the discriminator's ability to distinguish between real and fake data diminishes. It begins to classify the generated samples as real,

Grenze ID: 01.GIJET.10.1.550

resulting in a decrease in its accuracy. This can be seen as a measure of success for the generator, as it indicates that it is becoming more proficient at generating realistic data that closely resembles the real samples.

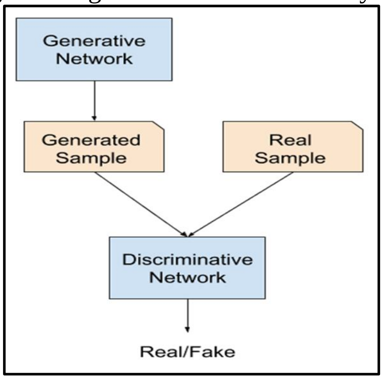

Figure 1: Overview of Generasive Adversial Network

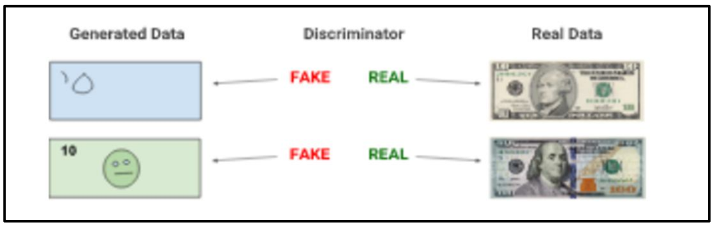

Figure 2: Understanding GAN

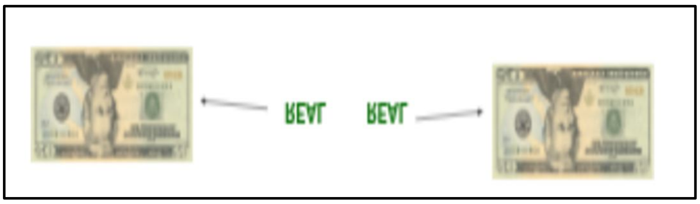

Figure 3: Understanding GAN

II. Literature Studies

Antonia Creswell eta, Tom White eta, Vincent Dumoulin eta, Kai Arulkumaran eta, Biswa Senguptats eta and Anil A Bharath in the paper "Genarative Adversial Networks" published on 19th October 2017 explains what are generators and discriminators in GAN's architecture, and how the data is trained and tested using GAN and the Mathematics behind the GAN is given networks.[1] The possible representations of a variety of applications may employ the knowledge gathered by GANs, image synthesis, semantic image editing, and fashion categorization, transmission, and image super-resolution. The objective of this review article is to give an overview of GANs.the signal processing community, utilizing well-known whenever feasible, use concepts and analogies. Additionally to recognising various training techniques and building GANs, we also highlight lingering issues with their theory and utilization.

Yilei Shi eta, Qingyu Li eta, and Xiao Xiang Zhu eta in the paper "Building Footprint Generation Using Improved Generative Adversarial Networks" published on 4th april 2019[3]. They Have Used the CGAN and WGAN for the Implementation part and training of the both models. And Collectively Proposed the term CWGAN. Image.Data regarding a building's footprint is crucial.a component of urban model rebuilding in three dimensions. it automatically building footprints can be generated using satellite photographs a significant issue given the complexity of the construction shapes. In this letter, we suggest more effective generative Using adversarial networks (GANs) to automatically produce satellite photos of the footprints of buildings. Utilizing a conditional GAN (CGAN) using a Wasserstein-derived cost function added a gradient penalty phrase and increased distance. The outcomes attained showed that the suggested strategy can greatly enhance the comparison to CGANs, the

building footprint generation's quality, and other networks, including the U-Net. Additionally, our approach for tweaking for all hyperparameters is removed.
Ian Goodfellow eta, Jean Pouget-Abadie eta, Mehdi Mirza eta, Bing Xu eta, David Warde-Farley eta, Sherjil Ozair eta, Aaron Courville eta, and Yoshua Bengio eta in the paper "Generative Adversarial Networks" published in the month of september 2018 shows how unsupervised learning is more convenient to use.[4] In generative modeling, training examples are drawn from an unknown distribution and then used for approximation which is based on density function.Generative Adversarial Networks having two networks viz.generator and discriminator network.Both are dependant on each other and uses backpropagation algorithm. One type of artificial intelligence is called generative adversarial networks.
created to tackle the generative model problem. A generative model's objective is to research a collection of practice examples and probability education distribution system that produced them. Adversarial Generative.
Then, networks (GANs) can provide more instances. based on the predicted probability distribution. Generative Deep learning models are ubiquitous, but GANs are some of the most effective generative models (particularly especially in terms of their capacity to provide believable high-picture resolution).
Jie Gui eta, Zhenan Sun eta, Yonggang Wen eta, Dacheng Tao eta, Jieping Ye eta in the paper "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications" published in 20th January 2020[5]
GANs, or generative adversarial networks, have developed a popular subject of study nowadays. The legendary Yann LeCun is GANs, according to deep learning, is the most
fascinating machine learning concept in the previous ten years There are a huge amount of papers linked to GANs accord-following Google Scholar. For illustration, there are roughly 11,800 publications on GANs published in 2018.

III. Problem Domain

Computer vision has been the key area of interest for GAN research and application. Convolutional neural networks (CNNs), a type of deep learning model, have had considerable success in the field of computer vision over the past 5 to 7 years, obtaining cutting-edge outcomes on difficult tasks like object detection and face recognition. The generation of new photos with a realistic appearance is the classic use of a GAN, which is most vividly illustrated in the example of creating photorealistic faces.

Figure 4: Human photorealistic faces

In the discipline of generative adversarial networks (GANs), a number of formulations have garnered considerable attention. These formulations deal with a variety of problems, including high-resolution picture synthesis (ProgressiveGAN), high-domain image conversion (CycleGAN), and producing images from textual descriptions. In this study, ConditionalGAN—which is pertinent to image-to-image translation tasks—is the main focus.
Let's look at some of the outcomes from ConditionalGAN. As you can see, the model has learnt how to change an image of a horse into a zebra, as well as an image from the summer to its corresponding image in the winter and vice versa.
Style GAN: Lets see can one able to recognize which image is generated and which is real from the following images:
It is true that StyleGAN is a GAN formulation that has made considerable strides in producing high-quality images, particularly ones with exceptionally high resolution. With StyleGAN, the objective is to construct a multi-layered architecture where the initial layers generate images at low resolutions (starting from 2x2), and subsequent layers

progressively enhance the resolution of the generated images. This layer-by-layer approach allows StyleGAN to create visually appealing and highly detailed images, even at resolutions as high as 1024x1024.

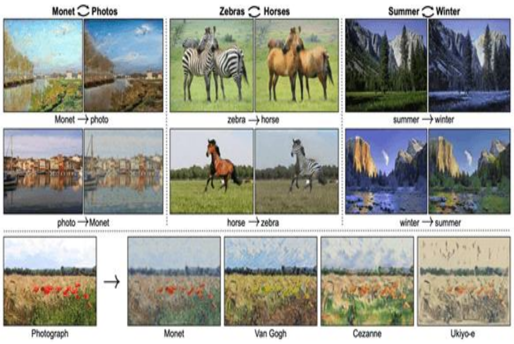

Figure 5: summer time image to the winter time

Figure 6: Fake human generation

Text-2-image:Random image generation is a skill that Generative Adversarial Networks excel at. A GAN trained on cat images, for instance, can generate random images of a cat with two ears, two eyes, and whiskers. The cat's color pattern, however, can be absolutely random. Therefore, using random images to answer commercial use cases is frequently useless. It is now quite tough to ask GAN to create an image based on our expectations.

We will discuss a GAN architecture in this part that has made substantial strides toward producing meaningful images from an explicit textual description. This GAN formulation generates an RGB image that corresponds to the description using a written description as input. For instance, if you enter "this flower has a lot of small round pink petals" it will create an image of a flower with round pink petals.

DiscoGAN: DiscoGAN recently gained enormous popularity due to its capacity to discover cross-domain relationships from unsupervised data.Cross-domain relationships come naturally to people. A human can determine the relationship between two separate domains given images of them. As an illustration, the photographs in the accompanying figure are from two separate domains, yet it is clear from a first glance at them that they are related due to the characteristics of their outer hue.

Figure 7: Disco GAN

IV. METHODOLOGY

The Pix2pix conditional generative adversarial network (cGAN) works by using an input image (also known as the source image) to produce an output image. To determine whether the generated target image is a realistic transformation of the supplied source image in the context of pix2pix, the discriminator is essential. The discriminator gains the ability to discern between realistic and unrealistic transformations by being given both the source and target images, which aids in the general training and quality assessment of the cGAN model.

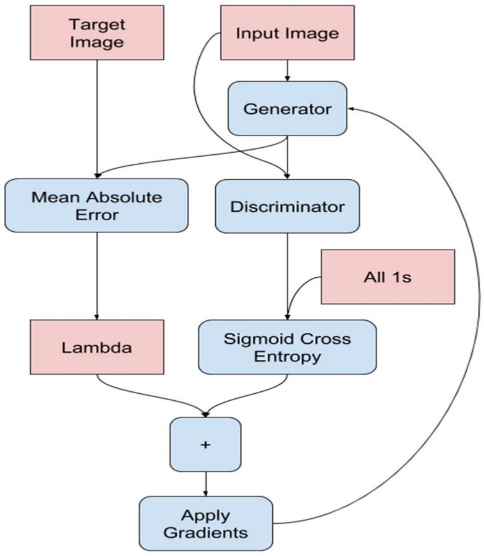

Figure 8: Generator Model Flow Chart

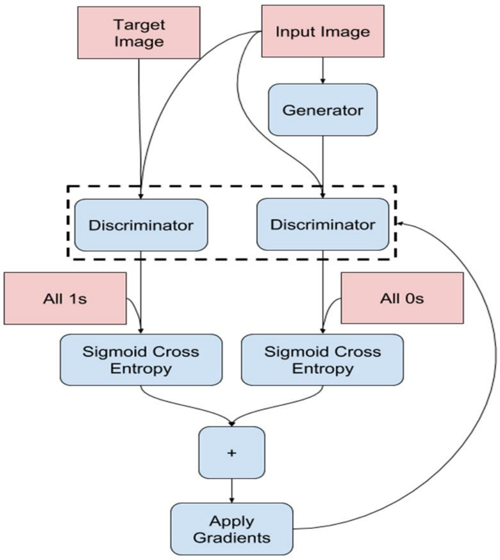

Figure 9: Discriminator Model Flow Chart

The Pix2Pix GAN drives the generator to produce accurate and realistic images in the target domain by combining adversarial loss training and L1 loss. In contrast to L1 loss, which evaluates the difference between the generated image and the desired output image and provides further advice for the generator's updates, adversarial loss forces the generator to produce plausible images by competing with the discriminator.

The Pix2Pix GAN has been successfully used for a variety of image-to-image translation tasks, including producing product shots from designs, converting maps to satellite images, and converting black-and-white photos to color. Now that we are familiar with the Pix2Pix GAN, we can develop a dataset that is intended exclusively for image-to-image translation.

Figure 10: Sample Image From the Maps Dataset Including Both Satellite and Google Maps Image.

How to develop and train the pix2pix model?
The Pix2Pix model uses both L1 (Mean Absolute error) loss which is produced by the fake map image generated by the generator and GAN loss (Binary Cross Entropy) which is normal generator loss in GAN architecture between real and fake pair of images. As per the formal official paper on GAN[1] The L1 has $\lambda 1$ value equal to 100 and GAN loss has $\lambda 2$ value equal to 1.
In the project the generator and the discriminator architecture are as follows:
Generator: The encoder-decoder architecture consists of:
encoder:

C64-C128-C256-C512-C512-C512-C512-C512

decoder:

CD512-CD512-CD512-C512-C256-C128-C64

U-Net decoder :

CD512-CD1024-CD1024-C1024-C1024-C512 -C256-C128

Discriminator:
C64-C128-C256-C512
receptive field = (output size - 1)* stride + kernel size (1)
C64: 34x34 out, 2x2 stride, 4x4 kernel
rf = (34-1)*2+4=70
C128: 16x16 rf, 2x2 stride, 4x4 kernel
rf=(16-1)*2+4=34
C256: 7x7 rf, 2x2 stride, 4x4 kernel
rf=(7-1)*2+4=16
C512: 4x4 rf, 1x1 stride, 4x4 kernel
rf = (4-1)*1+4=7
Last Layer: 1x1 out, 1x1 stride, 4x4 kernel
rf =(1-1)*1+4=4
The architecture described in the paper for image translation tasks involves a final convolutional layer followed by a 1-dimensional output mapping and a sigmoid activation function. The Torch deep learning framework provides an open-source implementation of this architecture, with detailed explanations in the main text and appendices of the paper.
In this section, the Keras deep learning framework will be used for implementation. This implementation is based on the model specified in the paper, which was designed to manage 256x256-pixel color images. It is comprised of two distinct models: the discriminator and the generator.
The discriminator's design is based on the model's effective receptive field, which determines the relationship between the model's output and the number of pixels in the input image. The generator uses the U-Net architecture, whereas the discriminator employs a PatchGAN classifier.
Using the Adam solver optimizer, both models are trained with a learning rate of 0.0002 during the training phase. The parameters for momentum are set to 1 = 0.5 and 2 = 0.999. The group size setting is set to 1.

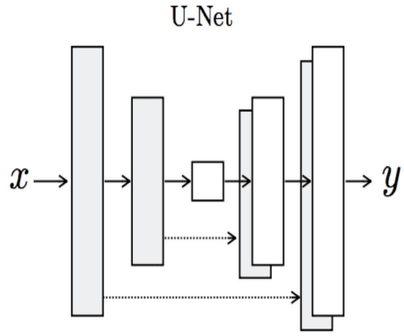

Figure 11U-Net Architecture

The 70x70 PatchGAN necessitates outputs that are distinct, easily discernible, and noise-free. The full 286x286 Image GAN generates results that are visually comparable to those of the 70x70 PatchGAN, but are of slightly inferior quality according to the FCN - score matrix (similar to the confusion matrix)[1].

Figure 12:Input Image (random 3 images out of the training dataset)

V. RESULTS

Using pix2pix image translation we have done the translation of satellite image i nto map so that it can be easy for one to find the path directly. And GANs giving it needed accuracy. So we can say GANs can also play a proper role in converting image to image compared with other available tools and technologies.
As the whole project is supposed to run on around 100 epochs but due to impossible amount of requirements and runtime requirements the model is tested on the 10 epochs only, and the results are very promising. The model is generating the output image very close to the expected output image with around of 5 hours of training. The following screenshots are the output images that are generated by the pip2pix model.

TABLE I: FCN SCORES FOR DIFFERENT LOSSES

Loss	Per-pixel acc.	Per-class acc.	Class IOU
L1	0.42	0.15	0.11
GAN	0.22	0.05	0.01
cGAN	0.57	0.22	0.16
L1+GAN	0.64	0.20	0.15
L1+cGAN	**0.66**	**0.23**	**0.17**
Ground truth	0.80	0.26	0.21

Figure 13:Output Image 1

Figure 14: Output Image 2

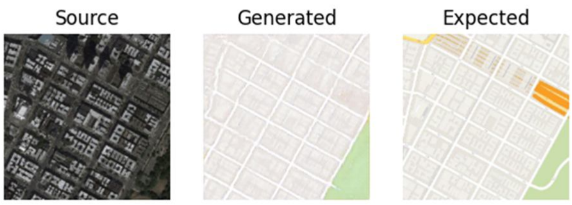

Figure 15: Output Image 3

VI. CONCLUSION

GANs are gaining popularity not only due to their ability to comprehend intricate, highly nonlinear mappings from a latent space to a data space, but also due to their ability to utilize massive amounts of un labeled image data that are inaccessible to deep replication sensation education. With regard to the nuances of GAN training, there are numerous potentials for theoretical advancements and algorithms, together with deep network power, There are several chances for brand-new applications.

VII. FUTURE SCOPE

As discussed earlier we have seen the uses of GANs. GAN is a very impressive technique in this new field, in photography we can say GAN will definitely play an important role as it can be use for aging deaging for resolution and many more.

REFERENCES

[1] Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta and A. A. Bharath, "Generative Adversarial Networks: An Overview," in IEEE Signal Processing Magazine, vol. 35, no. 1, pp. 53-65, Jan. 2018, doi: 10.1109/MSP.2017.2765202.

[2] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros : Image-to-Image Translation with Conditional Adversarial Networks. 2017 10.1109/CVPR.2017.632 arXiv:1611.07004.

[3] Y. Shi, Q. Li and X. X. Zhu, "Building Footprint Generation Using Improved Generative Adversarial Networks," in IEEE Geoscience and Remote Sensing Letters, vol. 16, no. 4, pp. 603-607, April 2019, doi: 10.1109/LGRS.2018.2878486.

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio : Generative Adversarial Networks September 2018 arXiv:1406.2661.

[5] J. Gui, Z. Sun, Y. Wen, D. Tao and J. Ye, "A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications," in IEEE Transactions on Knowledge and Data Engineering, doi: 10.1109/TKDE.2021.3130191.

[6] Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,”in Advances in Neural Information Processing Systems, 2014, pp.2672–2680.

[7] J.-Y. Zhu, P. Krähenbühl, E. Shechtman, and A. A. Efros, “Generative visual manipulation on the natural image manifold,” in European Conference on Computer Vision. Springer, 2016, pp.597–613.

[8] K. Bousmalis, N. Silberman, D. Dohan, D. Erhan, and D. Krish-nan, “Unsupervised pixel-level domain adaptation with generative adversarial networks” in IEEE Conference on Computer Vision and Pattern Recognition, 2016. arXiv:1612.05424.

[9] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks” in Proceedings of the International Conference on Computer Vision, 2017. arXiv:1703.10593.

[10] Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks” in Proceedings of the 5th International Conference on Learning Representations (ICLR) - workshop track, 2016. arXiv:1511.06434.

[11] Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta and A. A. Bharath, "Generative Adversarial Networks: An Overview," in IEEE Signal Processing Magazine, vol. 35, no. 1, pp. 53-65, Jan. 2018, doi: 10.1109/MSP.2017.2765202.

[12] E. Thirumagal, K. Saruladha, Chapter 2 - GAN models in natural language processing and image translation, Editor(s): Arun Solanki, Anand Nayyar, Mohd Naved, Generative Adversarial Networks for Image-to-Image Translation, Academic Press, 2021.

[13] X. Wang, K. Jin, K. Yu and Y. Cheng, "Asymmetric Training in RealnessGAN," in IEEE Transactions on Multimedia, doi: 10.1109/TMM.2022.3233307.

[14] V. Rakhimol, P. Uma Maheswari, Restoration of ancient temple murals using cGAN and PConv networks, Computers & Graphics, Volume 109, 2022.

[15] X. Qian, M. Zhang and F. Zhang, "Sparse GANs for Thermal Infrared Image Generation From Optical Image," in IEEE Access, vol. 8, pp. 180124-180132, 2020, doi: 10.1109/ACCESS.2020.3024576.

[16] Karim Armanious, Chenming Jiang, Marc Fischer, Thomas Küstner, Tobias Hepp, Konstantin Nikolaou, Sergios Gatidis, Bin Yang, MedGAN: Medical image translation using GANs, Computerized Medical Imaging and Graphics, Volume 79, 2020, 101684, ISSN 0895-6111.

[17] Conditional Image-to-Image Translation, Jianxin Lin, Yingce Xia, Tao Qin, Zhibo Chen, Tie-Yan Liu; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 5524-5532.

[18] High-Resolution Image Synthesis and Semantic Manipulation With Conditional GANs, Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, Bryan Catanzaro; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 8798-8807.

[19] Odena, A., Olah, C. & Shlens, J.. (2017). Conditional Image Synthesis with Auxiliary Classifier GANs. Proceedings of the 34th International Conference on Machine Learning, in Proceedings of Machine Learning Research, 70:2642-2651.

[20] Alankrita Aggarwal, Mamta Mittal, Gopi Battineni, Generative adversarial network: An overview of theory and applications, International Journal of Information Management Data Insights, Volume 1, Issue 1, 2021, 100004, ISSN 2667-0968.

[21] Deep Generative Adversarial Networks for Image-to-Image Translation: A Review by Aziz Alotaibi, College of Computers and Information Technology, Taif University, Taif 21974, Saudi Arabia Symmetry 2020, 12(10), 1705; https://doi.org/10.3390/sym12101705, Received: 3 August 2020 / Revised: 23 September 2020 / Accepted: 9 October 2020 / Published: 16 October 2020.

[22] Pix2pix Conditional Generative Adversarial Networks for Scheimpflug Camera Color-Coded Corneal Tomography Image Generation, Hazem Abdelmotaal; Ahmed A. Abdou; Ahmed F. Omar; Dalia Mohamed El-Sebaity; Khaled Abdelazeem, Translational Vision Science & Technology June 2021, Vol.10, 21. doi:https://doi.org/10.1167/tvst.10.7.21.

[23] Mavra Mehmood, Nasser Alshammari, Saad Awadh Alanazi, Asma Basharat, Fahad Ahmad, Muhammad Sajjad, Kashaf Junaid, Improved colorization and classification of intracranial tumor expanse in MRI images via hybrid scheme of Pix2Pix-cGANs and NASNet-large,Journal of King Saud University - Computer and Information Sciences,Volume 34, Issue 7,2022,Pages 4358-4374,ISSN 1319-1578.