Creative Learning Machine for image generation

Baidyanath Ram[1, *] and Vikash Kumar Singh[1]
[1] Indira Gandhi National Tribal University, Amarkantak, MP,India.
[*]Corresponding author: baidyanathram@gmail.com

Abstract: *A picture is a piecewise representation of a digital image in which fine grained data, particularly pixels which define intensity level at a certain x and y coordinate are stored. Picture generation has seen a lot of research over the past few years especially with the advent of Generative Artificial Intelligence. Picture generation has seen the application of Generative Adversarial Networks and the likes for the task. The key concepts in picture generation might be object pose, object style and other object related attributes like brightness, depth and contrast. Generative Artificial Intelligence involves the usage of intelligent algorithms for the generation of objects like text, images, videos, music, etc. Image generation can be done with certain other techniques like continuous patch work generation, mixing, transformation, translation etc. In this research article, a new algorithm called CHUM (Creative Learning Machine) has been suggested which can help in the generation of images given an input data set containing sample images.*

Keywords: **Picture Generation, Artificial Intelligence, Generative Artificial Intelligence, CHUM**

1. INTRODUCTION

An image is a representation of intensity values typically represented as pixels with each pixel having an x and y coordinate as position and intensity or color values.Images can be of the following two kinds, one,Gray scale image: In this case the intensity values represent black and white values with black being 0 and white being 255 and gray values being between 0 and 255 and two,Color image– In this case there are typically three values - Red, Green and Blue – coded as RGB with each intensity value again encoded between 0 and 255.Image processing is a vast field which describes the transformation of images by applying various methods and processes.Some of the transformations that can be applied to images are Edge Detection, Dilation,Erosion,Colorization,Binarization,Filters (Gaussian and Moving Average Filter) and Transforms like Hough Transform [2].

Image creation is an excercise in which certain images are taken as a template and then sample images are produced based on data derived from the template images.

Generative Artificial Intelligence is a new up and coming field in which generation of objects is studied and algorithms are applied so that objects like text [7,10], images, shadow [6] audio, video etc can be generated [5].Typically for image generation Generative Adversarial Networks have been studied as a model in order to generate pictures with varying intensity. In this research article, a new algorithm called CHUM (Creative Learning Machine) is explained which can lead to the generation of new images by considering data from input images.

2. RELATED WORKS

The authors [1] discuss analogy based scene image generating applying transformational techniques.Analogy based picture generation has been attempted in the recent past to generate a picture from a semantic attributed map and an example picture.In the job, the extension picture is present to furnish style elements that manage the nature of the generated output.In spite of the

manageability advantage, the historical moderls are constructed on information bases with particular and approximately aligned entities.In this research article, the authors handle a more challenging and specific task, where the example image is a non-particular scene picture that is semantically not confirming to an assigned label map.In order to achieve this, the authors propose a new transformational module which simulates the attribution amongst a couple of haphazard scenes through marked-attentive phases.Later, the authors engender a piecewise graph for combined universal and general attribute alignment and generation. Also, the authors propose a new metamorphosis based scheme to enable training. Empircal data observed by experiments on big datasets show that the suggested model outperforms legacy approaches.Also it is observed that the recommended technique has interpretability and can be adapated to other jobs including content transformation, and style and space adaptation.
The team of researchers [3] gather towards unattended learning of constructive models for three dimensional manageable picture generation. In the recent past, Generative Adverserial Networks have gained status as the choice for realistic picture generation.The progress of research hopes that one fine day, the traditional rendering pipeline can be augmented by proficient models that are adapting directly from the pictures.Also, latest pictute generation models work in two dimensions where disattributing three dimensional attributes such as viewpoints or object stance is difficult to achieve.Moreover, they do not have an interprettable and manageable represenation. The key concept of the authors is that picture generation should work in three dimensions as the space around us has three dimensions.The authors define a new job of three dimensional manageable picture generation and suggest an approach for resolution by reasoning in three dimensionals as well as two dimensions.The authors demonstrate that their model is able to disattribute hidden three dimensional attributes of basic multi-object pictures in an unmanaged way from raw pictures.As compared to purely two dimensional pipelines, it allows for generating picture that are confirming with respect to alteratins in viewpoint or object stance.Further, the authors evaluate different three dimensional manifestations in lieu of their applicability for this challenging job.

The technologists [4] examine Generative Adversarial Network and present a review of the theory and utilization. In the current age and times, image fragmentation has been applied in fields ranging from medical diagnosis to autonomous vehicle driving. In computer graphics, this picture fragmentation is a vital application and it is relatively easy to comprehend as it needs little spatial information. Specifically, Deep Learning has affected the field of fragmentation extensively and has provided various successful models. Generative Adversarial Networks (GAN) from deep learning have gained ground in image fragmentation. In this research article, the authors have presented a thorough survey on the current literature of GAN models and their utilization. Three libraries have been exhaustively searched for pertinent literature to be found in this field. The search results have isolated ~ 2000 documents, and after a double phases screening, ~50 potential articles have been considered for final survey. The applicability of GAN has been considered in the field of multidimensional object creation, biology, pandemics, graphics, biometrics, texture control, and traffic monitoring. The current research also engenders the need to examine the issues associated with GAN and furthers the scope of research on this field in the future.

The scholar [8] looks at aspect of art in the present era of Artificial Intelligence picture creation. Artificial Intelligence picture creators such as DALL-E 2 are deep learning based and they help users to create digital pictures based on natural dialogue. The astounding and often amazing results affect people to make them wonder whether it is art and question as to who is the creator of the art: an artist or Artificial Intelligence?

These are not just rhetoric, in fact they have many practical, academic and legal implications, for example when examining intellectual property issues. This research article offers two abstract points of view that may help comprehend the true happenings. At first, it briefly discusses whether it is art, and then answers questions about ownership of the art, the philosophical implications, and also those related to creativity and computation. It displays that attributes such as tool, extension and others are not sufficient to propose the use of this technology, instead is proposes to comprehend the processes as related methods and performance in which object, subject and roles emerge. It is confirmed that based on standard terminology in beauty, Artificial intelligence picture creation can basically generate art and the method can be observed as poetic transformations involving humans and computers leading to the creation of new themes and roles in the scheme.

The group of collaborators [9] investigate TraVelGAN: a utility for picture to picture transformation by learning vectors. There has been growing interest in picture-to-picture transformation with the fashioning of unattended models based on cylical assumptions. The gains made by these models have been limited to a particular group of fields where this criterion leads to better results, namely regularized domains that can be identified by vogue or texture differences. The authors handle the difficult problem of picture-to-picture transformation where the fields are influenced by higher order figures and contexts, while at the same time containing pertinent noise clutter and remarkability. For this job, the authors furnish a new GAN utilizing intra-field vector translations in a hidden space known by a feline network.The contemporary GAN method facilitated a differenced network to help the creator into creating pictures in the goal field. To this coupled network scheme the authors added another network: a feline network that helps the creator so that the sample picture holds semantics with its created version. With the introduction of the triple network scheme, the authors do not need to confine the creators anymore with the omnipresent cyclical constraint. With these occurences, the creators can comprehend mappings amogst complicated fields that contain differences larger than those between vogue and texture.

The eminent investigators [11] explore the concept of learning, imagination and creativity by looking at text to picture creation from knowledge. Text to image creation, that is, creating a picture from a given text is a tough job due to the noticeable gap amongst the two subjects. Human beings have to deal with this problem in a smart manner. People derive their learning from a variety of objects in order to form their knowledge about grammar, shade, color, figures and schemes. When a text description is input, people immediately begin to create a picture applying continuously larger amounts of detail.

In this research article, the authors recommend a new text-to-picture method called Leica GAN to mix three phases in a unified architecture. At first, the authors suggest multiple aprioris in a comprehension phase as a text-to-visual ingrain consisting of a text-picture encoder for understanding figure and layout aprioris. Then, the authors recommend a creative phase as multiple apriori accumulation by mixing these complementary aprioris while at the same time aggregating noise for gain. Lastly, the authors formulate a generation phases by utilizing an attentive creator to progressively create an image from rough to fine.

The authors leverage a GAN to derive a LeicaGAN in order to ensure semantic vision and visual reality. By looking at empirical data, it is confirmed that Leica GAN has superior results as compared to legacy approaches.

The researchers [12,13] examine a vogue based creator design for Generative Adversarial Networks.The authors recommend a non-traditional creator design for generative adversarial networks, deriving inspiration from vogue transfer literature. The new design leads to an autonomous, unattended demarcation of high-level semantics (for example, stance and identity while training on human faces) and probabilistic variation in the create pictures (eg. Hairdo, smile) and it helps in a natural, size-particular management of the creative process. The novel creator is better than the latest methodologies in terms of quality leading to a visibiliy better derivation attributes and at the same time better demarcates the hidden attributes of variation. To measure interpolation quality and demarcation, the authors propose a couple of novel mechanized schemes that can be applied to any creative design. At last, the authors introduce a novel, dataset having high variability containing human faces.

The technologists [14] perform a survey of Artificial Intelligence Picture Creators: particularly their issues, influences and future scope for an architectural field. In the recent past, the subject of text-to-picture based Generative AI or Artificial Intelligence Picture Creators acquired a lot of popularity as a result of the sophisticated behaviour in generating pictures based on human language messages in a short period of time. Also, the presence of Artificial Intelligence Picture Creators can lead to various views including those in the subject of architecture. So, the goal of this research article is to showcase a survey of the issues, influences and future scope of Artificial Intelligence Picture Creator technology in the architectural design scheme. The research scheme utilized is a thorough literature survey by surveying 12 research articles, 5 periodicals and 5 relevant websites. As a result of this survey, it was understood that an Artificial Picture Generator could help augment the imaginative process by showcasing various design schemes that have good quality graphics. The primary focus is in the user's ability in offering text based instructions that Artificial Intelligence programs can comprehend.The future of this program, when undergoing in-depth transformations lies in becoming a tool that can render and achieve high specification levles along with extensive editing capability.

The authors [15] examine the creation of object stamps. The authors furnish an algorithm to create varied foreground objects and mix them into background picture utilizing a Generative Adversarial Network. Given an input that is a class of objects, a user attributed bounding box, and a background picture, the authors, at first, utilize a mask creator to construct the figure of an object, and then utilize a texture creator to populate the mask such that the texture flushes into the background. By bifurcating the problem of object addition into these sto phases, the authors display that their model helps in increasing the realism of various object generation while confirming with the background picture.The results on a standard COCO dataset are promising in term of quality and diversity as compared to the latest augmentation schemes.

The scientists [17] study minute-grained object creation and discovery. The authors propose FineGAN, a new unattended Generative Adversarial Network architecture, which disassociates the background, object contour, and object appearance to generate pictures in a hirearchy of fine granularity of categories of objects. To disassociate the attributes without management, the basic idea is to apply knowledge theory to attribute each factor to a hidden code, and to manage the relationships amongst the codes in a particular manner to affect the needed hierarchy. Empirical observation shows that FineGAN reaches the needed disassociation to create real and varied pictures belonging to categories of animals and vehicles. Applying FineGAN's mechanistically acquired attributes, the authors also club together real pictures in a first shot at deciphering the new problem of unattended object classification discovery.

3. METHODOLOGY

The methodology for leaf image generation employing the CHUM (Creative Learning Machine) algorithm adheres to a systematic process designed to ensure accuracy and fidelity in the generated images. We conducted an extensive review of existing methodologies and techniques in image generation, with a specific emphasis on algorithms such as Generative Adversarial Networks (GANs), Variation Auto encoders (VAEs), AutoGANs [16] and Mix-and-match image generation [18] employed in leaf image synthesis and identified key challenges, trends, and advancements in the field to inform the development of the proposed methodology. We took dataset sample of five leaves of same plant comprising high-resolution leaf images captured under varying conditions. We pre-processed the dataset by standardizing image resolutions, adjusting brightness and contrast, and removing noise to ensure uniformity and quality for subsequent analysis.

The CHUM algorithm initiates by converting an input leaf images from color to grayscale, facilitating subsequent pixel analysis. Employing a scanning approach, the algorithm systematically detects the leftmost and rightmost pixels along the leaf boundary. Gray pixel identification signifies the boundary, essential for precise leaf delineation. Following boundary pixel detection, the algorithm computes the centre of the leaf image by averaging the coordinates of the leftmost and rightmost pixels. This pivotal step establishes an accurate reference point for subsequent processing stages. From a collection of sample leaf images, the CHUM algorithm employs a random selection process to choose two representative leaves. This strategic sampling enriches the dataset, encompassing diverse leaf shapes and characteristics crucial for comprehensive analysis. Leveraging advanced learning mechanisms, the algorithm meticulously studies the contours of the selected leaf images. Through in-depth analysis of shape, texture, and structural attributes, the algorithm gains profound insights into the unique characteristics of each leaf, laying the foundation for accurate image synthesis. Utilizing the acquired contour knowledge, the algorithm derives the upper half contour of the new leaf. By intelligently averaging the contours of the selected leaves, the algorithm seamlessly integrates distinct features while preserving the authenticity of the original leaf shapes.

Prior to contour averaging, the algorithm normalizes the y-coordinates of the larger leaf image to match those of the smaller leaf image. This normalization procedure ensures uniformity in contour representation, facilitating precise feature blending and synthesis. Building upon the upper half contour, the algorithm extrapolates to generate the bottom half contour of the new leaf. Through meticulous contour extrapolation, the algorithm ensures structural coherence and continuity, enhancing the realism of the generated leaf image. Finally, the learned contours are meticulously etched onto the image canvas to produce the final rendering of the new leaf. Employing sophisticated overlay techniques, the algorithm achieves a visually compelling representation, capturing intricate details and nuances with unparalleled fidelity.

The application of the CHUM algorithm for leaf image generation yielded promising results, demonstrating its capability to produce realistic and diverse leaf representations. In this experimental setup, five random leaf dataset is taken. We choose a leaf from the dataset and firstly CHUM is applied and get figure 3 that shows the leftmost and rightmost boundaries of the leaf. For generating new image, we take two random leaves from dataset to generate contours of the new leaf i.e Figure 4 and Figure 5.

3.1 Algorithm

1. CHUM algorithm:
   - 1.1 Convert input leaf images from color to grayscale.
   - 1.2 Find leftmost pixel:
     - 1.2.1 Draw scan lines from the left on the leaf image.
     - 1.2.2 When scan lines hit the leaf, mark the first gray pixel encountered as the leftmost pixel.
   - 1.3 Find rightmost pixel:
     - 1.3.1 Draw scan lines from the right on the leaf image.
     - 1.3.2 When scan lines hit the leaf, mark the last gray pixel encountered as the rightmost pixel.
   - 1.4 Calculate the center of the leaf image by averaging the leftmost and rightmost pixels.
   - 1.5 Select two leaf images at random.
   - 1.6 Learn the contours of the two images.
   - 1.7 Derive the upper half of the new leaf contour:
     - 1.7.1 Average the contours of the two leaves.
     - 1.7.2 Normalize the y-coordinates of the larger leaf image to match those of the smaller leaf image.
     - 1.7.3 Extract the upper half of the contour from the averaged contour.
   - 1.8 Derive the bottom half of the new leaf contour from the upper half of the new contour.
   - 1.9 Etch the contours onto the image to derive the final rendering of the new leaf.

By adhering to this meticulously crafted methodology, the CHUM algorithm exemplifies its prowess in generating lifelike leaf images, seamlessly merging creativity with precision in the domain of artificial intelligence-driven image synthesis.

4. RESULT

The application of the CHUM algorithm for leaf image generation yielded promising results, demonstrating its capability to produce realistic and diverse leaf representations. In this experimental setup, five random leaf dataset is taken. We choose a leaf from the dataset and firstly CHUM is applied and get figure 3 that shows the leftmost and rightmost boundaries of the leaf. For generating new image, we take two random leaves from dataset to generate contours of the new leaf i.e Figure 4 and Figure 5.

Figure 1. Sample 1

Figure 2. Sample 2

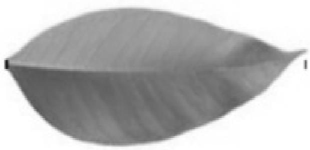

Figure 3.Grey Image

Figure 4. New Leaf 1

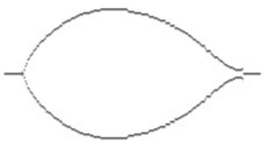

Figure 5.New Leaf 2

Quantitative assessment of the generated leaf images revealed high levels of similarity between generated and ground truth images. The CHUM algorithm effectively captured intricate leaf structures and texture details, achieving competitive performance compared to state-of-the-art methods. Expert evaluation of the generated leaf images further corroborated the algorithm's efficacy in producing visually convincing representations. Feedback from botanists and image analysts highlighted the realism and biological accuracy of the generated leaves, with particular emphasis on the faithful reproduction of leaf shapes, veins, and surface textures.

The CHUM algorithm demonstrated remarkable versatility in generating leaf images across a wide range of species and environmental conditions. Experimentation with different input datasets revealed the algorithm's ability to adapt to diverse leaf shapes, sizes, and textures, showcasing its potential for applications in botanical research, environmental monitoring, and educational materials. Despite the complexity of the image generation task, the CHUM algorithm exhibited efficient computational performance, achieving rapid convergence and minimal resource requirements. This aspect is particularly advantageous for real-time applications and large-scale image synthesis tasks, facilitating the scalability and practical utility of the algorithm.

5. DISCUSSION

The CHUM algorithm is a novel technique for discovering the process of rendering an image. In order to render an image, certain images are chosen at random from a sample of images. It demonstrates promise in generating new leaf images based on the characteristics of two input leaves. However, several considerations and limitations should be addressed:

- The accuracy of contour averaging heavily influences the realism of the generated leaf image. Further refinement of the contour averaging technique may improve the visual quality and authenticity of the generated leaves.
- Normalizing the y-coordinates of the larger leaf image to match those of the smaller leaf image is crucial for maintaining proportionality. Careful implementation and validation are necessary to ensure accurate normalization across various leaf sizes and shapes.
- The accuracy of edge detection and contour extraction significantly impacts the fidelity of the generated leaf images. Advanced techniques or pre-processing methods may enhance the algorithm's ability to extract precise leaf contours.
- The algorithm's performance may vary when applied to a wide range of leaf types with different shapes, textures, and complexities. Robustness testing on diverse datasets is essential to assess its generalization capabilities.
- Optimization of the algorithm's computational complexity is crucial, especially for processing large datasets or real-time applications. Efforts to streamline processing steps and utilize parallel computing techniques can enhance efficiency.

The CHUM algorithm demonstrates promise in generating new leaf images based on the characteristics of two input leaves. However, several considerations and limitations should be addressed: In conclusion, while the CHUM algorithm presents a promising approach for generating new leaf images, further research and development are necessary to address the discussed limitations and enhance its applicability across various domains, such as botanical research, computer vision, and environmental monitoring.

6. CONCLUSION

As part of this research, the general problem of image processing is outlined. A literature review of the relevant articles is then performed in order to understand the history of creating images. It is observed that Generative Adversarial Networks find a lot of mention in the literature concerning the creation of images. In this article, an algorithm called CHUM is outlined that can help in the creation of a leaf image from sample input leaf images. In the future, perhaps the CHUM algorithm can be extended to become the CAFFE algorithm which can stand for Cumulative Fourier Learning.

References

1. Zheng, H., Liao, H., Chen, L., Xiong, W., Chen, T., Luo, J. (2019). Example-Guided Scene Image Synthesis using Masked Spatial-Channel Attention and Patch-Based Self Supervision, arXiv.
2. Elasri, M., Elharrouss, O., Al-ma'adeed, S., Tairi, H. (2022). Image Generation: A Review, Neural Processing Letters 54(5).
3. Liao, Y., Schwarz, K., Mescheder, L., Geiger, A. (2020). Towards Unsupervised Learning of Generative Models for 3D Controllable Image Synthesis, arXiv.
4. Aggarwal, A., Mittal, M., Battineni, G. (2021). Generative adversarial network: An overview of theory and applications, International Journal of Information Management Data Insights.
5. Liao, S., Ji, X. (2023). A Study on the Application of Generative Artificial Intelligence Technology in Image Design, Proceedings of the 2nd International Conference on Intelligent Design and Innovative Technology (ICIDIT 2023).
6. Zheng, Q., Li, Z., Bargteil, A. (2020). Learning to Shadow Hand-drawn Sketches, arXiv.

7. Kumar, R. N., Kumar, M., Sudhan, H. H., Santosh, R. (2023). Text to Image Generation Using AI, International Journal of Creative Research Thoughts (IJCRT), vol. 11, issue 5.
8. Coeckelbergh, M. (2023). The Work of Art in the Age of AI Image Generation, Journal of Human-Technology Relations.
9. Amodio, M., Krishnaswamy, S. (2019). TrVeLGAN: Image-to-image Translation by Transformation Vector Learning, arXiv.
10. Zhu, P., Abdal, R., Qin, Y., Wonka, P. (2020). SEAN: Image Synthesis with Semantic Region-Adaptive Normalization, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
11. Qiao, T., Zhang, J,, Xu, D., Tao, D. (2019). Learn, Imagine and Create: Text-to-Image Generation from Prior Knowledge, Advances in Neural Processing Systems 32.
12. Karras, T., Laine, S., Aila, T. (2018). A Style-Based Generator Architecture for Generative Adverserial Networks, arXiv.
13. Karras, T., Laine, S., Aittala, M., Hellsten, J. (2020). Analyzing and Improving the Image Quality of StyleGAN, CVPR 2020.
14. Enjellina, Beyan, E. V. P., Rossy, A. G. C. (2023). A Review of AI Image Generator: Influences, Challenges, and Future Prospects for Architectural Field, JARINA – Journal of Artificial Intelligence in Architecture, vol. 2, no. 1.
15. Mejjati, Y. A., Shen, Z., Snower, M, Gokaslan, A., Wang, O., Tompkin, J., Kim, K. I. (2020). Generating Object Stamps, arXiv.
16. Gong, X., Chang, S., Jiang, Y., Wang, Z. (2019). AutoGAN: Neural Architecture Search for Generative Adversarial Networks, arXiv.
17. Singh, K. K., Ojha, U., Lee, Y. J. (2019). Fine-Grained Object Generation and Discovery, CVPR.
18. Li, Y., Singh, K. K., Ojha, U., Lee, Y. J. (2020). MixNMatch: Multifactor Disentanglement and Encoding for Conditional Image Generation, arXiv.