

MẠNG SÁNG TẠO ĐỐI NGHỊCH VÀ ỨNG DỤNG XÂY DỰNG NHÂN VẬT TRONG THỰC TẠI ẢO

Đỗ Thị Bắc*, Lê Sơn Thái, Mã Văn Thu, Đỗ Thị Chi, Hà Mỹ Trinh
Trường Đại học Công nghệ thông tin và Truyền thông – ĐH Thái Nguyên

TÓM TẮT

Bài báo tìm hiểu mạng sáng tạo đối nghịch (GAN) và ứng dụng sinh vật liệu tự động cho các nhân vật dạng con người trong thực tại ảo. Một tập dữ liệu thực của vật liệu do các nhà thiết kế 3D tạo ra được sử dụng để huấn luyện hai thành phần đối nghịch nhau trong mạng nơron là thành phần sinh dữ liệu và phân biệt dữ liệu. Kết quả thực nghiệm cho thấy mạng GAN cho phép sinh vật liệu tự động cho mô hình 3D. Vật liệu do GAN sinh ra đảm bảo các yêu cầu về mặt kỹ thuật khi trải lưới và hình ảnh để sử dụng cho mô hình nhân vật 3D. Đây là hướng nghiên cứu, ứng dụng trí tuệ nhân tạo nhiều tiềm năng trong quá trình sản xuất dữ liệu đa phương tiện nói chung và mô hình 3D nói riêng.

Từ khóa: Đồ họa máy tính; mạng sáng tạo đối nghịch; GAN; thực tại ảo; mô hình 3D.

Ngày nhận bài: 24/8/2020; Ngày hoàn thiện: 30/11/2020; Ngày đăng: 30/11/2020

GENERATIVE ADVERSARIAL NETWORKS AND APPLICATION FOR BUILDING CHARACTERS IN VIRTUAL REALITY

Do Thi Bac*, Le Son Thai, Ma Van Thu, Do Thi Chi, Ha My Trinh
TNU – University of Information and Communication Technology

ABSTRACT

The article explored Generative Adversarial Networks (GAN) neural networks and application of automated material for human characters in virtual reality. A real data set of materials created by 3D designers was used to train two opposing elements in a neural network, which are data generating and data differentiating. Experimental results show that the GAN network allows automatic material generation for 3D models. GAN generated materials meet the technical requirements of meshing and imagery for use in 3D character modeling. This is the direction of research and application of artificial intelligence with great potentials in the production of multimedia data in general and 3D models in particular.

Keywords: Computer graphics; generative adversarial networks; GAN; virtual reality; model 3D.

Received: 24/8/2020; Revised: 30/11/2020; Published: 30/11/2020

* Corresponding author. Email: dtbac@ictu.edu.vn

1. Giới thiệu

Với sự phát triển không ngừng của công nghệ, trí tuệ nhân tạo ngày càng được ứng dụng rộng rãi trong nhiều lĩnh vực khoa học, đời sống, kinh tế và xã hội. Thậm chí, ngay cả những lĩnh vực mà tưởng như chỉ dành riêng cho con người như vẽ tranh hay sáng tạo nghệ thuật cũng có sự góp mặt của máy tính. Mạng sáng tạo đối nghịch gọi tắt là GAN [1] mạng noron có khả năng sinh ra dữ liệu mới. Với khả năng này, GAN được nhiều nhà nghiên cứu quan tâm, phát triển và đã có nhiều ứng dụng khác nhau trong cuộc sống. Đặc biệt với các dữ liệu dạng đa phương tiện, đây là một mạng noron có nhiều ưu thế và cho kết quả tốt khi áp dụng sinh dữ liệu tự động.

Đối với ảnh GAN cho phép tăng chất lượng độ phân giải của hình ảnh. Với SRGAN [2] một mạng noron cải tiến từ GAN, cho phép nâng cao chất lượng hình ảnh lên tới bốn lần. Tiếp đó, ESRGAN [3] được đề xuất giúp tăng độ phân giải đồng thời có hiệu suất cao hơn SRGAN. Một cách tiếp cận khác, các nhà nghiên cứu ứng dụng cho các quá trình chuyển đổi hình ảnh.



Hình 1. Ứng dụng chuyển đổi hình ảnh

Hình 1 là việc chuyển một chú ngựa thường thành ngựa vằn với việc áp dụng CycleGAN [4]. Sự phát triển này có thể ứng dụng trong các vấn đề về xử lý video và kỹ xảo hình ảnh. Không chỉ dừng lại ở đó, các hình ảnh chuyển đổi có thể được thực hiện dựa trên việc học tập các phong cách nghệ thuật khác nhau và chuyển một ảnh chụp thành giống như tranh vẽ. Một ứng dụng nổi bật khác của GAN là việc tổng hợp hình ảnh khuôn mặt con người. Khi đó, các ảnh chân dung được tạo ra từ máy tính không phải là ảnh thực nhưng có chất lượng và đặc điểm gần như tương đồng với ảnh chụp.



Hình 2. Tổng hợp khuôn mặt

Hình 2 là ảnh khuôn mặt con người được tổng hợp bằng một cải tiến từ GAN [5] của các nhà phát triển. Hình ảnh tổng hợp được hoàn toàn có thể đánh lừa các hệ thống nhận dạng mặt người.

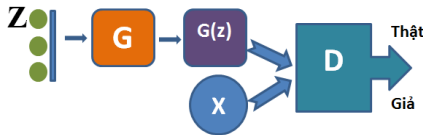
Y. Cui and W. Wang sử dụng mạng GAN cho việc tô màu video [6]. GAN được huấn luyện từ tập ảnh màu và tiến hành tô màu cho các video đen trắng hoặc các video đa cấp xám được quay từ thế kỷ trước. Kết quả cho thấy, mạng noron nhân tạo có khả năng tô màu tương đối tốt cho video. Tuy nhiên, nhiều vấn đề về xử lý ánh sáng, độ sâu của hình ảnh cũng như sự ổn định về màu sắc của video chưa được như mong muốn. Một số nhà nghiên cứu khác [7] áp dụng GAN cho việc nâng cao chất lượng cho video dựa trên các thước đo khác nhau, từ đó thu hút người xem hơn.

Gần đây, một số nhà nghiên cứu [8] áp dụng GAN khi xây dựng mô hình 3D từ ảnh 2D. Qua đó, mạng noron sinh các dữ liệu mô hình ba chiều từ tập các ảnh đầu vào. Mô hình sinh ra đảm bảo hình khối thực tế của đối tượng. Tuy nhiên, tồn tại một khoảng cách nhất định so với mô hình được thiết kế bởi con người hay máy quét. Nhưng kết quả này cũng cho thấy tiềm năng của việc ứng dụng GAN trong xây dựng, thiết kế mô hình 3D.

Nhận thấy tiềm năng của của GAN trong quá trình sinh dữ liệu, các nghiên cứu về mạng noron này thực nghiệm với nhiều loại dữ liệu khác nhau và các ứng dụng của chúng cũng tương đối phong phú. Nổi tiếp các nghiên cứu này, chúng tôi tìm hiểu về GAN và ứng dụng khi xây dựng nhân vật trong thực tại ảo. Điều này giúp rút ngắn thời gian và công sức khi xây dựng dữ liệu 3D. Cụ thể, phần hai trình bày các lý thuyết cơ bản về xây dựng mạng noron GAN và một số mô hình phát triển của nó. Phần ba là ứng dụng sinh vật liệu cho đối tượng dạng con người trong thực tại ảo. Cuối cùng là những đánh giá và kết luận của bài báo.

2. Một số nghiên cứu về mạng sáng tạo đối nghịch

Mạng sáng tạo đối nghịch (GAN) có khả năng tự động phát hiện và học các mẫu phổ biến từ dữ liệu đầu vào. Từ đó, mạng nơron này sáng tạo ra những dữ liệu mới tương tự như dữ liệu được học. Một cách tổng quan, GAN thường gồm hai thành phần: một thành phần sinh dữ liệu để sáng tạo dữ liệu mới và một thành phần để đánh giá dữ liệu mới sinh ra là thật hay giả. Hai thành phần được huấn luyện song song, đối nghịch nhau. Quá trình huấn luyện thường dừng lại khi thành phần đánh giá dữ liệu không phân biệt được đâu là dữ liệu thật và đâu là dữ liệu giả.



Hình 3. Mô hình tổng quan của GAN

Trong hình 3, Z chính là tập các Vector đầu vào có thể được khởi tạo một cách ngẫu nhiên. G chính là thành phần sinh dữ liệu. Dựa trên các Vector đầu vào Z sẽ tạo ra tập dữ liệu tổng hợp $G(z)$ là các dữ liệu giả. X là tập dữ liệu thật có thể được thu nhận từ thế giới thực hoặc do con người sáng tạo bằng công cụ thiết kế. Thành phần đánh giá, phân biệt dữ liệu D để đánh giá dữ liệu là thật hay giả. Khi D không thể phân biệt đâu là dữ liệu của $G(z)$ và đâu là dữ liệu của X , trạng thái tối ưu của mạng GAN sẽ đạt được và dừng việc huấn luyện. Cấu trúc này của GAN có điểm tương đồng với việc xây dựng các trò chơi đối kháng dạng min-max trong lý thuyết trò chơi với hai người chơi và người chơi nào cũng sẽ cố gắng để dành chiến thắng bằng cách tăng tối đa hàm mục tiêu của mình và hạn chế hàm mục tiêu của đối thủ. Việc tối ưu hóa GAN với hai thành phần là G tạo ra dữ liệu và D phân biệt dữ liệu có thể dựa trên hàm mục tiêu đối nghịch [1] như sau:

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

Trong công thức (1) ở trên: G, D là các thành phần của GAN như đã trình bày ở trên. Ký hiệu x, z là một phần tử thuộc tập X, Z . E biểu thị giá trị mong đợi, hàm p biểu thị xác suất của tham số kèm theo. D cho giá trị từ 0 đến 1 thể hiện độ thật của dữ liệu (0 tức là giả và 1 tức là thật). Khi huấn luyện thành phần phân biệt D có mục tiêu tối ưu theo công thức (2):

$$\max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (2)$$

Thành phần sinh dữ liệu G có mục tiêu tối ưu theo công thức (3):

$$\min_G V(D, G) = E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (3)$$

Sự phát triển của mạng GAN thể hiện bởi sự đối nghịch của hai thành phần này. Vì vậy, có thể gọi GAN là mạng nơron “sáng tạo đối nghịch”. Ở một cách áp dụng khác, thành phần sinh dữ liệu G có thể được huấn luyện nhằm mục tiêu hàm $\log D(G(z))$ đạt giá trị lớn nhất thay cho hàm $\log(1 - D(G(z)))$ đạt giá trị nhỏ nhất [1]. Quá trình huấn luyện cập nhật các tham số trong G và D được thực hiện một cách song song với nhau. Khi hai thành phần sinh G và phân biệt D đạt tới trạng thái cân bằng ($D(G(z)) = 0.5$), quá trình huấn luyện kết thúc.

Đối với thành phần phân biệt D , mục tiêu là cố gắng phân biệt dữ liệu thực và giả. Do đó, nó có thể sử dụng bất kỳ kiến trúc mạng nơron nào phù hợp với loại dữ liệu mà nó đang phân loại. Trong quá trình học tập, thành phần này kết nối tới hai hàm phản hồi. Một hàm trong đó phản hồi trực tiếp tới chính nó để nâng cao khả năng phân biệt dữ liệu. Hàm còn lại kết nối tới thành phần sinh dữ liệu G . Thành phần sinh dữ liệu dựa trên các phản hồi để nâng cao độ thật của dữ liệu được sinh ra. Các phản hồi này có thể dựa trên các hàm mất mát dữ liệu khác nhau và tùy theo loại dữ liệu khác nhau mà các hàm mất mát này cũng được lựa chọn khác nhau. Vì hai thành phần G và D có tính riêng biệt nhất định vì vậy GAN phải đảm bảo kết hợp huấn luyện được cả hai thành phần này. Quá trình huấn luyện luân phiên thường được sử dụng nhằm giải quyết vấn đề này.

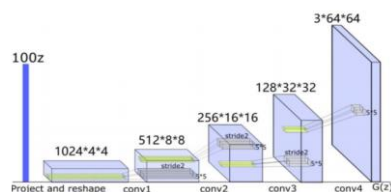
Từ khi hình thành, các nhà nghiên cứu đã có nhiều đột phá khác nhau trong việc phát triển GAN. Ban đầu, Multi-Layer Perceptron (MLP) được sử dụng trong kiến trúc của GAN khi xây dựng thành phần sinh và phân biệt dữ liệu. Tiếp đó, Convolutional Neural Network (CNN) [9] mô hình học tập có giám sát với những ưu thế lớn trong xử lý hình ảnh được đề xuất và sử dụng nhiều trong các ứng dụng của GAN.

Conditional GAN (CGAN) [10] cho phép cả thành phần sinh và thành phần phân biệt làm việc với các lớp được gán nhãn. Với các nhãn y được gán cho dữ liệu, mạng học tập và sinh các dữ liệu theo các lớp khác nhau. Khi đó, hàm mục tiêu đối nghịch thay đổi như sau:

$$\min_G \max_D V(D, G) = E_{(x, y) \sim p_{xy}} [\log D(x, y)] + E_{z \sim p_z, y \sim p_y} [\log(1 - D(G(z, y), y))] \quad (4)$$

CycleGAN [4] cho phép kết hợp đầu vào từ các miền khác nhau. Ví dụ, với một ứng dụng sáng tạo hình ảnh X theo một phong cách Y , ta có hai miền đầu vào phân biệt. Như vậy, tồn tại hai ánh xạ từ miền $X \rightarrow Y$ và ngược lại. Tương ứng với đó là hai thành phần sinh và hai thành phần phân biệt với mỗi ánh xạ. Hàm mục tiêu đối nghịch của CycleGAN phụ thuộc vào sự kết hợp của các hàm mất mát trong huấn luyện [4].

Sự kết hợp của GAN và mạng CNN [9] với hai thành phần sinh dữ liệu và phân biệt dữ liệu là hai mạng CNN ngược nhau được gọi là DCGAN [11]. Đây là một cải tiến của GAN được sử dụng nhiều trong quá trình xử lý hình ảnh do kế thừa các ưu thế của mạng CNN.



Hình 4. Kiến trúc thành phần sinh dữ liệu trong DCGAN [11]

Hình 4 mô tả kiến trúc thành phần sinh dữ liệu của DCGAN. Kiến trúc của thành phần

phân biệt dữ liệu được thiết kế ngược lại với thành phần sinh dữ liệu.

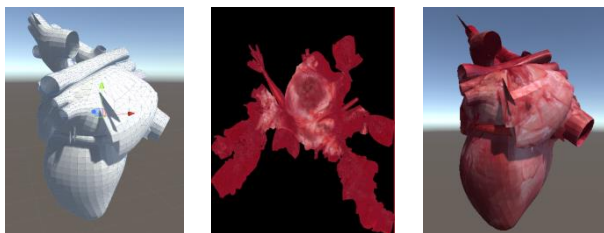
GAN là một lĩnh vực nghiên cứu nhiều tiềm năng và thu hút nhiều sự chú ý. Nhiều đề xuất, cải tiến khác nhau cho mạng GAN đã được các nhà nghiên cứu đưa ra trong thời gian gần đây. Tùy theo mục tiêu và yêu cầu khác nhau mà kiến trúc chi tiết bên trong có thể thay đổi và hình thành các dạng khác nhau của GAN.

3. Generative Adversarial Networks áp dụng trong xây dựng vật liệu cho mô hình nhân vật 3D

Như đã trình bày ở phần trên của bài báo, có nhiều ứng dụng khác nhau của GAN trong quá trình sinh dữ liệu tự động, đặc biệt là các dữ liệu dạng đa phương tiện. Loại dữ liệu này mang nhiều đặc trưng sáng tạo của con người và thường đòi hỏi nhiều thời gian và công sức trong quá trình tạo ra chúng. Các phần mềm thiết kế và chỉnh sửa thường được áp dụng trong quá trình sản xuất dữ liệu đa phương tiện (photoshop, adobe illustrator, premier, after effect, 3DsMax, Maya...). Nhược điểm chung của cách thức sản xuất này là đòi hỏi nhiều chi phí về thời gian và con người. Đây là một trong những khó khăn cơ bản khi nghiên cứu về một hệ thống tạo dữ liệu đa phương tiện tự động.

Như đã biết, mô hình 3D là một cấu trúc dữ liệu mô tả hình thái ba chiều của một đối tượng. Hiện nay, để tạo ra một mô hình 3D có nhiều cách khác nhau. Chúng thường được tạo ra từ các phần mềm thiết kế: 3Ds max, maya v.v.. thông qua các nhà thiết kế, hoặc từ các máy quét ba chiều.

Một mô hình 3D gồm có hai thành phần cơ bản: một là tập các đỉnh và tập các mặt hay thường được gọi là lưới của mô hình, hai là tập UV. Trong đó, tập UV thường kết hợp với một ảnh chất liệu để tạo ra hình ảnh của mô hình 3D với bề mặt giống với thực tế.

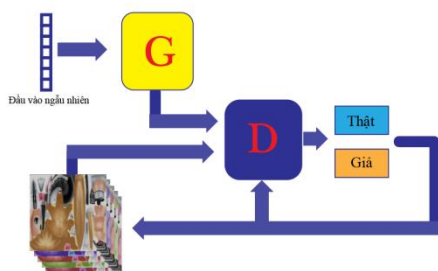


Hình 5. Mô hình 3D tim người

Hình 5 mô tả cấu trúc của một mô hình ba chiều. Ảnh bên trái là lưới của mô hình, ảnh giữa là vật liệu của mô hình, ảnh bên phải là hình ảnh ba chiều của mô hình khi kết hợp lưới và vật liệu. Có nhiều bài toán khác nhau xoay quanh quá trình xây dựng mô hình 3D với mục tiêu chính là tối ưu về thẩm mỹ, thời gian xây dựng và kết cấu mô hình. Như đã trình bày, để giảm thời gian xây dựng nhân vật chúng ta có hai hướng chính là tối ưu thời gian xây dựng lưới và tối ưu thời gian xây dựng vật liệu. Đây là hai thành phần chính của mô hình 3D. Đối với quá trình xây dựng lưới các máy quét giúp giảm thời gian thiết kế nhưng mô hình 3D thu được có số lượng lưới lớn cần tối ưu lại [12] trước khi sử dụng. Một cách tiếp cận khác, các nhà nghiên cứu dùng chính GAN để tạo lưới cho mô hình 3D từ ảnh 2D nhưng kết quả còn nhiều hạn chế.



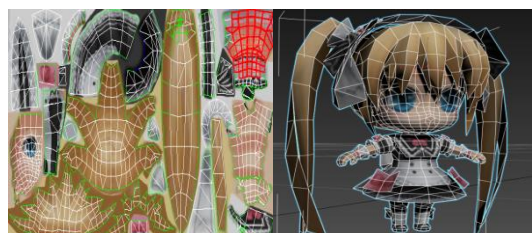
Hình 6. Mô hình 3D từ ảnh 2D



Hình 7. Kiến trúc GAN sinh vật liệu tự động

Hình 6 là kết quả của áp dụng GAN khi xây dựng mô hình 3D từ ảnh 2D [8]. Chúng tôi tiếp cận việc giảm thời gian, chi phí xây dựng mô hình thông qua việc sinh các vật liệu 3D một cách tự động.

Hình 7 mô tả kiến trúc của GAN khi sinh vật liệu tự động cho đối tượng 3D. Tập dữ liệu thực X là tập các vật liệu được xây dựng từ các nhà thiết kế 3D. Bằng cách thay đổi các vật liệu khác nhau cho cùng một lưới mô hình chúng ta thu được các nhân vật khác nhau. Thành phần sinh dữ liệu G có nhiệm vụ sinh các vật liệu mới có thể sử dụng được trên lưới mô hình. Thành phần phân biệt D có nhiệm vụ phân biệt đâu là vật liệu do người thiết kế, đâu là vật liệu được tự động sinh ra. Khi mạng trạng thái cân bằng thành phần sinh G thay thế cho con người tạo ra các vật liệu cho mô hình 3D.



Hình 8. Vật liệu trái UV do con người thiết kế

Hình 8 bên trái là ảnh vật liệu được trải UV tương ứng với lưới của mô hình. Hình ảnh bên phải là sự kết hợp giữa lưới của mô hình và vật liệu để tạo ra một mô hình 3D. Chúng tôi sử dụng 50 ảnh vật liệu do con người tạo ra làm tập dữ liệu thực áp dụng trong quá trình huấn luyện mạng GAN. Để đảm bảo tốc độ tính toán của mạng mỗi ảnh vật liệu có kích thước 256x256. Hình 9 bên dưới là ảnh một số vật liệu trong tập dữ liệu thực do con người tạo ra.

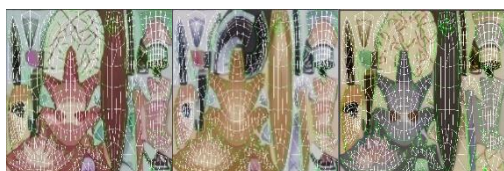


Hình 9. Một số vật liệu do con người thiết kế

Chúng tôi sử dụng kiến trúc của DCGAN [11], một cải tiến của GAN khi sử dụng CNN trong kiến trúc mạng. Khi đó, thành phần sinh dữ liệu G được thiết kế là một nghịch đảo của mạng CNN [9] với mục tiêu tạo ra một ảnh vật liệu có kích thước 256×256 . Mạng CNN thường cho kết quả là phân lớp của ảnh đầu vào (với một ảnh vào có thể cho ra phân lớp là người, cây cối hay động vật tùy theo mục đích thiết kế mạng). Một nghịch đảo của CNN cho kết quả ngược lại. Với vector một chiều làm đầu vào, mạng cho đầu ra là một ảnh màu được sử dụng làm vật liệu cho mô hình 3D. Thành phần phân biệt dữ liệu D được thiết kế là một mạng CNN thuận. Nó được thiết kế với đầu vào là ảnh vật liệu và đầu ra là xác định xem ảnh vật liệu là thực hay giả.

Để hỗ trợ cho quá trình cài đặt và thực nghiệm, chúng tôi sử dụng bộ thư viện mã nguồn mở Tensorflow do Google phát triển phục vụ cho các nghiên cứu về trí tuệ nhân tạo. Đây là bộ mã nguồn mở được nhiều nhà nghiên cứu về học máy nói chung và GAN nói riêng sử dụng rộng rãi khi cài đặt các ứng dụng của mạng nơron.

Quá trình đào tạo GAN được thực hiện theo quy trình đã trình bày ở phần 2 của bài báo. Khi mới bắt đầu, các kết quả tạo ra bởi thành phần sinh dữ liệu cho kết quả sử dụng không cao. Sau quá trình học tập cạnh tranh, các kết quả này được cải thiện dần. Kết quả cuối cùng khi mạng đạt trạng thái cân bằng tương đối khả quan. Hình 10 bên dưới là một số vật liệu do GAN tạo ra.



Hình 10. Một số vật liệu do GAN tạo ra

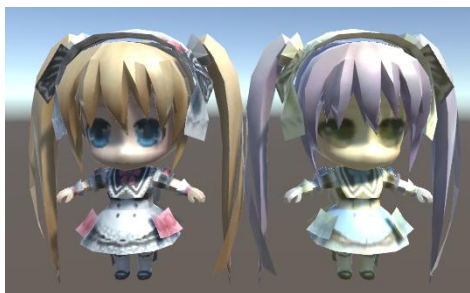
Mạng GAN có ưu thế với việc sáng tạo dữ liệu. Đối với các dữ liệu có tiêu chuẩn đánh giá cố định như video tham số đánh giá được sử dụng trực tiếp trong quá trình đào tạo từ đó

kết quả dữ liệu sinh ra có thể đánh giá thông qua các tiêu chuẩn này [7]. Một số dạng dữ liệu khác như khuôn mặt, là dữ liệu đã được nghiên cứu từ lâu. Có nhiều thuật toán nhận diện khuôn mặt được nghiên cứu, đề xuất. Khi GAN áp dụng sinh khuôn mặt tự động, các hệ thống nhận diện có thể được sử dụng với mục tiêu đánh giá kết quả sinh ra bởi GAN. Bên cạnh đó, có những dạng dữ liệu khác mang tính con người cao như tranh vẽ, các bản thiết kế khó có các tiêu chuẩn đánh giá cố định và hàm đánh giá chính xác, vì thế kết quả của GAN khi sinh các dữ liệu này được đánh giá bởi sự tương đồng giữa dữ liệu được tạo ra bởi con người và dữ liệu được tạo ra bởi máy.

Trong nội dung bài báo, chúng tôi áp dụng GAN sinh dữ liệu là vật liệu cho mô hình 3D. Các vật liệu này được đánh giá dựa trên hai tiêu chí cơ bản. Đầu tiên, các thành phần trong vật liệu đảm bảo vị trí tương ứng khi trải lưới và áp dụng vào mô hình 3D. Thứ hai, khi áp dụng vào mô hình 3D, kết quả hình ảnh của mô hình phải tương đồng với các vật liệu do con người thiết kế.

Khi tiến hành sử dụng vật liệu cho mô hình 3D kết quả cho thấy vật liệu mới được sinh ra đảm bảo cơ bản các yêu cầu kỹ thuật để có thể sử dụng được. Quan sát vật liệu do GAN sinh ra, các nhà thiết kế 3D nhận thấy vật liệu đảm bảo yếu tố về trải UV khi các thành phần được trải tương ứng với vị trí của nó trên mô hình. Đồng thời, các bộ phận cơ thể, họa tiết và hoa văn tương đối hợp lý. Khi áp dụng vật liệu vào mô hình ba chiều cho kết quả tương đối tốt khi hình ảnh nhân vật 3D mới được tạo ra có độ tương đồng cao với hình ảnh mô hình 3D do con người tạo ra.

Hình 11 là mô hình 3D có vật liệu do con người tạo ra và vật liệu được thiết kế bởi máy tính. Tuy chất lượng hình ảnh vật liệu do GAN tạo ra còn chưa thể so sánh với vật liệu do con người tạo ra nhưng đã đảm bảo các yếu tố kỹ thuật và tính hợp lý khi sử dụng.



Hình 11. Mô hình 3D với vật liệu do con người thiết kế và GAN tạo ra



Hình 12. Mô hình 3D với vật liệu do GAN tạo ra

Hình 12 là một số vật liệu do GAN tạo ra áp dụng vào mô hình 3D. Kết quả thử nghiệm cho thấy tiềm năng của việc ứng dụng GAN cho quá trình sinh vật liệu tự động. Các họa tiết hoa văn, chi tiết mắt, lông mày trên khuôn mặt, màu tóc được thay thế tương đối tốt. Tuy nhiên, do tập dữ liệu học tập còn ít đồng thời quá trình huấn luyện đòi hỏi nhiều thời gian và chi phí tính toán nên chất lượng về độ phân giải và đường nét của ảnh vật liệu sinh ra còn chưa cao.

4. Kết luận

Trong nội dung bài báo, chúng tôi trình bày về Generative Adversarial Networks (GAN) mạng nơron được sử dụng khi sinh dữ liệu trên máy tính và một số mô hình phát triển của nó. Đây là mạng nơron có nhiều tiềm năng khi áp dụng sản xuất dữ liệu đa phương tiện. Nhận thấy điều này, một mô hình mạng GAN được chúng tôi áp dụng cho quá trình tạo vật liệu tự động cho các mô hình 3D. Vật liệu là một trong hai thành phần cơ bản của một mô hình 3D tĩnh. Quá trình tự động tạo vật liệu giúp rút ngắn thời gian và công sức khi xây dựng mô hình 3D và có nhiều ứng dụng trong thực tại ảo. Kết quả của thực nghiệm cho thấy vật liệu được sinh ra bởi mạng GAN cho kết quả khả quan và có thể sử dụng được. Tuy nhiên, bộ dữ liệu huấn luyện

còn nhỏ, chất lượng vật liệu khi sinh ra vẫn còn khả năng tối ưu tốt hơn về hình ảnh. Điều này đòi hỏi nhiều công sức hơn cũng như những nghiên cứu sâu hơn về GAN và các đề xuất mới để cải tiến mạng nơron này khi áp dụng vào bài toán sinh vật liệu tự động cho mô hình 3D.

Lời cảm ơn

Nghiên cứu này được hỗ trợ bởi đề tài nghiên cứu khoa học cấp cơ sở năm 2020 - Trường Đại học Công nghệ thông tin và Truyền thông - Đại học Thái Nguyên (Mã số: T2020-07-22).

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1]. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets", *Advances in Neural Information Processing Systems*, vol. 27, pp. 2672-2680, 2014.
- [2]. C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 105-114, 2017.
- [3]. X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. C. Loy, "Esrgan: Enhanced super-resolution generative adversarial networks", *The European Conference on Computer Vision Workshops (ECCVW), Won Region 3 in the PIRM2018-SR Challenge*, 2018.
- [4]. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks", *IEEE International Conference on Computer Vision (ICCV), Venice*, pp. 2242-2251, 2017.
- [5]. T. Zhang, W. Tian, T. Zheng, Z. Li, X. Du, and F. Li, "Realistic Face Image Generation Based on Generative Adversarial Network", *16th International Computer Conference on Wavelet Active Media Technology and Information Processing, Chengdu, China*, pp. 303-306, 2019.
- [6]. Y. Cui, and W. Wang, "Colorless Video Rendering System via Generative Adversarial Networks", *IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Dalian, China*, pp. 464-467, 2019.

-
- [7]. K. Gopan, and G. S. Kumar, "Video Super Resolution with Generative Adversarial Network", *2nd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli*, pp. 1489-1493, 2018.
- [8]. S. Lunz, Y. Li, A. Fitzgibbon, and N. Kushman, "Inverse Graphics GAN: Learning to Generate 3D Shapes from Unstructured 2D Data", *Computer Vision and Pattern Recognition (cs.CV); Machine Learning (cs.LG)*, pp. 1-11, 2020.
- [9]. A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks", *International Conference on Learning Representations, Soumith Chintala, New York*, 2016.
- [10]. M. Mirza, and S. Osindero, "Conditional Generative Adversarial Nets", *Simon Osindero, San Francisco, CA 94103*, 2014.
- [11]. Y. Du, W. Zhang, J. Wang, and H. Wu, "DCGAN Based Data Generation for Process Monitoring", *IEEE 8th Data Driven Control and Learning Systems Conference (DDCLS), Dali, China*, pp. 410-415, 2019.
- [12]. S. T. Le, V. T. Ma, N. D. T. Tran, and T. N. Phung, "Guad mesh optimization in controlling 3d objects applied for Vietnamese sign language performance", *TNU Journal of Science and Technology*, vol. 178(02), pp. 91-96, 2018.