

Diseño de calzado con redes neuronales artificiales

Shoe design with artificial neural networks

Dandy Delvi Montoya Santos¹, Jorge Juan Egúsquiza Gorriti², Serge Justinian Rayn Chávez Feria³, Cristhian Andre Mariños Legendre⁴, Juan Carlos Mariños Legendre⁵, Fiorella Carolina Morillo Méndez⁶, Ana Isabel Huamán Jiménez⁷, Miguel Elías Pinglo Bazán⁸.

Resumen

La investigación presentada se basa en el uso de inteligencia artificial para la automatización y personalización del diseño de calzado, optimizando el proceso de diseño de acuerdo a gustos y preferencias de los clientes. Mediante la utilización de redes neuronales artificiales, la plataforma logró ofrecer opciones de calzado según una paleta de colores, tendencias y catálogos que mostraron inicialmente un diseño en 2D el cual, tras aprobación, pasó a ser a un diseño 3D. Mediante la inteligencia artificial, el usuario obtuvo el diseño de su calzado bajo modalidades diferentes: un diseño aleatorio donde se generaron tres posibles opciones mediante la selección de alguna propuesta mostrada en un catálogo predefinido de calzados y una modalidad definida por gustos y preferencias, donde el usuario selecciona los subconjuntos de calzados que se desea combinar, y el sistema genera una matriz de calzados con todas las posibles combinaciones. Recalcar que el usuario puede modificar el calzado seleccionado según cuatro parámetros ubicados en los “slide bars”. Cuando el usuario haya terminado de modificar las características del diseño este puede ser descargado en local como un archivo JPG. Finalmente, el sistema toma como entrada esta imagen 2D y genera tanto la geometría 3D como la textura del modelo de forma semiautomática.

Palabras claves: Red neuronal artificial; diseño de calzado; realidad virtual; calzado personalizado, innovación del diseño.

¹ Grupo Carusso S.A.C., Trujillo, Perú, pmontoya@grupocasusso.com

² Universidad Privada Antenor Orrego, Trujillo, Perú, jegusquiza@upao.edu.pe

³ Grupo Carusso S.A.C., Trujillo, Perú, chavezferia.justinian@gmail.com

⁴ Universidad Privada Antenor Orrego, Trujillo, Perú, cmarinosl@upao.edu.pe

⁵ Instituto Tecnológico de la Producción, Trujillo, Perú, <https://orcid.org/0000-0002-4716-8244>, jmarinos@itp.gob.pe.

⁶ Instituto Tecnológico de la Producción, Trujillo, Perú, fmorillo@itp.gob.pe.

⁷ Instituto Tecnológico de la Producción, Trujillo, Perú, ahuaman@itp.gob.pe.

⁸ Instituto Tecnológico de la Producción, Trujillo, Perú, <https://orcid.org/0000-0001-7130-0156>, mpinglo@itp.gob.pe.

Abstract

The research presented is based on the use of artificial intelligence for the automation and personalization of footwear design optimizing the design process according to customer tastes and preferences according to a palette of colors, trends and catalogs that initially showed a 2D design which, after approval, became a 3D design. Through artificial intelligence, the user obtained the design of their footwear under different modalities: a random design where three possible options were generated by selecting a proposal shown in a predefined catalog of footwear and a modality defined by tastes and preferences, where the user selects the subsets of footwear that you want to combine and the system generates a shoe matrix with all possible combinations. Emphasize that the user can modify the selected footwear according to four parameters located in the "slide bars". Cuando el usuario haya terminado de modificar las características del diseño este puede ser descargado en local como un archivo JPG. Finally, the system takes this 2D image as input and generates both the 3D geometry and the texture of the model semi-automatically.

Keywords: Artificial neural network; shoe design; virtual reality; custom footwear, design innovation.

Introducción

El principal problema identificado ha sido el tiempo de espera necesario para obtener un diseño de calzado y el alto costo que ello implica, puesto que se tiene que esperar 17 días y gastar en múltiples pruebas antes de obtener un diseño de calzado para fabricación; asimismo, las herramientas y equipos usados para el diseño demuestran que el 98 por ciento efectúan el diseño de forma manual y el 90 por ciento no usa ninguna herramienta ni equipo tecnológico para diseñar el calzado; razones por las cuales se planteó el reto de diseñar de forma automática y personalizada de la mano del cliente a través de la inteligencia artificial. Un estudio de Accenture a tres mil consumidores del sector retail evidenció que existe una demanda de personalización: el 63 por ciento de los consumidores mostró su interés en contar con productos personalizados y su disposición a compartir su información personal a cambio de éstos productos. En tanto, el 56 por ciento confirmó su deseo de probar el producto antes de la compra. Conocedores de estos requerimientos, se buscó disminuir el porcentaje de tercerización de la fase de diseño con mejora en el proceso, reduciendo los tiempos de diseño, automatizándolo y generando que el consumidor adquiriera lo que realmente quiere de acuerdo a sus gustos y preferencias. Para tal fin, ha sido integrado un módulo de calce virtual, el cual permitió obtener las medidas de forma rápida para poder realizar el diseño del calzado personalizado, evitando la devolución a razón de una talla distinta o mal calce que se ve reflejado al tener de cada 10 compras 5 son devueltas por adquirir un calzado que no genera el ajuste deseado.

Material y métodos

La metodología aplicada contempló las siguientes actividades secuenciales.

a. Diseño y construcción de la red neuronal artificial:

El módulo de diseño 2D

Su finalidad ha sido la generación de los diseños de calzados en dos dimensiones, es decir imágenes. Dada una imagen de entrada al sistema, el usuario ha sido capaz de modificar diferentes parámetros del calzado mediante una interfaz gráfica. El resultado de este módulo fue el diseño original.

El módulo de generación 3D

Se encargó de la generación del modelo 3D dado un conjunto de imágenes. El modelo generado agrupó la información geométrica del sistema. En tanto, el módulo de agregación de textura fue el responsable de añadir textura a los modelos 3D que recibe como input.

Modelos Generativos

Los modelos generativos a diferencia de los modelos discriminativos tratan de aprender cómo se encuentran distribuidos los datos en el mundo real a partir de un subconjunto de ellos. Una vez estimada esta distribución de probabilidad, estos modelos brindan una probabilidad alta a muestras que se parezcan a las muestras que fueron utilizadas durante el proceso de entrenamiento, mientras que estima una probabilidad baja para aquellas muestras que por el contrario han sido completamente diferentes a las encontradas en el *dataset* de entrenamiento.

El proceso se inició con la selección de un vector de números aleatorios denominado *vector latente*. La red generativa tomó como entrada este vector y generó una imagen con la resolución y la calidad deseada con el objetivo de poder engañar al discriminador; el discriminador tomó como entrada tanto las imágenes reales como las imágenes ficticias y clasificó si la imagen es real o no. Utilizando estas predicciones y las etiquetas reales se calculó el error que ha cometido el modelo, modificando los parámetros de ambos modelos para asegurar menos errores tanto en las tareas de discriminación como en la tarea de generación; el entrenamiento finalizó cuando el generador fue capaz de engañar al discriminador la mayor parte del tiempo.

Previo a la exploración de las posibles fuentes de datos fue necesario realizar una acotación de las capacidades del modelo con el objetivo de llegar a ser más eficaces durante la colección de las imágenes. Las siguientes suposiciones / restricciones han sido tomadas en cuenta:

1. El modelo a entrenar solo será capaz de generar dos tipos de zapatos en la primera fase. Esto con el objetivo de reducir la complejidad del modelo y el esfuerzo en la colección del dataset. En fases posteriores, se recolectaron imágenes de diferentes tipos de calzado.
2. A priori no ha sido posible conocer cuáles son los atributos que se pueden modificar de las imágenes. Esto solo logró ser posible una vez entrenado el modelo, donde de manera experimental se varía cada parámetro y se visualiza el efecto final sobre la imagen.

Características del dataset

Como se ha mencionado el *dataset* se conformó de solo dos categorías o tipo de calzado. Las imágenes contienen sólo un elemento, las imágenes de pares de zapatos no han sido consideradas; asimismo, se utilizó solo un perfil del calzado, dando preferencia a la vista lateral exterior.

La resolución de las imágenes utilizada para el entrenamiento fue de 256x256 píxeles. El primer dataset estuvo conformado por 5,000 imágenes, siendo el dataset final un conjunto de 50,025 imágenes.

La acotación con este *dataset* es la diversidad en las resoluciones de las imágenes. Para entrenar los modelos de redes GAN's fue necesario que las imágenes tengan una relación de aspecto de 1:1 y que la resolución sea un número exponencial de 2, por ejemplo, 128x128, 256x256 ó 512x512. Debido a esta variedad en las resoluciones se tuvo mayor cuidado durante el redimensionamiento de las imágenes para evitar deformaciones muy pronunciadas.

StyleGAN a diferencia de otras redes GAN no inyecta el vector latente a la red de forma directa, por el contrario, crea una nueva red de neuronas en paralelo que se encarga de procesar este vector antes de que pueda realizar modificaciones al generador e influenciar el proceso de generación de imágenes. La hipótesis detrás de esta decisión en el diseño del modelo se basó en que mediante el uso de esta red auxiliar el modelo podría desacoplar la influencia que tiene cada valor en el vector latente sobre las características de la imagen. Es decir, imaginemos que tenemos un vector latente con 100 posibles valores, cada uno de estos valores en el caso ideal debe poder modificar solo una característica en la imagen generada, sin embargo, uno de los principales problemas en los sistemas de generación de imágenes es que un valor en estos vectores cambia más de una característica a la vez. En el caso específico de diseño de calzados, una modificación de estos valores en el vector latente provocaría un cambio no solo en la forma de la suela si no también una modificación en el alto del taco.

StyleGAN-Ada

Se utilizó la versión StyleGAN-Ada que incluye un artificio para la aumentación de datos en escenarios donde no se tienen datasets los suficientemente grandes para poder entrenar el modelo desde cero.

El aumento de los datos ha sido controlado por el hiper parámetro p . A mayor valor de p , mayor es la probabilidad de aplicar una serie de transformaciones a las imágenes de nuestro dataset, que fueron finalmente las imágenes que nuestro modelo utilizó durante el entrenamiento.

Las transformaciones aplicadas fueron:

1. Cambios del tipo espejo.
2. Rotaciones utilizando diferentes ángulos.
3. Cambios de color.
4. Cambios de contraste.
5. Cambios en la escala.
6. Deformaciones afines, entre otros.

b. Análisis de Híper parámetros del modelo

StyleGAN incluye una serie de híper parámetros que debieron ser ajustados y experimentados durante el proceso de entrenamiento. La siguiente lista describe los hiper parámetros que se tuvieron que ajustar para conseguir resultados óptimos para nuestro problema:

1. --aug: Deshabilita el proceso de aumentación de datos.
2. --mirror: Amplifica el dataset con transformaciones del tipo flip. Normalmente es beneficio activar este tipo de transformaciones.
3. --gamma: Controla la cantidad de regularización que aplicamos al modelo. La regularización permite evitar el sobreajuste del modelo a los datos. Puede tomar valores entre 0 y 1. Valores altos de gamma, regularizan más el modelo.
4. --p: Controla la probabilidad de utilizar transformaciones en el dataset. Puede tomar valores entre 0 y 1.
5. --augpipe: Especifica el tipo de transformaciones a ser aplicadas a las imágenes del dataset en el entrenamiento.
6. --king: Controla la duración del tiempo de entrenamiento.
7. --batch: Controla el tamaño del minibatch.

c. Entrenamiento de Redes Neuronales

El entrenamiento de las redes neuronales consistió en el ajuste de sus parámetros con el objetivo de generar el resultado deseado a la salida de la red. Al igual que en otros algoritmos de aprendizaje automático la forma de realizar estos ajustes se dio a través de un algoritmo de optimización matemática. Antes de aplicar algún algoritmo de optimización se definió nuestra función de coste o pérdida. La función de pérdida es aquella expresión matemática que deseamos minimizar o maximizar dependiendo del problema que estemos abordando. Una función de coste típica es la definida por la ecuación (1), donde se utiliza la media de la suma de los errores cuadráticos para evaluar el rendimiento de la red. Las predicciones generadas fueron comparadas con los valores reales obtenidos. Este proceso se realizó por cada ejemplo en el dataset de entrenamiento, siendo ***m*** el número de muestras. El objetivo fue el de minimizar este error.

Ecuación 1: Función de Pérdida.

$$J(W, b, x, y) = \frac{1}{2m} \sum_{i=1}^m \|f_i(W, b, x, y) - y_i\|^2 + \lambda R(W)$$

Donde:

w = Matriz de pesos de la red.

x = Vector de datos de entrada.

b = Vector de sesgos de la red.

y = Vector de etiquetas.

f_i = Salida de la red para el dato *i*.

y_i = Etiqueta del ejemplo *i*.

El descenso del gradiente ha sido el método utilizado para la optimización de redes de neuronas. Esta técnica realiza actualizaciones de manera iterativa de cada parámetro de la red en aquella dirección que minimice más la función de coste. Para obtener esta dirección se calculó la derivada parcial de la función de coste con respecto a cada parámetro de manera individual. Finalmente, el grado de actualización dependió del hiper parámetro ***α*** o tasa de aprendizaje, el cual pudo ser modificado dinámicamente a medida que fuimos convergiendo hacia el mínimo. Las ecuaciones 2 y 3 describen este proceso.

Ecuación 2: Actualización de pesos de la red neuronal.

$$w := w - \alpha * \frac{\partial J(W, b)}{\partial w}$$

Ecuación 3. Actualización de sesgos de la red neuronal.

$$b := b - \alpha * \frac{\partial J(W, b)}{\partial b}$$

Previo al realizar el entrenamiento ha sido necesario inicializar los parámetros de la red, siendo este un factor importante que logró determinar la convergencia o no. En la práctica, la inicialización se puede realizar de manera aleatoria, utilizando una distribución Gaussiana, o mediante el método de Xavier, siendo este último uno de los más utilizados a la fecha.

Red neuronal artificial permitió la generación de modelos de calzados en 2D de forma automática; el diseño puede ser generado utilizando 3 posibles opciones. La primera opción a partir de la generación de un diseño completamente aleatorio (Ver Figura 1 y 2), una vez generado este diseño fue posible modificar ciertos aspectos de la imagen mediante 04 parámetros que son controlados por los “slide bars” mostrados en la interfaz gráfica.

Figura 1. Generación aleatoria.

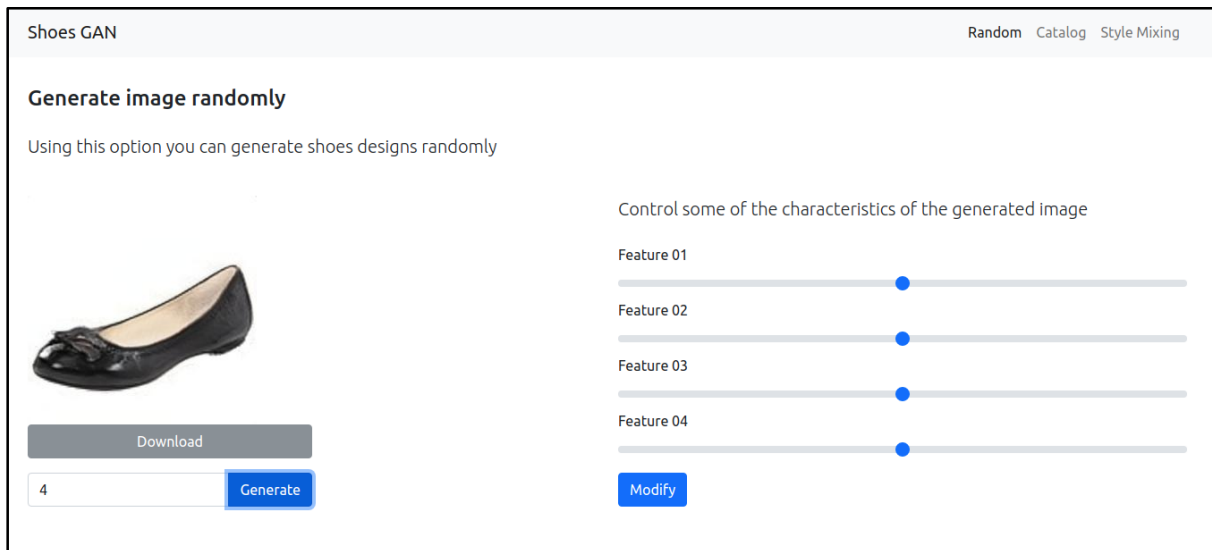
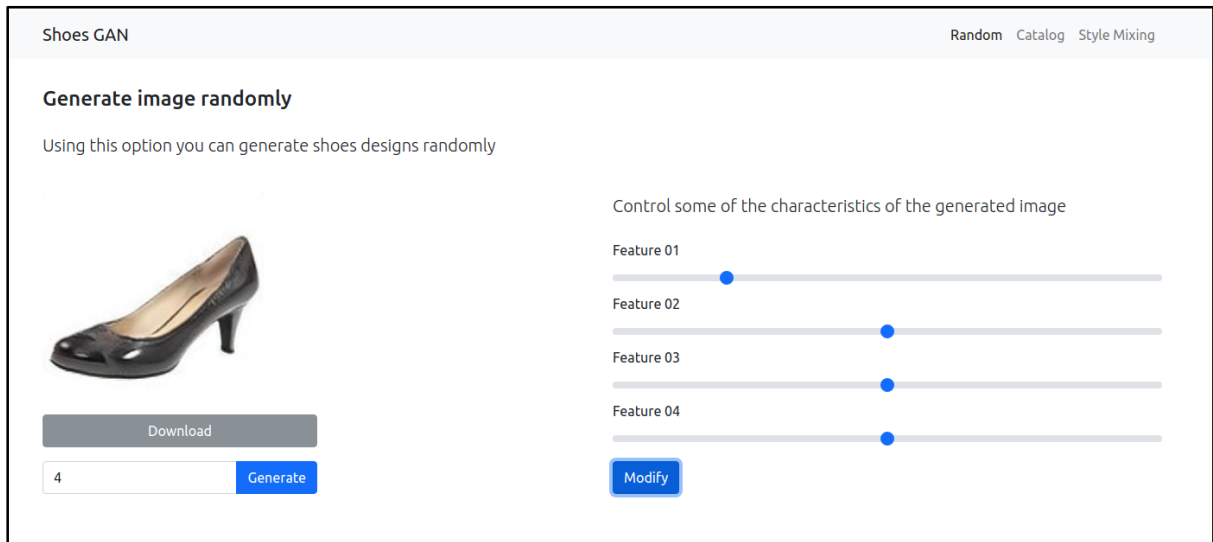


Figura 2. Modificación de imagen aleatoria.



La segunda opción consistió en seleccionar alguna de las opciones que se muestran en un catálogo predefinido de calzados (Ver Figura 3 y 4). El calzado seleccionado logró ser modificado mediante el mismo conjunto de parámetros mencionados previamente.

Figura 3. Diseño de catálogo

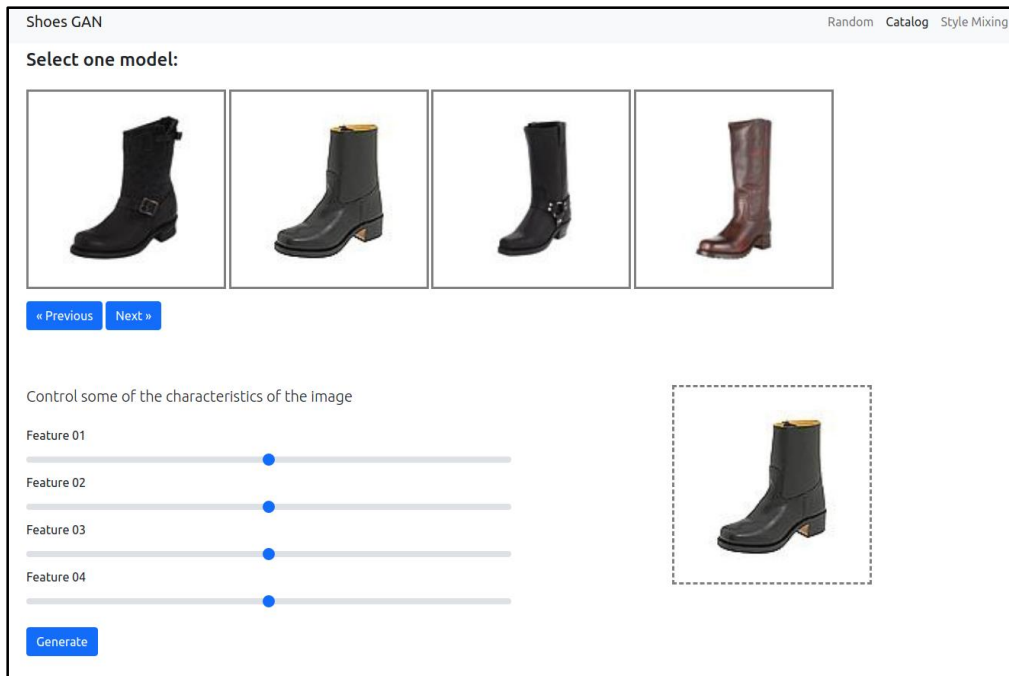
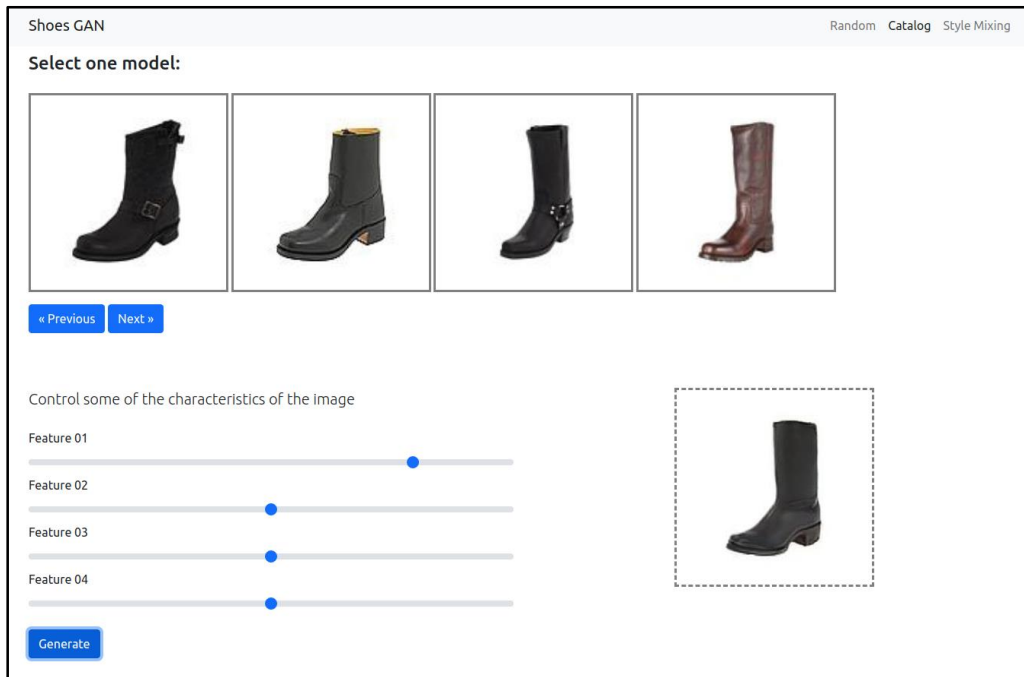


Figura 4. Modificación de imagen de catálogo



La tercera opción por su parte, consistió en generar combinaciones de diseños de calzados a partir de la selección de los subconjuntos de calzados que se deseaban combinar, logrando que el sistema generara un matriz de calzados con todas las posibles combinaciones (Ver Figuras 5 y 6).

Figura 5. Combinación de estilos

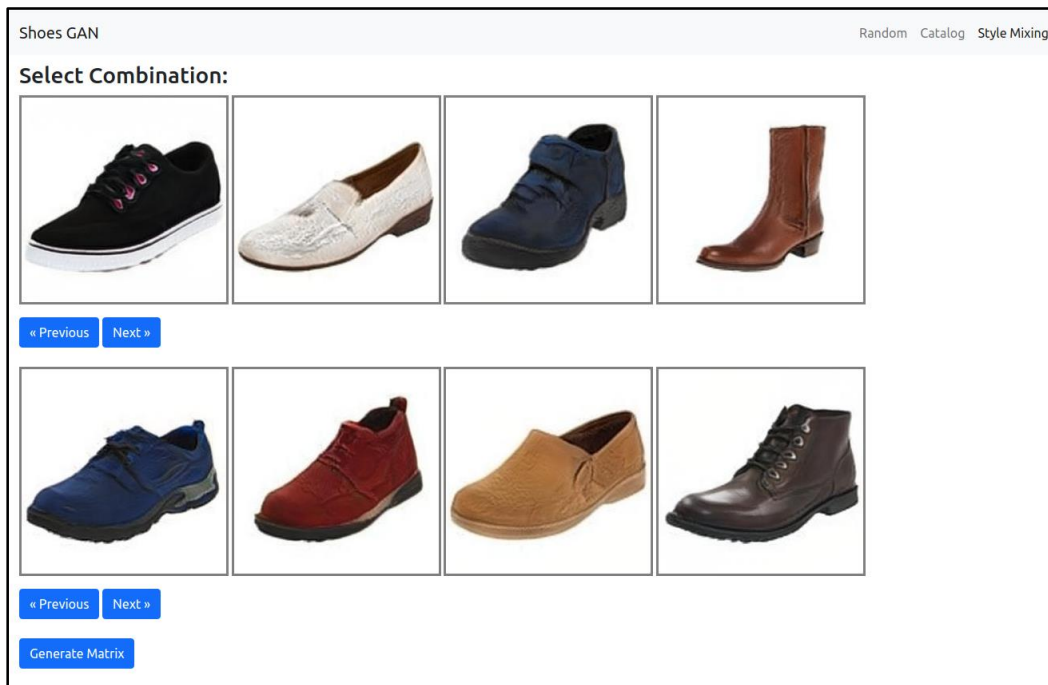


Figura 6. Estilos generados



Cuando el usuario terminó de modificar las características del diseño éste logró ser finalmente descargado en local como un archivo JPG.

Discusión

Entre los modelos identificados que resultaron útiles para nuestro propósito se encuentra el modelo *Dib-Render*, capaz de proporcionar como salida no solo la información geométrica del modelo (*mesh format*), sino también extraer la textura del objeto, y la aplicación de esta textura utilizando diferentes ángulos de visión, creando nuevas vistas del objeto, y todo gracias a una sola imagen.

Otro de los métodos probados fue *Mesh-RCNN* que también al recibir una imagen en 2D de entrada mostró su capacidad de generar la representación en malla del objeto. Sin embargo, el flujo de procesamiento de este modelo resultó ser un poco más complicado que en el caso de *Dib-Render*, donde se tiene que primero ubicar al objeto en la imagen y mediante la rama de detección de objeto y luego realizar el proceso de inferencia de la geometría 3D. Es ideal para aplicaciones donde estamos interesados en estimar los modelos 3D de dos o más objetos a la vez, pero debido a que nuestro problema resultó un poco más restringido se determinó que no resultaba ser la mejor opción. Otra de las restricciones de este método resulta ser su falta de capacidad de agregar textura y color a los objetos 3D por lo que requeriría realizarse un paso adicional para la obtención de las imágenes renderizadas desde diferentes ángulos.

En este artículo, proponemos un modelo de fusión novedoso utilizando una red antagónica generativa de triple discriminador, que puede lograr el equilibrio. La imagen de diferencia obtenida mediante la sustracción de imágenes puede resaltar la información de diferencia, extraer detalles de la imagen y obtener los contornos del objetivo en algunas escenas. Por lo tanto, además del discriminador visible y discriminador de infrarrojos, se agrega un nuevo discriminador de imágenes de diferencia para retener la diferencia entre las imágenes infrarrojas y visibles, mejorando así el contraste de los objetivos infrarrojos y manteniendo los detalles de la textura en las imágenes visibles. Las características multinivel extraídas por los discriminadores se utilizan para la medición de la información y, como resultado, se derivan pesos de fusión perceptivos para la fusión adaptativa. También se introducen la función de pérdida de SSIM y la pérdida de mejora de borde objetivo para mejorar la calidad de la imagen fusionada. En comparación con los métodos de fusión de vanguardia existentes en conjuntos de datos públicos, se demuestra que nuestro modelo tiene un mejor rendimiento en métricas cuantitativas y efectos cualitativos.

Específicamente, entrenamos una red neuronal basada en GAN para regresionar una estimación de movimiento. El modelo resultante recibe un par de imágenes y estima el movimiento relativo entre ellas. Entrenamos la red neuronal utilizando un enfoque semi-supervisado. A diferencia de los sistemas monoculares tradicionales basados en geometría, nuestro método basado en *Deep Learning* es capaz de estimar la escala absoluta de la escena sin información extra ni conocimiento previo. Evaluamos WGANVO en el conocido conjunto de datos KITTI. Demostramos que nuestro sistema funciona en tiempo real y la precisión obtenida alienta a seguir desarrollando sistemas de localización basados en Aprendizaje Profundo.

La moda desde sus inicios fue un diferenciador para las personas, ya que con la fusión de los diferentes campos se demostró que se pueden obtener resultados relevantes y funcionales como lo fue el *Cognitive Dress* y que no solo eso, sino que también con la ayuda de la inteligencia artificial se pueden optimizar los procesos de diseño ya que el análisis de los datos se vuelve más preciso y rápido. La red neuronal convolucional cuenta con la habilidad para clasificar con una variedad bastante amplia en cuanto a identificación de imágenes se refiere, teniendo así la capacidad de llegar a aplicaciones mucho más complejas.

Finalmente, cabe destacar la rapidez con que la red es entrenada y probada, tardando mucho menos tiempo en realizar el proceso completo sin tener una capacidad de cómputo muy elevada.

Conclusiones

Las redes neuronales permiten la generación de modelos de calzado en 2D para finalmente generar el modelo en 3D, obteniendo un proceso de diseño de calzado optimizado y personalizado que incluye la obtención de las medidas de los pies y el modelo de calzado. La automatización de del proceso de diseño para calzado personalizado ha generado los conocimientos para posibilitar la automatización en otras áreas productivas.

Agradecimientos

Un especial agradecimiento al Programa Nacional de Desarrollo Tecnológico e Innovación - Proinnóvate por el cofinanciamiento para el desarrollo de la presente investigación y por su apoyo en el fomento de la innovación, que permita al país fortalecer sus capacidades tecnológicas y de innovación para de esta manera contribuir al crecimiento de la competitividad y productividad con fondos para la generación de proyectos innovadores realizando el acompañando técnico y de monitoreo oportuno que finalmente permita lograr un impacto positivo en el desarrollo productivo y el fortalecimiento de las empresas de todos los sectores y niveles, emprendedores innovadores, entidades de soporte al emprendimiento, asociaciones de productores y gremios empresariales, además de entidades del sector público y entidades académicas.

Contribución de autoría

Juan Carlos Mariños Legendre, Jorge Juan Egusquiza Gorriti, Serge Justinian Rayn Chavez Feria, Cristhian Andre Mariños Legendre, y Fiorella Carolina Morillo Mendez concibieron la idea, diseñaron el estudio, recogieron los de datos y realizaron el análisis e interpretación de datos. Dandy Delvi Montoya Santos aporó y recogió los de datos, Ana Isabel Huamán Jiménez realizó la búsqueda bibliográfica, revisó el borrador y elaboró el primer borrador para envío, realizando en las subsanaciones requeridas posteriormente. Juan Carlos Mariños Legendre revisó el primer borrador del manuscrito y elaboró la versión final. Miguel Elías Pinglo Bazán revisión crítica de su contenido y aprobación final de la versión que va a publicarse.

Conflictos de interés

Ninguno.

Referencias bibliográficas

- Atenas F. et al (2017) Redes Neuronales Adversarias Convolucionales para Generación de Imágenes.
http://games.cmm.uchile.cl/media/uploads/courses/2017/Informe/Atenas_Maldonado__Felipe_Eduardo.pdf
- Castejón, F. (2020). Diseño _Automatico fde circuitos electrónicos analógicos mediante algoritmos evolutivos.
http://e-spacio.uned.es/fez/eserv/tesisuned:ED-Pg-SisInt-Fcastejon/CASTEJON_LAPEYRA_FEDERICO_Tesis.pdf
- Cifuentes A. et al (2019). Desarrollo de una red neuronal convolucional para reconocer patrones en imágenes. Universidad Simón Bolívar.
- Cremona, J. A. et al (2021) «WGANVO: odometría visual monocular basada en redes adversarias generativas», Revista Iberoamericana de Automática e Informática industrial. doi: 10.4995/riai.2022.16113.
<https://polipapers.upv.es/index.php/RIAI/article/view/16113>
- Image Completion with Deep Learning in TensorFlow (2017, 25 Junio). Recuperado de: <http://bamos.github.io/2016/08/09/deep-completion/>
- Improving the way neural networks learn (2017, 27 Junio). Recuperado de: http://neuralnetworksanddeeplearning.com/chap3.html#the_cross-entropy_cost_function.
- María C. Pastás (2021). El Vestido Tecnológico: El Diseño De Vestuario Y Las Digital Tech.
- Padilla, L. Bonifaz, N. (2021) Desarrollo de un sistema de recomendación inteligente para selección de prendas utilizando análisis de imágenes con redes neuronales. <https://repository.usta.edu.co/handle/11634/34732>
- Song, A. et al (2022). Red antagonica generativa de triple discriminador para fusión de imágenes visibles e infrarrojas.
<https://www.sciencedirect.com/science/article/pii/S0925231222001679>
- Villagrán M. & Villa V. (2019). LA INTELIGENCIA ARTIFICIAL Y LA ALTA COSTURA El caso del “cognitive dress”.
<https://madgu.unison.mx/index.php/madgu/article/view/31/34>