




Article

StyleGANs and Transfer Learning for Generating Synthetic Images in Industrial Applications

Harold Achicanoy ^{1,2,*} , Deisy Chaves ^{2,3,*}  and Maria Trujillo ^{2,t} ¹ Alliance of Bioversity International and CIAT, Km 17 Recta Cali-Palmira, Palmira 763537, Colombia² School of Computer and Systems Engineering, Universidad del Valle, Cali 760001, Colombia; maria.trujillo@correounivalle.edu.co³ Department of Electrical, Systems and Automation, Universidad de León, 24007 León, Spain

* Correspondence: h.a.achicanoy@cgiar.org (H.A.); deisy.chaves@unileon.es (D.C.)

† These authors contributed equally to this work.

Abstract: Deep learning applications on computer vision involve the use of large-volume and representative data to obtain state-of-the-art results due to the massive number of parameters to optimise in deep models. However, data are limited with asymmetric distributions in industrial applications due to rare cases, legal restrictions, and high image-acquisition costs. Data augmentation based on deep learning generative adversarial networks, such as StyleGAN, has arisen as a way to create training data with symmetric distributions that may improve the generalisation capability of built models. StyleGAN generates highly realistic images in a variety of domains as a data augmentation strategy but requires a large amount of data to build image generators. Thus, transfer learning in conjunction with generative models are used to build models with small datasets. However, there are no reports on the impact of pre-trained generative models, using transfer learning. In this paper, we evaluate a StyleGAN generative model with transfer learning on different application domains—training with paintings, portraits, Pokémon, bedrooms, and cats—to generate target images with different levels of content variability: bean seeds (low variability), faces of subjects between 5 and 19 years old (medium variability), and charcoal (high variability). We used the first version of StyleGAN due to the large number of publicly available pre-trained models. The Fréchet Inception Distance was used for evaluating the quality of synthetic images. We found that StyleGAN with transfer learning produced good quality images, being an alternative for generating realistic synthetic images in the evaluated domains.

Keywords: data augmentation; fine-tuning; generative models; StyleGAN; transfer learning

Citation: Achicanoy, H.; Chaves, D.; Trujillo, M. StyleGANs and Transfer Learning for Generating Synthetic Images in Industrial Applications. *Symmetry* **2021**, *13*, 1497. <https://doi.org/10.3390/sym13081497>

Academic Editor: José Carlos R. Alcantud

Received: 27 June 2021

Accepted: 6 August 2021

Published: 16 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Deep learning methods, a subset of machine learning techniques, have achieved outstanding results on challenging computer vision problems, such as image classification, object detection, face recognition, and motion recognition, among others [1]. However, the use of deep learning requires a large volume of representative annotated data to learn general models that achieve accurate results [2]; data are still scarce with asymmetric distributions, i.e., disproportionate number of examples between classes, in most applications related to healthcare, security and industry, due to legal/ethical restrictions, unusual patterns/cases, and image annotation costs [3–5].

As an alternative, image data augmentation has emerged to create training data with symmetric distributions by increasing data and reducing overfitting in deep learning models [6]. Data augmentation has been used through simple transformations, such as rotations, mirroring, and noise addition [4]. However, simple transformations produce a reduced number of valid data which usually are highly correlated and produce overfit models with poor generalisation capacity.

Generative Adversarial Networks (GANs) [7] have emerged as an alternative to create synthetic images by learning the probability distribution from data and generating images with high diversity and low correlation that can be used to build deep learning models [4,5,8–13]. GANs are used in medical applications, such as CT image segmentation [4,14] and disease/injury detection [12,13,15–17]. Nevertheless, since GANs are deep-learning-based models, they also require a significant amount of data and computational time to be trained from scratch. This drawback limits the use of GANs in generating images in applications where data are scarce, such as security and industry. A way to cope with this disadvantage is the use of transfer-learning techniques, which allow building new models from pre-trained ones in other applications or source domains with an abundance of training data by transferring the main features and reducing the training time [2].

Transfer learning has been widely used to address image classification [18–20] and segmentation [21,22] problems. However, the effect of StyleGAN-transfer learning on generating image quality is poorly reported. Wang et al. [23] evaluated the transferability of features, using different source and target domains to build generative models applying transfer learning, with some limitations, such as the generation of low-resolution images and the lack of evaluation of the impact of content variability in target domains.

In this paper, we evaluate a data augmentation strategy based on transfer learning and StyleGAN [24]. We use the first version of StyleGAN due to the large number of publicly available pre-trained models. In particular, we evaluate the capability of StyleGAN and transfer learning to generate synthetic images (data augmentation) considering variability levels on content. Thus, we assess quantitatively and visually the quality of the generated images, using three target domains with fine-tuned StyleGANs from five pre-trained models—source domains. The evaluated target domains correspond to three image sets derived from industrial processes with different levels of content variability, shown in Figure 1: bean seeds (low variability), faces of people aged between 5 and 19 years (medium variability), and chars obtained during coal combustion (high variability). The assessed source domains, to transfer features and build generative models, correspond to five pre-trained StyleGANs with images of paintings, portraits, Pokémon, bedrooms, and cats. Distinct from the commonly used transfer learning strategy, which consists of using related source and target domains, the evaluation is focused on source and target domains that are completely different. Obtaining the results shown, StyleGAN with transfer learning is suitable for the generation of high-resolution images in industrial applications due to having a good generalisation capability regarding content variability of the target image.

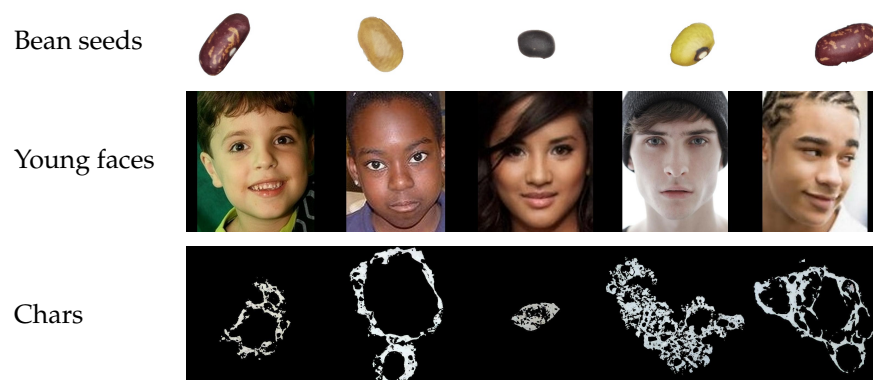


Figure 1. Illustration of target domains for image generation.

The rest of the paper is structured as follows: Section 2 presents the theoretical background on StyleGAN, GANs assessment, and transfer learning. Section 3 summarises the relevant related works. Section 4 details the data augmentation strategy used as an evaluation pipeline. Section 5 describes the performed experiments and results. Section 6 presents the discussion on the obtained results, focusing on the effect of pre-trained models for synthetic image generation; Section 7 depicts the conclusions and future research lines.

2. Theoretical Background

2.1. StyleGAN

StyleGAN [24] combines the architecture of Progressive Growing GAN [25] with the style transfer principles [26] in Figure 2. StyleGAN's architecture addresses some limitations of the GAN models, such as stability during training and lack of control over the images generated.

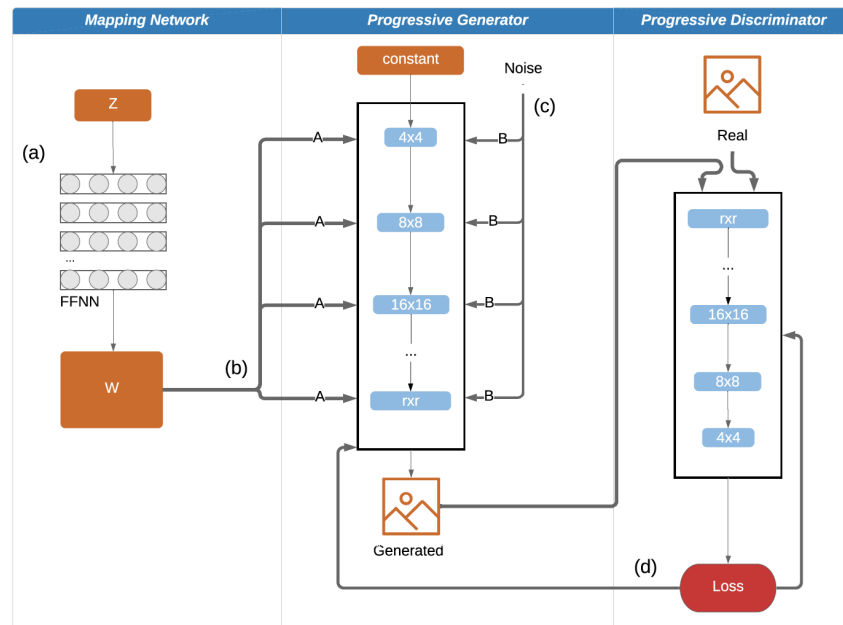


Figure 2. The StyleGAN architecture [27] is composed of three neural networks. (a) Mapping network, which converts a random into a style signal. (b,c) Progressive generator network, which receives the style signal (A) and random noise (B), and produces images progressively. (d) The progressive discriminator network, which compares real and generated images to update all the weights for the three networks, improving their performance.

In a traditional GAN model, a generative network receives as input a random vector Z , or latent vector, to generate a new image. In contrast, in the StyleGAN architecture, a latent vector Z (512-dimensional) feeds an 8-layer neural network, called a mapping network, that transforms the latent vector into an intermediate space W (512-dimensional), which defines the style of the image to be generated; see Figure 2a.

An image style defined in the intermediate space W is transferred to the progressive generative network (Figure 2b), where the technique AdaIN (Adaptive Instance Normalization) [26] transforms the latent vector W into two scalars (scale and bias) that control the style of the image generated at each resolution level. In addition to the style guide provided by AdaIN, the progressive generator network has as an input a constant argument. This constant corresponds to an array of $4 \times 4 \times 512$ dimensions, i.e., an image of 4×4 pixels with 512 channels, which is learned during network training and contains a kind of sketch with the general characteristics of the training set images. Furthermore, StyleGAN has noise sources injected at each resolution level to introduce slight variations in the generated images (Figure 2c). The improvements in the StyleGAN generative network allow optimising the quality of the generated synthetic images. Finally, the back propagation algorithm is applied to adjust the weights of the three networks, improving the quality of the images generated during the following iterations (Figure 2d).

2.2. GANs Model Evaluation

Evaluating GAN architectures is particularly hard because there is not a consensus on a unique metric that assesses the quality and diversity of generated images [4,23].

However, a widely used metric in the literature is the Fréchet Inception Distance (FID) [28], defined as follows:

$$FID = \|\mu_r - \mu_g\|^2 + \text{Tr}\left(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{\left(\frac{1}{2}\right)}\right), \quad (1)$$

where μ_r, μ_g are mean vectors and Σ_r, Σ_g are variance–covariance matrices. Equation (1) is a distance measurement between real and generated images. $X_r \sim N(\mu_r, \Sigma_r)$ and $X_g \sim N(\mu_g, \Sigma_g)$ are multidimensional normal distributions of 2048 dimensions of real and generated images, respectively, which are extracted from the third layer of pooling of the Inception-v3 network [29].

The closer the distributions, the lower the metric value. FID values close to zero correspond to a larger similarity between real and generated images, resulting in the differences between the distributions.

2.3. Transfer Learning

Transfer learning models involve applying the knowledge learned in a source domain—where a large amount of training data are available—to a target domain that has a reduced amount of data, as illustrated in Figure 3. The objective is to transfer most characteristics obtained from a source domain into a target domain in order to reduce training time and use deep learning models with limited data [30].

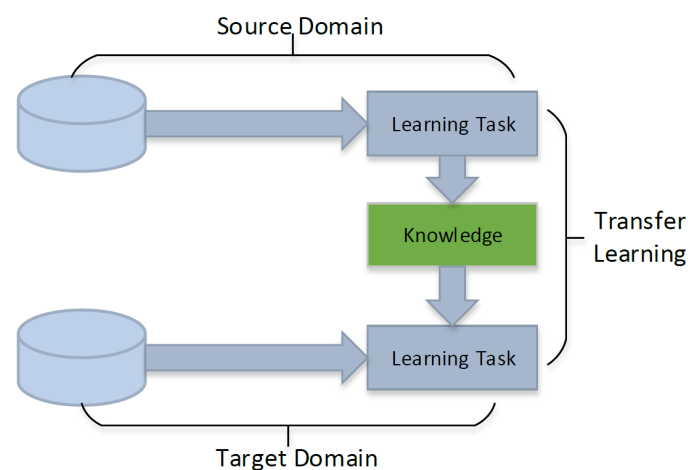


Figure 3. Transfer learning scheme [2]. The learning task to be accomplished in the source domain is provided by a vast number of training data. The target domain usually has a limited number of training data.

The transfer learning is defined as follows: given a source domain D_S , a source learning task T_S , a target domain D_T , and a target learning task T_T , the transfer learning aims to improve the learning of the target predictive function $f_T(\cdot)$ in D_T using the knowledge learned in D_S and T_S , where $D_S \neq D_T$ or $T_S \neq T_T$. A domain D is defined by two components: a feature space X and a marginal probability distribution $P(X)$, $D = \{X, P(X)\}$. Similarly, a learning task T consists of two components: a labels space Y and a target predictive function $f(\cdot)$, $T = \{Y, f(\cdot)\}$.

In unsupervised models, such as GANs, the Y labels space does not exist, and the learning task objective is to estimate the generative distribution of data. For this purpose, a particular transfer learning technique, known as fine tuning, is used. Fine tuning is a transfer learning technique in which a model that has been trained for a specific task is used to perform a new task, with similar characteristics to the first task, by adjusting model parameters. This process means that a model is not built from scratch, but rather takes advantage of the characteristics learned from the original task.

3. Related Works

GANs are capable of image generation in two categories: low-resolution [4,8,12,13,15,16,23] and high-resolution [14,17,31–41]. A summary of these approaches is presented in Table 1. The first group comprises mostly studies between 2017 and 2018. A predominance of medical image augmentation focus on cancer detection [4,15], cerebral diseases [16] and COVID-19 [12,13] is observed. The image generation is motivated by the high cost of medical images acquisition that translates into a limited number of samples to train deep learning models. Publications include studies that evaluated the effectiveness of GANs data augmentation, using computer vision benchmarks datasets [8,23]. Particularly, the work of Wang et al. [23] corresponds to the only attempt to use transfer learning with generative models. This study concluded that it is possible to apply transfer learning in a Wasserstein GAN model [42], using source and target domains with low-resolution images.

The second group includes studies from 2019 to date, driven by new GAN architectures, such as StyleGAN [24], generating images with high-resolution. There are three application areas: agriculture [31,32], medical [14,17,36,37], and electrical domains [35]. Particularly, Fetty et al. [17] presented a complete analysis of StyleGAN models trained from scratch for data augmentation of pelvic malignancies images.

Transfer learning on generative models for limited data has been the subject of study for the last three years [33,34,38–41], focusing on evaluating the impact of freezing the lower generator layers [33,34], the lower discriminator layers [39], and both the generator and discriminator lower layers [40], using mainly general purposes datasets of indoors (e.g., LSUN, Bedrooms) and faces (e.g., CelebHQ, FFHQ, CelebA). The results show a reduction in the overfitting derived from the knowledge transfer and training time. However, transfer learning in conjunction with generative models has not been evaluated with a focus on the capability to generate synthetic images with high-resolution, considering the variability levels of content.

We aim to fill this literature gap using the proposed pipeline to evaluate the transfer capability of the knowledge obtained from certain source domains to target domains with different levels of content variability, such as bean seed images (with simple shape, texture, and colour), young faces and chars images (with more complex visual features). These target domains correspond to real industrial applications.

Table 1. Summary of GANs approaches for image generation.

Application Description				Model Description			
Year	Context	Images Number	Images Quality	Generative Model	Transfer Learning	Training Time	Evaluation Metric
2017 [8]	Data augmentation in alphabets, handwritten characters, and faces	Alphabets: 1200, Characters: 3400, Faces: 1802	Low-resolution	Conditional GAN	Yes	Not available	Not available
2018 [4]	Data augmentation using computed tomography images	Different variations	Low-resolution	Progressive Growing GAN	No	~36 GPU hours	Not available
2018 [15]	Data augmentation using GANs for liver injuries classification	Computed tomography images: 182	Low-resolution	Deep Convolutional GAN (DCGAN)	No	Not available	Not available
2018 [16]	Medical image data augmentation using GANs	Alzheimer’s Disease Neuroimaging Initiative: 3416	Low-resolution	Pix2Pix GAN	No	Not available	Not available
2018 [23]	Transfer learning in GANs for data augmentation	Varies from 1000 to 100,000	Low-resolution	Wasserstein GAN with Gradient Penalty (WGAN-GP)	Yes	Not available	FID minimum: 7.16, FID maximum: 122.46
2020 [12]	Transfer learning in GANs for COVID-19 detection on chest X-ray images	Normal cases: 79, COVID-19: 69, Pneumonia bacterial: 79, Pneumonia virus: 79	Low-resolution	Shallow GAN	Yes	Not available	Not available
2020 [13]	COVID-19 Screening on chest X-ray images	Normal cases: 1341, Pneumonia: 1345	Low-resolution	Deep Convolutional GAN (DCGAN)	No	GPU Titan RTX; 100 epochs	Not available
2019 [31]	Plant diseases detection	Plant leaves: 79,265	High-resolution	StyleGAN	No	Not available	Not available
2019 [32]	Data augmentation on plant leaf diseases	Plant leaf disease: 54,305	High-resolution	DCGAN & WGAN	No	1000 epochs	Not available
2019 [14]	Data augmentation using GANs to improve CT segmentation	Pancreas CT: 10,681	High-resolution	CycleGAN	No	3 M iterations	Qualitative evaluation
2019 [33]	Image generation from small datasets via batch statistics adaptation	Face, Anime, and Flowers: 251–10,000	High-resolution	SNGAN, BigGAN	Yes	3000, 6000–10,000 iterations	FID: 84–130

Table 1. Cont.

Application Description				Model Description			
Year	Context	Images Number	Images Quality	Generative Model	Transfer Learning	Training Time	Evaluation Metric
2019 [34]	Mind2Mind: transfer learning for GANs	MNIST, KMNIST, and CelebHQ: 30,000–60,000	High-resolution	MindGAN	Yes	Not available	FID: 19.21
2020 [35]	Data augmentation using GANs for electrical insulator anomaly detection	Individual insulators: 3861	High-resolution	BGAN, AC-GAN, PGAN, StyleGAN, BPGAN	No	Not available	Not available
2020 [17]	Data augmentation using StyleGAN for pelvic malignancies images	17,542	High-resolution	StyleGAN	No	One GPU month	FID: 12.3
2020 [36]	Data augmentation for Magnetic Resonance Brain Images	50,000	High-resolution	StyleGAN and Variational autoencoders	No	Not available	Not available
2020 [38]	MineGAN: effective knowledge transfer from GANs to target domains with few images	MNIST, CelebA, and LSUN: 1000	High-resolution	Progressive GAN, SNGAN, and BigGAN	Yes	200 iterations	FID: 40–160
2020 [39]	Freeze the discriminator: a simple baseline for fine-tuning GANs	Animal face, Flowers: 1000	High-resolution	StyleGAN, SNGAN	Yes	50,000 iterations	FID: 24–80
2020 [40]	On leveraging pretrained GANs for generation with limited data	CelebA, Flowers, Cars, and Cathedral: 1000	High-resolution	GP-GAN	Yes	60,000 iterations	FID: 10–80
2021 [37]	Data augmentation for Cardiac Magnetic Resonance	6000	High-resolution	Conditional GAN-based method	Yes	Not available	Not available
2021 [41]	MineGAN++: mining generative models for efficient knowledge transfer to limited data domains	MNIST, FFHQ, Anime, Bedroom, and Tower: 1000	High-resolution	BigGAN, Progressive GAN, and StyleGAN	Yes	200 iterations	FID: 40–100

4. Evaluation Synthetic Images Generation Pipeline

We proposed a five-step pipeline based on the fine tuning of StyleGAN pre-trained models from five source domains, as is shown in Figure 4—paintings, portraits, Pokémon, bedrooms, and cats—in order to generate synthetic images in three target domains: bean seeds, young faces, and chars. Although, there is a new version of StyleGAN, called StyleGAN2 [43], we selected the first version of StyleGAN, due to the large number of publicly available pre-trained StyleGAN models. Moreover, in practice, training a StyleGAN model from scratch requires a huge number of images, computational resources (preferably with multiple GPUs) and processing time.

In the proposed image generation pipeline, first, we select the images target domain of bean seeds, young faces, or chars, as input. Second, images from the target domain are pre-processed to improve the transferability of features by adjusting image resolution. Third, pre-trained StyleGAN models are fine tuned with pre-processed images. Fourth, the FID evaluation metric is used for selecting the best source domain. Fifth, the synthetic images for the input target domain are generated with the best source domain.

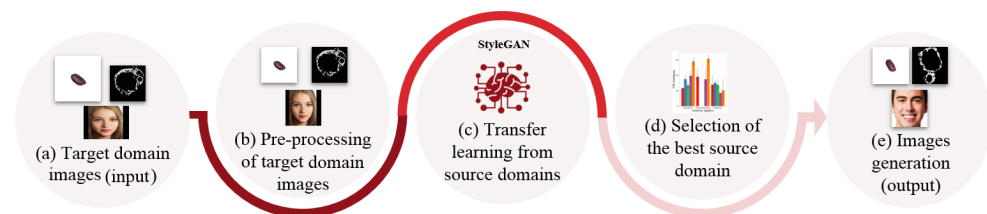


Figure 4. Evaluation pipeline: (a) Input images from a target domain. (b) Pre-processing of target domain images to improve the transferability of the features. (c) Transfer learning using StyleGAN pre-trained models from source domains. (d) Selection of the best source domain using the FID metric. (e) Generation of synthetic images for the target domain.

4.1. Input Target Domain Images

We used images from three application domains—bean seeds, young faces and chars—with different variability levels of content and potential industrial applications; see Figure 1.

Bean seeds: The dataset [44] has 1500 seed images from 16 bean varieties. Bean seed images are classified as low content variability since shape, colour, and texture characteristics are homogeneous for the analysed seed varieties, corresponding to oval shapes with limited range of red, cream, black, and white colours. In addition, the acquired images share the same background colour. Synthetic images of bean seeds are valuable in developing evaluation tools of genetic breeding trials. These tools are used to preserve the genetic pedigree of seeds over time, accomplish market quality requirements, and increase production levels [45].

Young faces: The images set consists of 3000 images randomly selected from publicly available datasets with reference to age estimation problems: IMDB-WIKI [46], APPA-Real [47], AgeDB [48]. Images correspond to individuals aged between 5 and 19 years. This range of ages was selected because it presents the lowest frequencies in the considered facial datasets. In addition, young faces are crucial in cybersecurity applications, such as access control, the detection of Child Sexual Exploitation Material or the identification of victims of child abuse [49–51]. Young facial images are considered to be of medium variability content since faces have a similar shape structure; the StyleGAN model was originally designed for faces generation.

Chars: The dataset contains 2928 segmented char particle images from coals of high, medium, and low reactivity. Char images are considered to be of high variability content due to the complex particle shapes and lack of colours. Synthetic images of char particles are useful to train models to estimate the combustion parameters in power generation plants [52].

Table 2 contains a summary of the target domains, including the source, number of images, number of classes, and content variability type.

Table 2. Description of target domain datasets.

Target Domain	# of Images	# of Classes	Content Variability
Bean seeds [44]	1500	16	Low
Young faces [46–48]	3000	14	Medium
Chars [52]	2928	3	High

4.2. Pre-Processing Target Domain Images

Input images are pre-processed to improve the transfer of features from the source domain images into the target domain. The pre-processing consists of equally adjusting the images resolution to the resolution of the source domain. In short, operations of resizing, up-scaling, or down-scaling are applied, depending on the difference between the target and source images' dimensions.

4.3. Transfer Learning from Source Domains

We used five pre-trained StyleGAN models—paintings, portraits, Pokémon, bedrooms, and cats—shown in Figure 5. The selection of pre-trained models was based on the (i) public availability of models, and (ii) diversity of images used to build models. Table 3 presents relevant information about pre-trained models: the images source, required images resolution, and the number of iterations used for training.

**Figure 5.** Illustration of generated images using the training source domain models.**Table 3.** Description of pre-trained StyleGAN models (source domains).

Source Domain	Image Resolution	Number of Iterations
Paintings [53]	512 × 512	8040
Portraits [54]	512 × 512	11,125
Pokemon [55]	512 × 512	7961
Bedrooms [56]	256 × 256	7000
Cats [56]	256 × 256	7000

Transfer learning is performed by fine tuning the pre-trained StyleGAN models (source domains) with images of the target domain to build new image generators. During the fine tuning of StyleGAN models, the learning rate is set to 0.001 and the number of minibatch repetitions is set to 1 based on those reported in [57]. The selected values for the learning rate and minibatch repetition increase the stability and speed during training.

4.4. Selection of the Best Source Domain

Once the StyleGAN models are fine tuned, the best source domain is selected based on the best FID metric value to generate images of a target domain. In particular, the StyleGAN model with the lowest FID value has a better performance since the model generates images with a distribution similar to the target domain.

4.5. Synthetic Images Generation (Output)

The StyleGAN with the best performance is used to generate as many images as needed for the target domain (data augmentation).

5. Experimental Evaluation








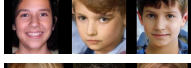

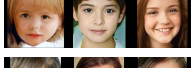

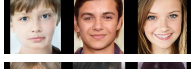




We assess the transfer learning capability of pre-trained models from five source domains—paintings, portraits, Pokémon, bedrooms, and cats—to build StyleGAN models for generating images of unrelated target domains with different levels of content variability: bean seeds, young faces, and chars. Pre-processed images from target domains are used to fine tune pre-trained StyleGAN models (source domains) over 1000 iterations, using the hyperparameters described in Section 4.3. We run the experiments on a GNU/Linux machine with a GPU Nvidia TITAN Xp 11GB, Cuda 10.1, and CuNDD 7. The source code is available at https://github.com/haachicanoy/stylegan_augmentation_tl (accessed on 28 July 2021).

Table 4 presents the FID values obtained for the fine-tuned StyleGAN models, and Table 5 illustrates the generated images by the target domain.

Table 4. FID values for StyleGAN models built for target domains by fine tuning pre-trained models from different source domains. The best source domain to generate images of a target domain by transfer learning is highlighted in bold. Lower FID values indicate better performance.

Source Target	Bean Seeds	Young Faces	Chars
Paintings	23.26	27.77	38.13
Portraits	35.04	30.11	—
Pokémon	27.06	27.56	—
Bedrooms	39.31	16.98	34.81
Cats	57.92	20.48	61.52

Table 5. Generated images for target domain using fine tuned StyleGAN models.

Source Target	Bean Seeds	Young Faces	Chars
Original image			
Paintings			
Portraits			—
Pokémon			—
Bedrooms			
Cats			

The results show that StyleGAN models are able to generate bean seed images (low content variability) through transfer learning with excellent performance. FID values range between 23.26 and 57.92, corresponding to the source domains of paintings and cats,

respectively. Hence, the best source domain to generate bean seed images is paintings. It indicates that the colour pattern in these images is more similar to beans in comparison to the other source domains.

Regarding the generation of young face images (medium content variability), the bedrooms source domain achieves the best results (FID of 16.98). Bedrooms are one of the most complete source domains, standing out for colour and shape features that are efficiently transferred to generate facial images. It is essential to highlight that the five source domains yield FID values lower than 30.11. This performance may be related to the fact that the StyleGAN architecture was specifically developed for generating synthetic face images.

During the fine tuning of models to generate char images, we observed that the source domains of Pokémon and portraits do not converge, leading the training to fail. This is presumably because of the high content variability of chars with complex shapes, varying sizes, and changes in colour intensities that make it difficult to adapt the features from these two source domains. The remaining source domains (paintings, bedrooms, and cats) have features that can be successfully transferred to the generation of char images. In particular, bedrooms achieve the best performance (FID of 34.81).

In most of the cases, fine tuned models generate images of bean seeds, young faces and chars with visual characteristics that are similar to the original ones (see Table 5). However, in some cases, the generated images have visual defects; see Figure 6. Defects on bean seed images comprise bean shape deformation, stains, and changes in colour intensities. Defects on young facial images correspond to colour spots and alterations in the hair, skin, and smile. Defects on chars images include blurring and undefined shapes. Therefore, the generated images have to be filtered to remove images with defects before using them in any application, e.g., the training of an image classifier. A sample of 1000 synthetic images of target domains generated from the best source domains is available at <https://doi.org/10.7910/DVN/HHSJY8> (accessed on 23 July 2021).

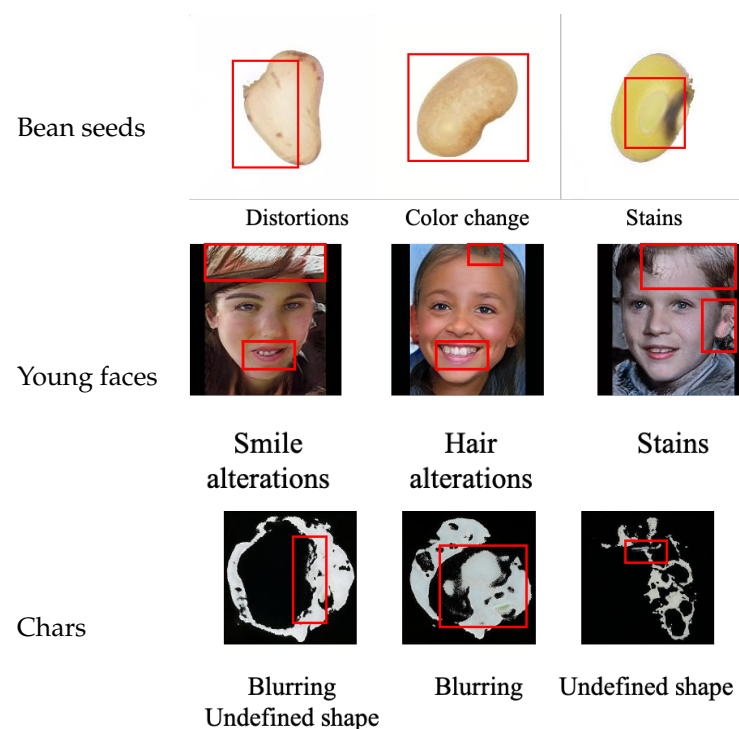


Figure 6. Defects in generated images by target domain.

6. Effect of Pre-Trained Models on Synthetic Images Generation

Figure 7 shows the evolution of the FID metric across the iterations of fine tune StyleGAN models for evaluating target domains (bean seeds, chars and young faces). The

results show that, in most cases, the trained models for a target domain—regardless of the source domain—tend to stabilise and converge to a constant FID value, after a certain number of iterations. This occurs for all cases, except for target domains with high content variability (chars) where some source domains do not converge—portraits and Pokémon. Hence, we conclude that transfer learning and generative models, such as StyleGAN, can be successfully used to build generators of images with low and medium content variability, such as seeds and faces. However, the generation of synthetic images with high content variability is limited by the characteristics of source domains. In particular, the best FID values are obtained for the source domains of paintings (bean seeds) and bedrooms (chars and young faces).

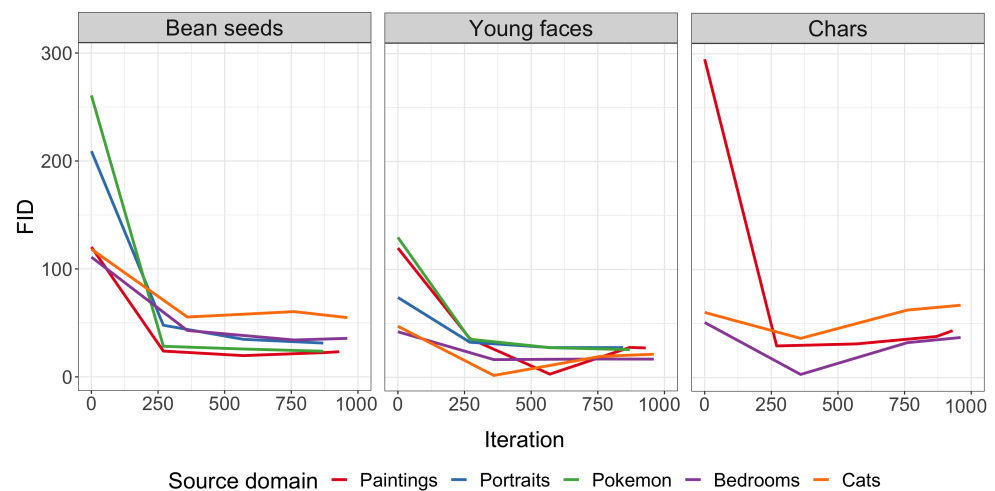


Figure 7. FID metric vs. StyleGAN training iterations for the evaluated source domains grouped by target domain.

We also analysed the loss score values of the generator and discriminator networks during training, shown in Figure 8. Similar to the observed for the FID values, the loss scores of the source domains exhibit a steady behaviour, except for the cats' domain, indicating an instability around 500 iterations.

Furthermore, the use of transfer learning reduces significantly the number of images (up to 1500 for beans) and iterations (up to 1000 or 2 days) required to build StyleGAN models, in comparison to models trained from scratch (70,000 images and 14 days) [24].

Regarding the quality of the synthetic images, Figure 9 presents a bar graph of the FID values by source domains—paintings, portraits, Pokémon, bedrooms, and cats—grouped by the target domain: bean seeds, chars, and young faces. The bar highs correspond to median values of the FID obtained during the fine tuning of StyleGAN models by a source domain, while the black line on the bars represents dispersion of the median FID values. The length of the line denotes the range of the dispersion. Larger lines indicate that the images generated from a source domain differ significantly from the target domain.

The source domains with the best performance—lower median FID value and dispersion—are bedrooms, for the young faces and chars, and paintings, for the bean seeds. Despite bedrooms yielding the lowest FID value in the target domain of chars, this source domain is the one with the largest dispersion, indicating a possible generation of char images with defects.

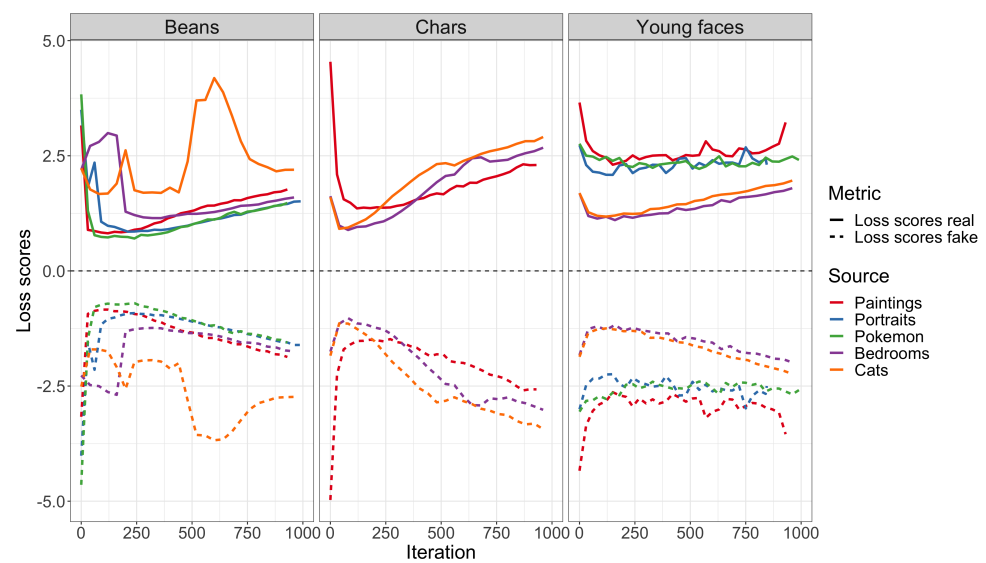


Figure 8. Loss scores vs. StyleGAN training iterations for the evaluated source domains grouped by target domain. Solid and dotted lines correspond to discriminator and generator network scores, respectively; following the source domains colours.

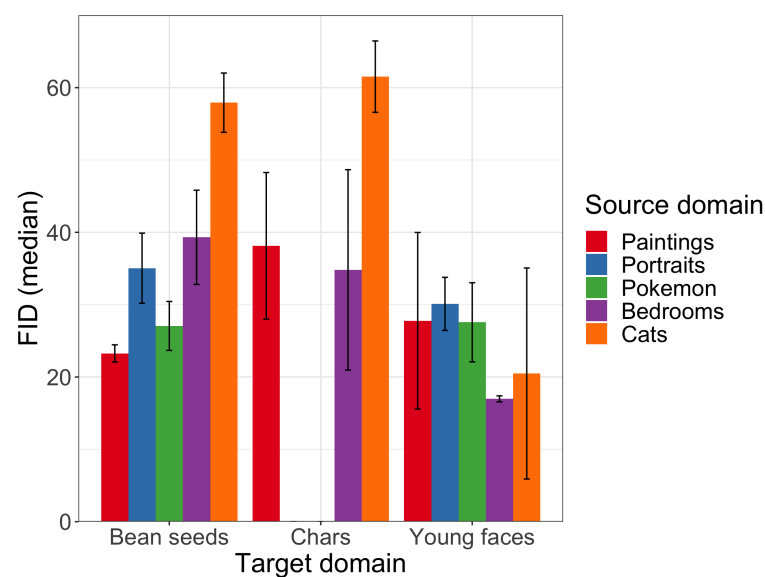


Figure 9. FID (median) for target domains and their source domains.

7. Conclusions

StyleGAN with transfer learning is a strategy for generating synthetic images with a limit number of images from the target domain. We evaluated the application of StyleGAN with transfer learning on generating high-resolution images by a pipeline based on the fine tuning of StyleGAN models. The evaluation was conducted using three target domains from industrial applications with different content variability (bean seeds, chars, and young faces) and five source domains from general applications (paintings, portraits, Pokémon, bedrooms, and cats) to perform transfer learning.

The experimental evaluation confirmed the potential of StyleGAN with transfer learning for generating synthetic images for industrial applications. The proposed pipeline performed better with target domains with low and medium content variability in terms of colour and shape, such as bean seeds and young faces. Moreover, the time and number of images required to build the models were reduced in all cases, which validates the use of StyleGAN with transfer learning for generating synthetic images.

As future work, strategies to optimise the fine tuning hyper-parameters will be evaluated to improve the performance of image generators with high content variability and reduce the defects in synthetic images. General purpose datasets will be assessed, such as FFHQ and LSUN.

Author Contributions: Conceptualisation, H.A., D.C. and M.T.; methodology, H.A., D.C. and M.T.; software, H.A.; validation, H.A. and D.C.; formal analysis, H.A., D.C. and M.T.; investigation, H.A., D.C. and M.T.; resources, H.A., D.C. and M.T.; data curation, H.A.; writing—original draft preparation, H.A.; writing—review and editing, D.C. and M.T.; visualisation, H.A. and D.C.; supervision, D.C. and M.T.; project administration, D.C. and M.T.; funding acquisition, H.A., D.C. and M.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded with support from the European Commission under the 4NSEEK project with Grant Agreement 821966. This publication reflects the views only of the author; the European Commission cannot be held responsible for any use that may be made of the information contained therein.

Data Availability Statement: Generated data can be found at <https://doi.org/10.7910/DVN/HHSJY8> (accessed on 28 July 2021).

Acknowledgments: This work was supported by the framework agreement between the Universidad de León and INCIBE (Spanish National Cybersecurity Institute) under Addendum 01. Additionally, the authors thank Alliance Bioversity-CIAT for providing the bean seed images.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study.

References

- Dargan, S.; Kumar, M.; Ayyagari, M.R.; Kumar, G. A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning. *Arch. Comput. Methods Eng.* **2019**, *27*, 1071–1092. [CrossRef]
- Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. A Survey on Deep Transfer Learning. *arXiv* **2018**, arXiv:1808.01974.
- Haixiang, G.; Yijing, L.; Shang, J.; Mingyun, G.; Yuanyue, H.; Bing, G. Learning from class-imbalanced data: Review of methods and applications. *Expert Syst. Appl.* **2017**, *73*, 220–239. [CrossRef]
- Bowles, C.; Chen, L.; Guerrero, R.; Bentley, P.; Gunn, R.; Hammers, A.; Dickie, D.A.; Hernández, M.V.; Wardlaw, J.; Rueckert, D. GAN augmentation: Augmenting training data using generative adversarial networks. *arXiv* **2018**, arXiv:1810.10863.
- Tanaka, F.H.K.d.S.; Aranha, C. Data Augmentation Using GANs. *arXiv* **2019**, arXiv:1904.09135.
- Shorten, C.; Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6*, 60. [CrossRef]
- Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In *Advances in Neural Information Processing Systems 27*; Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q., Eds.; Curran Associates, Inc.: New York, NY, USA, 2014; pp. 2672–2680.
- Antoniou, A.; Storkey, A.; Edwards, H. Data augmentation generative adversarial networks. *arXiv* **2017**, arXiv:1711.04340.
- Garcia Torres, D. Generation of Synthetic Data with Generative Adversarial Networks. Ph.D. Thesis, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Karlskrona, Sweden, 2018.
- Zeid Baker, M. Generation of Synthetic Images with Generative Adversarial Networks. Master's Thesis, Department of Computer Science and Engineering, Blekinge Institute of Technology, Karlskrona, Sweden, 2018.
- Ma, Y.; Liu, K.; Guan, Z.; Xu, X.; Qian, X.; Bao, H. Background Augmentation Generative Adversarial Networks (BAGANs): Effective Data Generation Based on GAN-Augmented 3D Synthesizing. *Symmetry* **2018**, *10*, 734. [CrossRef]
- Loey, M.; Smarandache, F.; Khalifa, N.E.M. Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning. *Symmetry* **2020**, *12*, 651. [CrossRef]
- Zulkifley, M.A.; Abdani, S.R.; Zulkifley, N.H. COVID-19 Screening Using a Lightweight Convolutional Neural Network with Generative Adversarial Network Data Augmentation. *Symmetry* **2020**, *12*, 1530. [CrossRef]
- Sandfort, V.; Yan, K.; Pickhardt, P.J.; Summers, R.M. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. *Sci. Rep.* **2019**, *9*, 16884. [CrossRef]
- Frid-Adar, M.; Diamant, I.; Klang, E.; Amitai, M.; Goldberger, J.; Greenspan, H. GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification. *Neurocomputing* **2018**, *321*, 321–331. [CrossRef]
- Shin, H.C.; Tenenholtz, N.A.; Rogers, J.K.; Schwarz, C.G.; Senjem, M.L.; Gunter, J.L.; Andriole, K.; Michalski, M. Medical Image Synthesis for Data Augmentation and Anonymization using Generative Adversarial Networks. *arXiv* **2018**, arXiv:1807.10225.
- Fetty, L.; Bylund, M.; Kuess, P.; Heilemann, G.; Nyholm, T.; Georg, D.; Löfstedt, T. Latent Space Manipulation for High-Resolution Medical Image Synthesis via the StyleGAN. *Z. Med. Phys.* **2020**, *30*, 305–314. [CrossRef]

18. Coulibaly, S.; Kamsu-Foguem, B.; Kamissoko, D.; Traore, D. Deep neural networks with transfer learning in millet crop images. *Comput. Ind.* **2019**, *108*, 115–120. [[CrossRef](#)]
19. Abu Mallouh, A.; Qawaqneh, Z.; Barkana, B.D. Utilizing CNNs and transfer learning of pre-trained models for age range classification from unconstrained face images. *Image Vis. Comput.* **2019**, *88*, 41–51. [[CrossRef](#)]
20. Liu, X.; Wang, C.; Bai, J.; Liao, G. Fine-tuning Pre-trained Convolutional Neural Networks for Gastric Precancerous Disease Classification on Magnification Narrow-band Imaging Images. *Neurocomputing* **2020**, *392*, 253–267. [[CrossRef](#)]
21. Ferguson, M.; Ak, R.; Lee, Y.T.T.; Law, K.H. Detection and Segmentation of Manufacturing Defects with Convolutional Neural Networks and Transfer Learning. *Smart Sustain. Manuf. Syst.* **2018**, *2*, 137–164. [[CrossRef](#)]
22. Abdalla, A.; Cen, H.; Wan, L.; Rashid, R.; Weng, H.; Zhou, W.; He, Y. Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure. *Comput. Electron. Agric.* **2019**, *167*, 105091. [[CrossRef](#)]
23. Wang, Y.; Wu, C.; Herranz, L.; van de Weijer, J.; Gonzalez-Garcia, A.; Raducanu, B. Transferring GANs: Generating images from limited data. *arXiv* **2018**, arXiv:1805.01677.
24. Karras, T.; Laine, S.; Aila, T. A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 4401–4410.
25. Karras, T.; Aila, T.; Laine, S.; Lehtinen, J. Progressive Growing of GANs for Improved Quality, Stability, and Variation. *arXiv* **2018**, arXiv:1710.10196.
26. Huang, X.; Belongie, S. Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization. *arXiv* **2017**, arXiv:1703.06868.
27. Oeldorf, C. Conditional Implementation for NVIDIA’s StyleGAN Architecture. 2019. Available online: <https://github.com/cedricoeldorf/ConditionalStyleGAN> (accessed on 13 November 2019).
28. Heusel, M.; Ramsauer, H.; Unterthiner, T.; Nessler, B.; Hochreiter, S. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. In *Advances in Neural Information Processing Systems 30*; Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc.: New York, NY, USA, 2017; pp. 6626–6637.
29. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. *arXiv* **2015**, arXiv:1512.00567.
30. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
31. Arsenovic, M.; Karanovic, M.; Sladojevic, S.; Anderla, A.; Stefanovic, D. Solving Current Limitations of Deep Learning Based Approaches for Plant Disease Detection. *Symmetry* **2019**, *11*, 939. [[CrossRef](#)]
32. Arun Pandian, J.; Geetharamani, G.; Annette, B. Data Augmentation on Plant Leaf Disease Image Dataset Using Image Manipulation and Deep Learning Techniques. In Proceedings of the 2019 IEEE 9th International Conference on Advanced Computing (IACC), Tiruchirappalli, India, 13–14 December 2019; pp. 199–204. [[CrossRef](#)]
33. Noguchi, A.; Harada, T. Image generation from small datasets via batch statistics adaptation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27–28 October 2019; pp. 2750–2758.
34. Fregier, Y.; Gouray, J.B. Mind2Mind: transfer learning for GANs. In Proceedings of the International Conference on Geometric Science of Information. Paris, France, 21–23 July 2021; Springer: 2021; pp. 851–859.
35. Luo, L.; Hsu, W.; Wang, S. Data Augmentation Using Generative Adversarial Networks for Electrical Insulator Anomaly Detection. In Proceedings of the 2020 2nd International Conference on Management Science and Industrial Engineering, Osaka, Japan, 7–9 April 2020; Association for Computing Machinery: Osaka, Japan, 2020; pp. 231–236. [[CrossRef](#)]
36. Hirte, A.U.; Platscher, M.; Joyce, T.; Heit, J.J.; Tranvinh, E.; Federau, C. Diffusion-Weighted Magnetic Resonance Brain Images Generation with Generative Adversarial Networks and Variational Autoencoders: A Comparison Study. *arXiv* **2020**, arXiv:2006.13944.
37. Xia, Y.; Ravikumar, N.; Greenwood, J.P.; Neubauer, S.; Petersen, S.E.; Frangi, A.F. Super-Resolution of Cardiac MR Cine Imaging using Conditional GANs and Unsupervised Transfer Learning. *Med. Image Anal.* **2021**, *71*, 102037. [[CrossRef](#)] [[PubMed](#)]
38. Wang, Y.; Gonzalez-Garcia, A.; Berga, D.; Herranz, L.; Khan, F.S.; Weijer, J.V.D. Minegan: effective knowledge transfer from gans to target domains with few images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 9332–9341.
39. Mo, S.; Cho, M.; Shin, J. Freeze the discriminator: A simple baseline for fine-tuning gans. *arXiv* **2020**, arXiv:2002.10964.
40. Zhao, M.; Cong, Y.; Carin, L. On leveraging pretrained GANs for generation with limited data. In Proceedings of the 37th International Conference on Machine Learning, Virtual Event, 13–18 July 2020; pp. 11340–11351.
41. Wang, Y.; Gonzalez-Garcia, A.; Wu, C.; Herranz, L.; Khan, F.S.; Jui, S.; van de Weijer, J. MineGAN++: Mining Generative Models for Efficient Knowledge Transfer to Limited Data Domains. *arXiv* **2021**, arXiv:2104.13742.
42. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein Generative Adversarial Networks. In Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; Precup, D., Teh, Y.W., Eds.; International Convention Centre: Sydney, Australia, 2017; Volume 70, pp. 214–223.
43. Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Lehtinen, J.; Aila, T. Analyzing and Improving the Image Quality of StyleGAN. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020.
44. Diaz, H. Bean seeds images calibration dataset. Alliance of Bioversity International and CIAT. Unpublished raw data, 2019.
45. Beebe, S. Breeding in the Tropics. In *Plant Breeding Reviews*; John Wiley & Sons, Ltd.: New York, NY, USA, 2012; Chapter 5, pp. 357–426.

46. Rothe, R.; Timofte, R.; Gool, L.V. DEX: Deep EXpectation of apparent age from a single image. In Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW), Santiago, Chile, 7–13 December 2015; pp. 10–15.
47. Agustsson, E.; Timofte, R.; Escalera, S.; Baró, X.; Guyon, I.; Rothe, R. Apparent and real age estimation in still images with deep residual regressors on APPA-REAL database. In Proceedings of the FG 2017—12th IEEE International Conference on Automatic Face and Gesture Recognition, Washington, DC, USA, 30 May–3 June 2017; pp. 1–12.
48. Moschoglou, S.; Papaioannou, A.; Sagonas, C.; Deng, J.; Kotsia, I.; Zafeiriou, S. AgeDB: The First Manually Collected, In-the-Wild Age Database. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 1997–2005.
49. Ye, L.; Li, B.; Mohammed, N.; Wang, Y.; Liang, J. Privacy-Preserving Age Estimation for Content Rating. In Proceedings of the 2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP), Vancouver, BC, Canada, 29–31 August 2018; pp. 1–6.
50. Anda, F.; Lillis, D.; Kanta, A.; Becker, B.A.; Bou-Harb, E.; Le-Khac, N.A.; Scanlon, M. Improving Borderline Adulthood Facial Age Estimation through Ensemble Learning. In Proceedings of the 14th International Conference on Availability, Reliability and Security, Canterbury, UK, 26–29 August 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 1–8. [[CrossRef](#)]
51. Chaves, D.; Fidalgo, E.; Alegre, E.; Jáñez-Martino, F.; Biswas, R. Improving Age Estimation in Minors and Young Adults with Occluded Faces to Fight Against Child Sexual Exploitation. In Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications—Volume 5: VISAPP, Valletta, Malta, 27–19 February 2020; INSTICC; SciTePress: Setúbal, Portugal, 2020; pp. 721–729. [[CrossRef](#)]
52. Chaves, D.; Fernández-Robles, L.; Bernal, J.; Alegre, E.; Trujillo, M. Automatic characterisation of chars from the combustion of pulverised coals using machine vision. *Powder Technol.* **2018**, *338*, 110–118. [[CrossRef](#)]
53. StyleGAN Trained on Paintings (512 × 512). Available online: <https://colab.research.google.com/drive/1cFKK0CBnev2BF8z9BOHxePk7E-f7TtUi> (accessed on 9 March 2020).
54. StyleGAN-Art. Available online: <https://github.com/ak9250/stylegan-art> (accessed on 15 March 2020).
55. StyleGAN-Pokemon. Available online: <https://www.kaggle.com/ahsenk/stylegan-pokemon> (accessed on 15 March 2020).
56. StyleGAN—Official TensorFlow Implementation. Available online: <https://github.com/NVLabs/stylegan> (accessed on 19 December 2020).
57. Gwern. Making Anime Faces With StyleGAN. 2019. Available online: <https://www.gwern.net/Faces> (accessed on 9 March 2020).