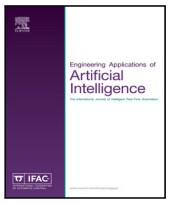
Engineering Applications of Artificial Intelligence 143 (2025) 110019

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai



Research paper

Synthetic image generation for effective deep learning model training for ceramic industry applications

Fábio Gaspar [a], Daniel Carreira [a], Nuno Rodrigues [a], Rolando Miragaia [a], José Ribeiro [a], Paulo Costa [a], António Pereira [a,b,*]

[a] Computer Science and Communications Research Centre, School of Technology and Management, Polytechnic of Leiria, 2411-901 Leiria, Portugal
[b] INOV INESC Inovação, Institute of New Technologies, Leiria Office, 2411-901 Leiria, Portugal

ARTICLE INFO

Keywords:
Ceramic industry
Computer vision
Deep learning
Image classification
Synthetic data

ABSTRACT

In the rapidly evolving field of machine learning engineering, access to large, high-quality, and well-balanced labeled datasets is indispensable for accurate product classification. This necessity holds particular significance in sectors such as the ceramics industry, in which effective production line activities are paramount and deep learning classification mechanisms are particularly relevant for streamlining processes; but real-world image samples are scarce and difficult to obtain, hindering dataset building and consequently model training and deployment. This paper presents a novel approach for dataset building in the context of the ceramic industry, which involves employing synthetic images for building or complementing datasets for image classification problems. The proposed methodology was implemented in *CeramicFlow*, an innovative computer graphics rendering pipeline designed to create synthetic images by employing computer-aided design models of ceramic objects and incorporating domain randomization techniques. As a result, a fully synthetic image dataset named *Synthetic CeramicNet* was created and validated in real-world ceramic classification problems. The results demonstrate that synthetic images provide an adequate basis for datasets and can significantly reduce reliance on real-world data when developing deep learning approaches for image classification problems in the ceramic industry. Furthermore, the proposed approach can potentially be applied to other industrial fields.

1. Introduction

Computer Vision (CV) is a field focused on enabling machines to interpret and analyze digital images and video (Voulodimos et al., 2018; Chai et al., 2021). In recent years image classification has made remarkable advances, thanks to the progress in graphical processing units (GPUs) and deep learning (DL) – which has proven to be extremely effective in image classification tasks (de Melo et al., 2022; Gaidon et al., 2018) – and the availability of open source libraries like PyTorch (Paszke et al., 2019). Nevertheless, DL approaches typically rely on large amounts of data for training, even when applying transfer learning techniques (Sun et al., 2017; Dawson et al., 2023), and are often hindered by the unavailability of adequate samples that can be used to build datasets for specific problem areas.

While CV enables the easy classification of pieces, it also poses challenges, particularly in the data required for DL techniques. To ensure the accurate classification of new types of ceramic pieces, it becomes essential to capture new images to retrain the machine learning (ML) model; unfortunately, acquiring these images is a time-consuming and challenging task, which can significantly impede the production workflow. Additionally, the process involves the labor-intensive task of labeling and organizing the captured images.

The proposed approach for mitigating the challenges posed by the difficulty of acquiring real-world images for building DL datasets involves the integration of CAD 3D models; by leveraging the available 3D data files, synthetic images (artificially generated images that simulate the characteristics and variations present in real-world images) of products can be generated and used as training data for image classification models — without the need for physically producing the products or manually capturing real-world images.

As shown in Fig. 1, CAD 3D models are employed not only for the production of real-world ceramic pieces (that can be photographed and labeled for posterior inclusion in DL datasets) but also, and most importantly, for the generation of synthetic images that can be employed as training samples — complementing or even replacing real-world

* Corresponding author at: Computer Science and Communications Research Centre, School of Technology and Management, Polytechnic of Leiria, 2411-901 Leiria, Portugal.
E-mail addresses: fabio.c.gaspar@ipleiria.pt (F. Gaspar), daniel.s.carreira@ipleiria.pt (D. Carreira), nunorod@ipleiria.pt (N. Rodrigues), rolando.miragaia@ipleiria.pt (R. Miragaia), jose.ribeiro@ipleiria.pt (J. Ribeiro), paulo.costa@ipleiria.pt (P. Costa), apereira@ipleiria.pt (A. Pereira).

https://doi.org/10.1016/j.engappai.2025.110019
Received 27 February 2024; Received in revised form 24 July 2024; Accepted 5 January 2025
Available online 14 January 2025

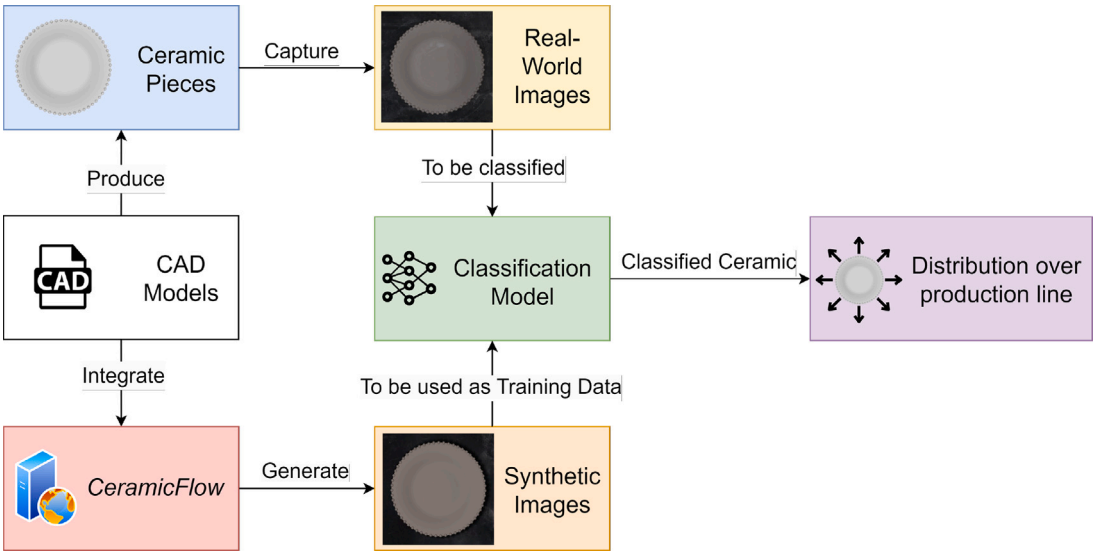


Fig. 1. Abstract representation of a classification process in a ceramic industry relying on synthetic data.

images. Thus, Real-world images are obtained on location, by photographing real ceramic pieces in a controlled environment; whereas synthetic images are generated automatically with basis on CAD 3D models using the computer graphics rendering pipeline developed for this purpose. This study demonstrates that synthetic images can significantly reduce the time and cost associated with collecting and labeling data for DL problems while improving the accuracy of classification models, and ultimately leading to more efficient and productive production lines in the ceramic industry.

A real-world set of images was made available for the purpose of research by a ceramic industry representative (Section 3.1), as well as the 3D CAD models utilized for generating those pieces (Fig. 7). However, the common scenario quickly became apparent in this industry; the number of real-world images available was insufficient to support the training of DL models as the limited quantity of images is very likely to lead to overfitting; What is more, the process of acquiring further real-world images is laborious, time-consuming, and even inexecutable in the early stages of the production process as no real-world ceramic pieces exist at that point.

Several relevant contributions to the field of CV linked to the use of synthetic data for solving a real-world problem are presented, namely:

1. the development of a computer graphics rendering pipeline designed for the creation of synthetic data;
2. the creation of a synthetic dataset for the purpose of image classification in the ceramic industry, which provides a top-down perspective of ceramic products;
3. the validation of the proposed approach by training state-of-the-art DL models using the synthetic dataset, and assessing their performance when classifying real-world industrially manufactured ceramic images obtained in a production environment.

This article is divided into six Sections. Section 2 provides an in-depth analysis of related research that has explored synthetic data and the techniques that have been applied in these contexts; Section 3 provides a description of the scenarios created to evaluate the proposed solution; Section 4 explains the *CeramicFlow* computer graphics rendering pipeline, as well as the techniques applied and the assembled architecture to create and classify ceramic pieces; Section 5 describes the synthetic ceramic dataset as well as the techniques applied to its creation; and Section 6 discuss the obtained results, comparing the different techniques used; finally, in Section 7 highlights the benefits that the use of synthetic data brings to the ceramics industry.

2. Related work

The lack of data in industries has become a significant challenge that hinders the development and deployment of effective ML models and systems.

Synthetic and simulated data has become increasingly popular in recent years due to its potential to overcome the limitations of traditional data collection methods, such as cost, time, privacy concerns, and data quality issues (Nikolenko, 2019; Xu et al., 2022, 2023). This type of data can be used in various ML tasks, such as image classification, object detection, text generation, and others.

In the context of image classification, synthetic data can be used as standalone training data or as a data augmentation technique to improve the performance of the model. One approach is by using game engines along with scripting and 3D models, which provide a virtual representation of real-world objects. This technique has been used in various studies from different domains and has shown promising results in improving the classification accuracy of models (Öztürk and Erçelebi, 2021; Müller et al., 2018; Abu Alhaija et al., 2018; Solovyev et al., 2022; Gaidon et al., 2016; Melo et al., 2020; Osinski et al., 2020). Another approach to synthetic data generation is the use of generative models such as Generative Adversarial Networks (GANs); however, this approach requires a large and diverse training dataset to produce high-quality synthetic images (Aranha et al., 2019; Jain et al., 2022; Rather and Kumar, 2024).

Recent research has also placed significant emphasis on harnessing high-quality synthetic images for diverse applications (Mahmood et al., 2018; Spindler et al., 2020; Barth et al., 2020). In Shrivastava et al. (2016), an unsupervised model was utilized to elevate the authenticity of synthetic images produced by a simulator incorporating unlabeled real data, while ensuring the retention of annotations information from the simulator. Similarly, in another study (Ren and Lee, 2017), an unsupervised feature space domain adaptation method based on adversarial learning was employed, enabling the assimilation of insights from synthetic images and facilitating the adaptation to real images.

Moreover, several methodologies have been introduced to support the precision and resilience of image classification models when integrating synthetic data. One notable approach involves a two-stage classification technique, where objects are categorized as "parents", and viewing angles images generated via domain randomization techniques are considered as "children", empowering the model to effectively manage instances of overclassification (Iwasaki and Yoshioka, 2019).

Furthermore, some researchers have focused on developing specific layers to improve the accuracy of image classification. One such layer is the corner detection and nearest point search (CDNTS), as introduced by Öztürk and Erçelebi (2021), which improves feature extraction from both synthetic and real images, ultimately enhancing the performance of DL networks.

Our approach for the generation of synthetic images is adapted for the ceramic industry, emphasizing the significance of only requiring the ceramic pieces' CAD files that are used in the industry for their production, unlike other methods that still rely on a substantial amount of real-world data to produce high-quality synthetic images. While it is adapted to the ceramic, it also offers the versatility to be utilized in other scenarios. Unlike previous methods that relied on complex setups and specialized hardware, our approach is entirely self-contained. It can be used with any device, eliminating the need for dependencies and making it highly practical for ceramic manufacturers of all scales.

One of the key advantages of the proposed approach is its ability to generate high-quality synthetic images on demand in a short period of time. This fast generation process enables manufacturers to have an immediate and diverse group of images at their disposal for training and testing ML models, improving the accuracy and effectiveness of the models in ceramics manufacturers. With a vast number of images covering a wide range of ceramic variations, ML models can start learning more quickly and effectively which translates to faster deployment and implementation of improved models, leading to enhanced productivity and quality control. Additionally, the flexibility of our approach allows for the customization and adaptation of synthetic images according to specific requirements and characteristics. By providing immediate access to a diverse dataset of synthetic images in real-time, our approach expands the knowledge base and advances the field of ceramics manufacturing and ML applications within it.

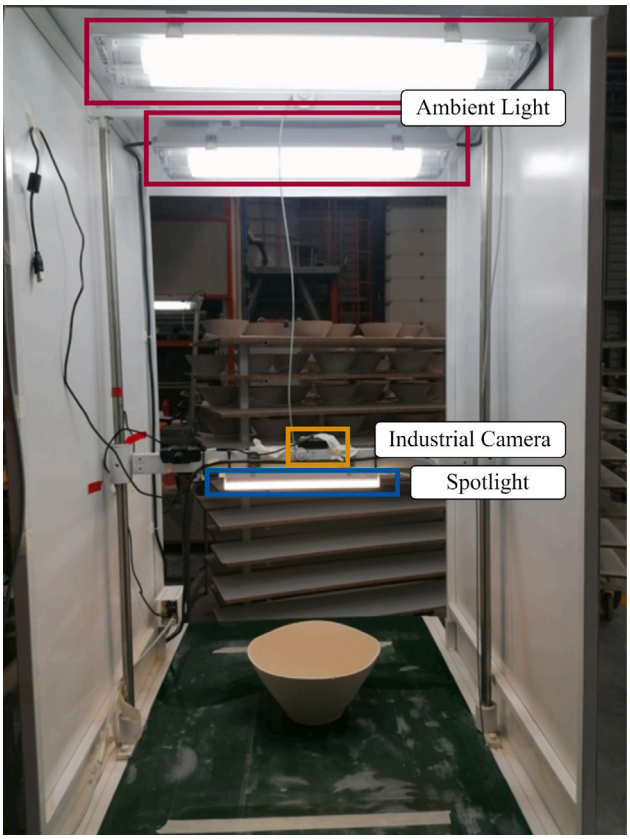


Fig. 2. Industrial scenario layout.

3. Collecting ceramic images in real-world scenarios

The conventional methods of capturing and labeling ceramic piece images pose significant challenges and conditions for training ML models, often resulting in time-consuming delays in production. Typically, this task is carried out manually, necessitating human intervention to photograph and label each ceramic piece. Consequently, employees must invest substantial effort and attention to detail, leading to potential bottlenecks in the production line. These issues are a cause for concern within various industries.

The monotonous nature of the task can also play a role in slowing down the production process. As humans repeat the same actions repeatedly, their efficiency can decrease, leading to longer processing times and ultimately causing delays.

Furthermore, the possibility of errors during manual image labeling is a common problem. When mistakes occur, it can result in confusion and wasted time as the mislabeled pieces need to be identified and corrected, further contributing to production delays.

This section aims to provide an overview of scenarios and techniques commonly employed in real ceramic industries helping in the understanding of the *CeramicFlow* framework and in the need for an automated approach.

3.1. Industrial scenario

In Fig. 2, a real-world representation of an industrial scenario is presented that highlights the critical need for an automated process for ceramic classification. Apart from the previously discussed challenges, this scenario poses numerous complex tasks related to image classification within the ceramic industry; even though all types of ceramic pieces are manufactured in the same environment using identical materials, generating a diverse dataset that comprehensively covers all possible variations in ceramic appearance presents significant difficulties. These variations include different textures, colors, and defects such as cracks or chips. Moreover, the images captured during this process may exhibit various lighting conditions and shadows, further complicating the classification task.

Additionally, there is the issue of data imbalance, as not all ceramic pieces are produced at the same rate or in similar quantities. This imbalance can impede the system's ability to effectively categorize new or distinct types of ceramic pieces, underscoring the importance of a fair and balanced dataset for accurate classification.

In this scenario, real-world ceramic pieces are used; they are transported using an industrial treadmill, which serves as the carrying system for the different distribution points. To capture high-quality images of the produced ceramic pieces, an industrial camera is positioned 56 cm above the carrying system. In terms of illumination, it utilizes the natural ambient light present in the environment and a spotlight pointed toward the surface of the treadmill, ensuring clear and well-defined images of the ceramic pieces.

3.2. Simulation scenario

To better understand the problem and conduct preliminary tests, we recreated the industrial scenario (Fig. 2) by developing a simulation scenario (Fig. 3). We ensured the authenticity of our setup by acquiring actual ceramic pieces from the ceramic industry, guaranteeing that the pieces utilized in our simulation scenario accurately represent those encountered in the industrial one. For the carrying system, we opted for black cardboard to simulate the background of an industrial setup. To capture the images, we employed an RGB camera positioned 56 cm above the carrying system. In terms of illumination, we aimed to replicate the lighting conditions of the industrial environment; thus, we utilized the natural ambient light present in the surroundings and we incorporated a spotlight with the same color temperature as the one used in the industrial scenario.

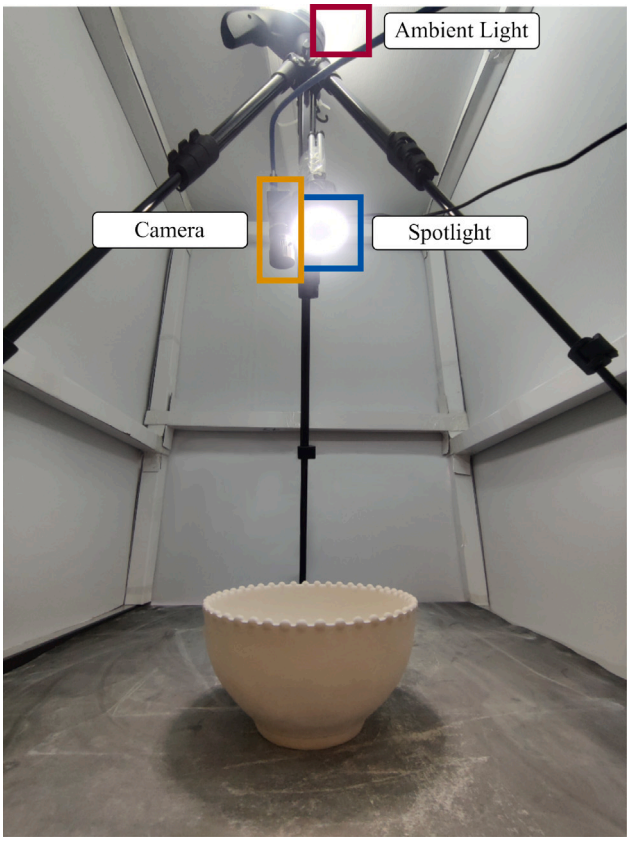


Fig. 3. Simulation scenario layout.

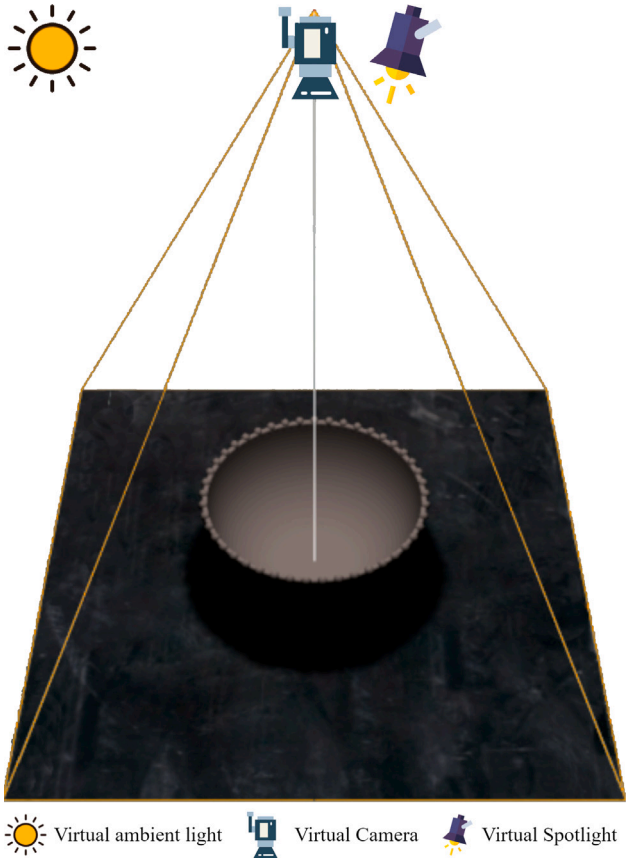


Fig. 4. *CeramicFlow* scene layout.

One of the challenges in capturing images in this synthetic scenario was the dust generated by the ceramic pieces settled on the 1 m by 1 m cardboard surface, which could potentially affect the quality of the images. To address this, the cardboard surface was managed and cleaned regularly to maintain similarity with the larger treadmill used in the industrial setting. Within the construction of the simulation model, it is essential to acknowledge the potential disparities that may arise between the simulated scenario and the authentic industrial setting. While both the industrial and simulation scenarios undergo regular maintenance to manage dust accumulation and diligent efforts have been made to replicate the camera characteristics, positions, illumination, and other conditions of the industrial setting, differences may persist.

Despite the nuanced disparities, this scenario remained a valuable source for preliminary testing of multiple aspects of this research.

4. Creating synthetic images — the *CeramicFlow* framework

To reduce reliance on real-world images in the ceramic industry, a virtual scenario named *CeramicFlow* was created with the objective of generating synthetic images of ceramic products, greatly reducing the dependence on extensive real-world image datasets as explained in Section 3.

CeramicFlow (Fig. 4) can be easily accessed from anywhere, at any time, and is low in resource demands, making it a practical solution for the industry engineers. It was built using Three.js,[1] a popular library that utilizes a Web Graphics Library to create 3D graphics-powered apps straight from web browsers that allow real-time rendering of complex 3D scenes, making it an ideal solution for the production of synthetic images.

[1] https://threejs.org/.

CeramicFlow is designed to compile a varied selection of images modeling various conditions common in industrial environments. This is accomplished by dynamically changing the characteristics of various elements using domain randomization techniques. It is crucial to point out that using these domain randomization approaches does not compromise data quality. Instead, it increases the dataset's diversity and composition through the inclusion of controlled variations within the ceramic environment's established standards. *CeramicFlow*'s necessary components for its functionality include firstly, the CAD models of ceramic pieces as the visual representation of the objects to be recognized and used for generating synthetic images. The carrying system is represented with a geometric plane designed to simulate a real industrial treadmill and, as usual, serves as the platform upon which the ceramic pieces are positioned.

Furthermore, a perspective camera is used and set up at a height that replicates real-world industrial conditions enabling the capture of images within the simulated scenario with a top-down perspective. *CeramicFlow* also incorporates virtual ambient lighting in the environment, along with a virtual spotlight that illuminates the surface of the treadmill.

The following subsections will elucidate CeramicFlow's approach in applying domain randomization techniques, as well as the proposed ceramic pipeline for a more automated process. The pipeline is structured into two main processes and further divided into multiple sections, each dedicated to serving a specific purpose.

4.1. Domain randomization

When generating synthetic data, the concept of the reality gap arises (Tobin et al., 2017a; Valtchev and Wu, 2021). This refers to the challenge of creating synthetic data that perfectly mimics real-world

Fig. 5. Domain randomization example of synthetic data in a fixed placement industry environment with *CeramicFlow*.

data. Achieving photorealism in synthetic data requires significant computational resources and time (Movshovitz-Attias et al., 2016). To narrow this reality gap and achieve desired results, *CeramicFlow* must generate a wide range of quality synthetic images according to industrial needs and user preferences, while introducing diversity to optimize the performance of DL models. To address this requirement, we have incorporated the concept of Domain Randomization (Tremblay et al., 2018; Tobin et al., 2017b). This technique helps create varied synthetic datasets that improve the robustness and generalization of the models.

In order to generate synthetic data for ceramic applications, it is essential to consider various factors, such as variations in color, texture, defects that exist among ceramic pieces, and others. As previously discussed, even when utilizing the same material to manufacture identical ceramics, individual pieces may have distinct characteristics. Furthermore, during the fabrication process of ceramics, the presence of dust is common, and it can settle on both the surfaces of the pieces and the carrying systems, ultimately affecting their appearance.

To start using *CeramicFlow*, the user is required to import the 3D file (*.obj* file extension) of the ceramic piece into the framework. This file contains the 3D model of the ceramic object, which serves as the foundation for generating synthetic images. Once the model is imported, the user has the option to apply textures to both the ceramic piece itself and to the treadmill on which it will be placed. The variation of treadmill textures serves a dual purpose: it not only allows for the realistic simulation of dust residues left on the ceramic but also helps prevent potential issues associated with a static background, which can lead to a mistrained model (Elhabian et al., 2010).

Moreover, users have the flexibility to choose from a diverse range of material types for the ceramic piece, including options for roughness, reflectivity, metalness, and more. They can fine-tune the illumination intensity and specify the position of the lighting source to achieve the desired visual effects. For added versatility, the dynamic lighting feature can be enabled, allowing the light source to dynamically change its position during the image generation process. This creates a variety of lighting conditions that simulate different times of the day, resulting in varying shadow effects as well (Fig. 5). Additionally, users must also choose how the division of the dataset is done, between training and validation data, according to their specific requirements.

In typical industry scenarios, there are two possibilities to consider with the placement of ceramics in the carrying system. Fixed placement where ceramic pieces are always positioned in the exact same place on the treadmill; or random placement where ceramic pieces are spread randomly along the treadmill. *CeramicFlow* has this into account, making it possible to create synthetic images for both types. In the first scenario, the position of the ceramic pieces is static and only its rotation can change. The second scenario involves placing the ceramic piece objects on top of the treadmill randomly and rotating it to create different perspectives.

The incorporation of randomization techniques, such as motion, rotation, surface attributes, materials, lighting changes, and ceramic defects, provides a viable route for generating a broad range of images. Ensuring the quality of synthetic images is vital in this challenge, as these images will be used as real-world virtual representations to train classification models used in the categorization of real-world production. Therefore, these images must be generated in controlled settings to ensure error-free data, which is achieved by maintaining precise information about all virtual components and scene details, such as bounding boxes. This procedure involves the setup of the camera, including optimal positioning and distance from the ceramic piece, and adjustment of the camera field of view to capture realistic perspectives that mimic industrial scenarios. Additionally, the ceramic element is randomly but accurately placed along positions of the geometric plane (representing the carrying system) within the defined capturing boundaries, ensuring complete object capture. Randomly incorporated elements, such as the ceramic piece and treadmill textures, are selected from a predefined set that approximates realistic appearances. Surface properties, including metalness, roughness, and reflectivity, are carefully adjusted to achieve accurate texture representation and lighting variations are meticulously controlled, resulting in shadows that closely approximate those in a real environment. Although generated in controlled settings, once the images are produced, they undergo a thorough visual inspection to ensure they meet the necessary quality standards and conditions.

By fine-tuning these characteristics these virtual environments can accurately represent real-world circumstances in accordance with the restrictions and expectations of ceramic manufacture. Because of this adaptability, they may accommodate the individual needs of different industries, resulting in the generation of high-quality datasets designed for real-world CV applications.

The pseudo-code for *CeramicFlow* can be found in Algorithm 1, providing a more comprehensive understanding of its behavior.

Algorithm 1 *CeramicFlow* process.

```
function GenerateSamples(numberOfSamples)
    imageArray ← []
    for i ← 1 to numberOfSamples do
        RandomTexture(ceramicPiece)
        RandomTexture(threadmill)
        if Type = Randomplacement then
            RandomMove(ceramicPiece)
        end if
        RandomRotation(ceramicPiece)
        RandomMove(ambientLight)
        imageArray[i] ← CaptureImage()
    end for
    DownloadDataset(imageArray)
end function
```

The combination of movement, rotation, textures, materials, light variations, and defects on the ceramics allows the creation of multiple, unique, and diverse virtual scenarios that closely emulate real-world conditions in the industry. The use of these techniques can result not only in more realistic images but also in the creation of a synthetic dataset for defect classification and tracking. To create more realistic synthetic images of the produced ceramic objects, effects like dust, and defects such as cracks and chips can be added to the model and to the virtual treadmill without affecting their physical properties. To do this firstly a UV mapping technique is used to map the 3D geometry of an object to a 2D texture representation, and then a normal mapping (a technique used to add surface detail to a 3D object without increasing the complexity of its underlying geometry) is applied on top of this one.

4.2. Creation of the ceramic classification model

The process shown in Fig. 6(a) depicts the required steps when introducing *CeramicFlow* for the first time in the industry or when introducing a new type of ceramic piece to the production line. The main steps are the following:

1. **Materials and Data Acquisition:** The Materials and Data Acquisition step involves utilizing CAD 3D models, used by the CAM to produce ceramic pieces. These 3D models serve as the input for *CeramicFlow* to generate multiple scenarios through domain randomization techniques. This technique enables the creation of new synthetic images, which are subsequently utilized for model training purposes.
2. **Model Training:** In this step, various NN architectures use the newly generated images from the previous step to develop a model capable of accurately classifying the different ceramic pieces in production. During this process, careful consideration is given to multiple hyperparameters, including the selection of the optimizer, determination of the optimal number of epochs, appropriate learning rate, and other factors.
3. **Model Creation:** After completing the preceding stages, the model exhibiting the highest performance is selected to undertake the classification task. The chosen model serves as the starting point for the production of the new ceramics.

4.3. Classification of ceramic pieces

After successfully completing the process detailed in Fig. 6(a), the classification model will be generated, enabling the classification of real-world ceramic pieces through the implementation of CV strategies as illustrated in Fig. 6(b).

By implementing this approach the ceramics industry can incorporate the production of new pieces alongside their existing product lines.

Consequently, the manufacturing of current products can proceed without interruption, while newly designed pieces can be quickly produced once a new model is applied, ensuring efficient and uninterrupted workflow. The main steps involved in this are:

1. **Materials and Data Acquisition:** Through the utilization of CAD 3D models and CAM, the industry can produce the required ceramic pieces. After being produced the ceramic pieces will be transported via the industrial treadmill, while simultaneous capture of images occurs as the products traverse along.
2. **Evaluation:** Using the DL model created in the first step (Fig. 6(a)) the images taken in the previous step will be classified;
3. **Distribution:** After successfully classifying the ceramics, they can be efficiently distributed to their respective destinations. This means sending them to ovens for baking, allocating them to designated storage areas, and delivering them to customers or other appropriate distribution channels based on their intended locations.

5. The *Synthetic CeramicNet* dataset

The demand for training ML algorithms with multiple images has led to the development of the *Synthetic CeramicNet* dataset. This comprehensive dataset comprises authentic images captured from various settings within the ceramic industry. It encompasses 12,000 2D images, each generated using *CeramicFlow*, featuring a resolution of 950 × 950 pixels and maintaining an aspect ratio of 1:1.

The *Synthetic CeramicNet* dataset offers a wide range of synthetic images that serve as valuable resources for testing and training diverse ML models. These images exhibit diverse perspectives of ceramic products, encompassing distinct textures, lighting conditions, and shadow effects. It classifies into four distinct types of ceramic pieces: PES161, PEP241, LSS151, and PEP282), as depicted in Fig. 7.

It employs a top-down perspective, commonly referred to as Bird's Eye View. This perspective provides a comprehensive overview of the objects' contours, shapes, and sizes, effectively facilitating easy differentiation among them. The dataset construction methodology involved the random placement of the ceramic pieces, adhering to the industry-standard practice in our industrial scenario.

Although derived from legal constraints there are no immediate plans to release the *Synthetic CeramicNet* dataset, we actively explore avenues to make it available to the research community in the future. In the meantime, there is the possibility of accessing the dataset upon request.

In the following sections, we will delve into the techniques employed for enhancing the quality and consistency of both synthetic and real-world datasets in Section 6. Additionally, we will discuss the evaluation metrics utilized to assess the performance of the synthetic images.

5.1. Preprocessing

In this section, we will delve into the preprocessing techniques presented in Fig. 8.

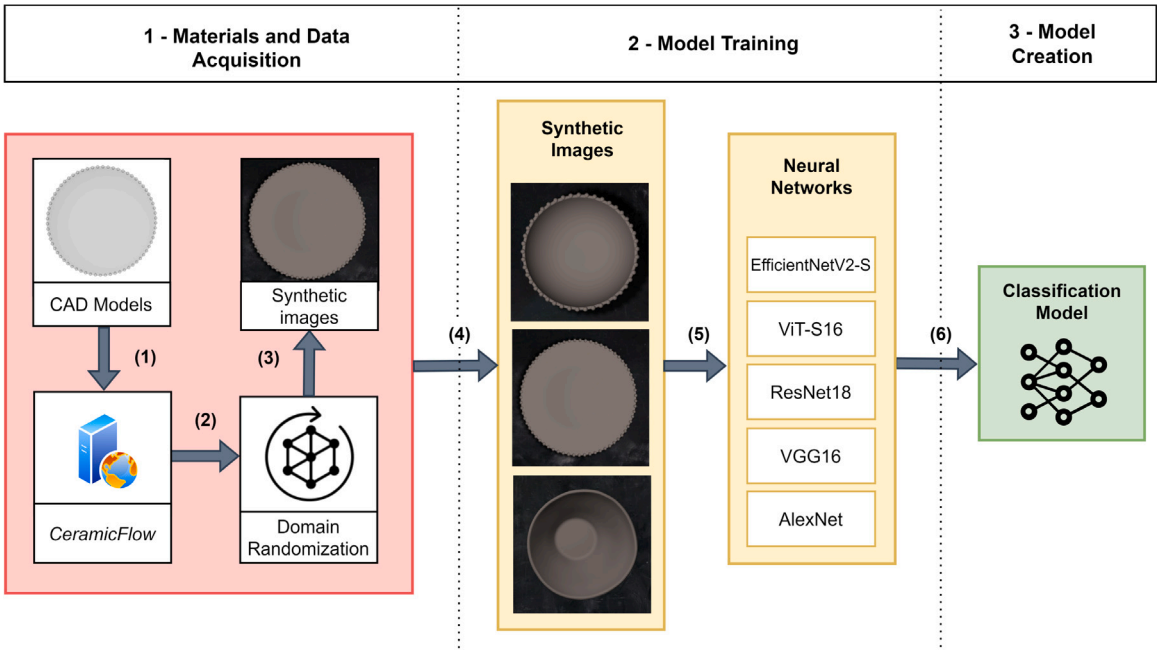
These techniques are used with the aim of maximizing the efficiency of data, ultimately culminating in the acquisition of optimal results within Section 6.

5.1.1. Region of Interest (ROI) selection

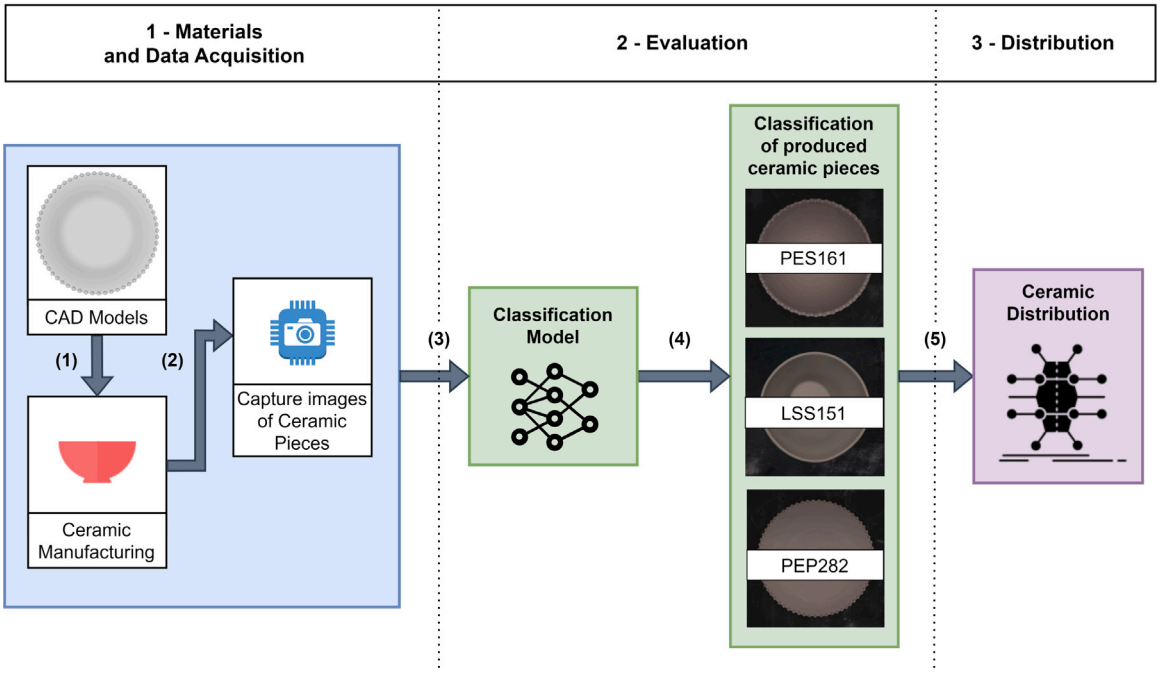
In the scenarios mentioned above, captured images often include a lot of empty space on the treadmill, making it difficult for the model to accurately classify the ceramic pieces. This step is therefore crucial in the process, as its integration in the preprocessing pipeline automatically eliminates any distracting information that could potentially mislead the ML model thus negatively affecting its accuracy. When the ROI is selected, the model can focus on the relevant features, resulting in improved recognition rates (Fig. 9) (Chu et al., 2012). To address this issue, an adaptive threshold algorithm is employed to automatically select the images and concentrate solely on the ceramic piece. This type of algorithm makes this process fast, efficient, and accurate, ensuring that the focus remains on the ceramic piece of interest, thus eliminating the need for manual selecting or even an AI-based image cropping algorithm, saving valuable time and resources.

In pursuit of accurate image cropping, multiple techniques involving the application of a mask to the image to facilitate ROI identification were employed (Fig. 10). The techniques included defining a global thresholding value, using an adaptive mean thresholding approach, and applying adaptive Gaussian thresholding (Mandyartha et al., 2020; Liao et al., 2020). The global thresholding method uses a single threshold value for the entire image to differentiate foreground objects from the background. However, this method may not be effective when there are variations in lighting conditions and background intensity across the image, leading to inaccurate image ROIs. The adaptive mean thresholding technique attempts to overcome this limitation by computing the threshold value for each pixel based on the mean intensity of its neighboring pixels. While this approach is an improvement over the global thresholding method, it still may not be effective in images with non uniform lighting conditions or complex backgrounds. In contrast, the adaptive Gaussian thresholding approach computes the threshold value locally for each pixel, taking into account the local intensity variation within the image, making this method more robust and generalizable to be applied to other production lines or environments.

For testing these different techniques, we opted to use the Intersection over Union (IoU) metric. The IoU metric measures how well the predicted ROI overlaps with the ground truth ROI by calculating



(a) Creation of the classification model.



(b) Classification of ceramic pieces.

Fig. 6. Process architecture: Synthetic image generation, modeling, and real-time classification architecture for ceramic production workflow.

the ratio of their overlapping area to their combined area showing how accurate the predicted regions match the ground truth ROI. The IoU score ranges from 0 to 1, where 1 indicates a perfect match. In preliminary tests, we used a set of images with the best possible manual cropping for comparison with the predicted ROI. The best performance was achieved using the adaptive Gaussian thresholding approach with an IoU of 0.91, followed by adaptive mean thresholding with an IoU of 0.86, and global thresholding with an IoU of 0.62.

Regarding the quality of the images, all images undergo visual inspection after this pre-processing phase to ensure they meet the necessary conditions for inclusion in the dataset. This includes verifying that the element to be classified is clearly visible and has not been affected by the ROI processing.

5.1.2. Data augmentation

While domain randomization can help increase the diversity of the training dataset, other types of data augmentation techniques may have a different impact on the model's robustness. Data augmentation (Perez and Wang, 2017; Shorten and Khoshgoftaar, 2019) is a powerful technique used in ML to increase the diversity and quantity of data available for training a model. The applied data augmentation techniques to understand the impact on the model's performance were the following:

Fig. 7. Ceramic piece types example images from left to right and top to bottom: PES161, PEP241, LSS151, and PEP282.

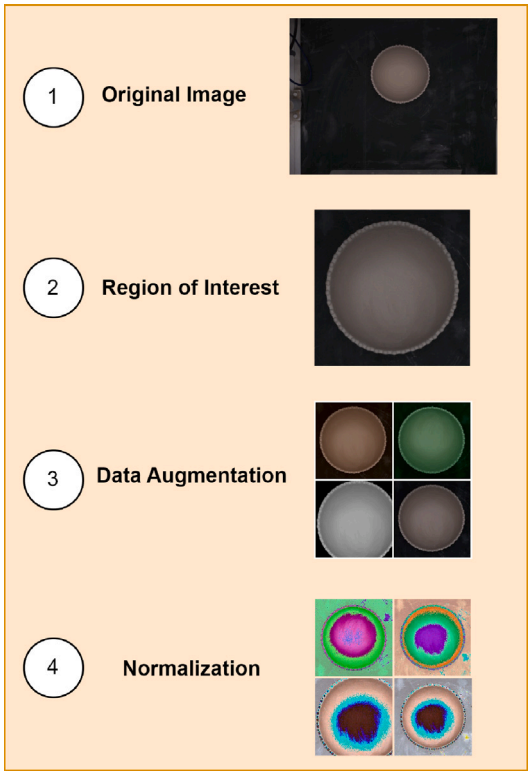


Fig. 8. Preprocessing techniques applied to the data.

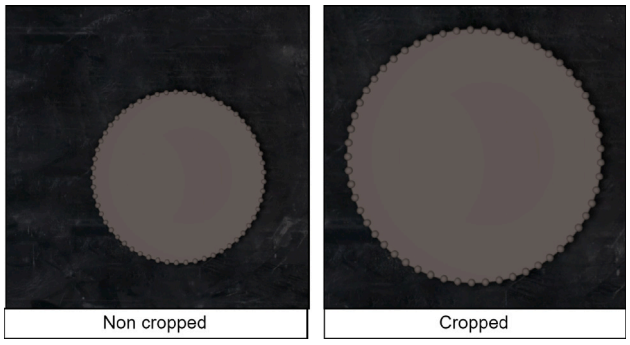


Fig. 9. ROI selection of synthetic image rendered in *CeramicFlow*.

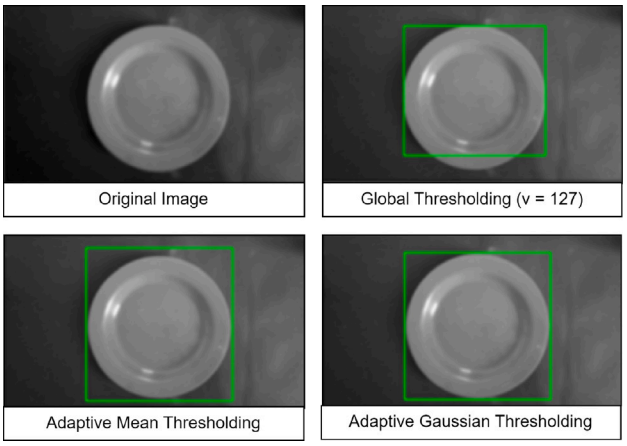


Fig. 10. Comparison of ROI selection techniques.

- **Horizontal and vertical flips** — Flipping the image can help the model learn to recognize ceramics that are oriented in different directions;
- **Random cropping** — Randomly selecting a portion of the image and resizing it to the original size can help recognize objects that are partially visible or occluded (which may occur in ceramic where the ceramic is randomly placed);
- **Color jitter** — Randomly adjusting the brightness, contrast, saturation, and hue of the image can help the model learn to recognize objects in different lighting conditions;
- **Rotation** — Rotating the image by a certain degree can help the model learn to recognize objects that are rotated.

5.1.3. Normalization

When using synthetic images, image normalization becomes critical. Synthetic images mimic real-world data, but may not fully represent the natural variations found in real images. Therefore, to ensure that the synthetic images can be used to classify real images accurately, they can be normalized to match the statistical properties of the real data (Ioffe and Szegedy, 2015).

To normalize the dataset we used the z-score normalization function (Eq. (1)) (Anggoro and Supriyanti, 2019).

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

This form of normalization includes subtracting the mean value of the dataset from each pixel value in the picture and then dividing the result by the estimated standard deviation of the dataset, producing a normalized picture with pixel values within a specific range Fig. 11.

5.2. Evaluating

A variety of measures were used to assess how well different DL architectures performed across various test cases using synthetic images from *Synthetic CeramicNet*. These metrics are frequently used to evaluate the performance of multiclass classification models, offering a comprehensive understanding of the model's performance by taking both accurate and incorrect predictions into account.

1. **Accuracy** (Acc.) — the percentage of correctly classified samples out of the total number of samples in the dataset (Eq. (2));

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

2. **Precision** (Prec.) — the percentage of true positives (correctly classified positive samples) out of all positive predictions (Eq. (3));

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

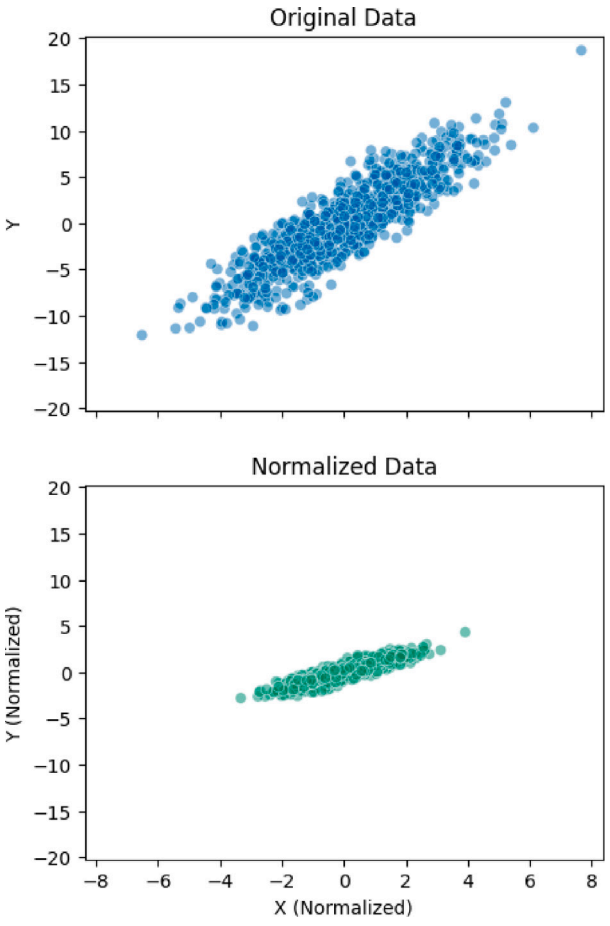


Fig. 11. Comparison of randomly generated 2D data before and after normalization.

3. **Recall** (Rec.) — the percentage of true positives out of all actual positive samples (Eq. (4));

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

where, TP means "True Positive", TN "True Negative", FP "False Positive", FN "False Negative" respectively.

4. **F1-score** (F1) — a weighted average of precision and recall, which provides a more balanced evaluation of a classifier's performance (Eq. (5)).

$$F_{score} = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{5}$$

6. Experimental results

In order to assess the effectiveness of our solution, we employed a DL-based approach for testing the neural network's response to synthetic images, classifying ceramic pieces across different domain conditions (de Melo et al., 2022; Tobin et al., 2017b). When using synthetic data to train classification models, there is a risk that the models may learn patterns and features specific to the synthetic domain. This can lead to poor generalization to real-world scenarios and different domains, as the model may not effectively translate synthetic patterns to real-world variations. Ensuring good generalization requires careful attention to the realism and representativeness of the training dataset. This could involve enhancing the quality and diversity of the synthetic data or incorporating a mix of synthetic and real data. This way, the study employed two different approaches to assess the performance of synthetic data:

- **Synthetic Modelization**: involves using only synthetic images generated by *CeramicFlow* as the training data, and only real-world images from the industrial scenario for testing (Section 6.1).
- **Hybrid Modelization**: adopted a hybrid model, incorporating real-world images (from the industrial scenario) and synthetic images for training (20% and 80% respectively) and only real-world images from the industrial scenario for testing (Section 6.2).

For the synthetic images, a distribution of the *Synthetic CeramicNet* datasets was created using only three classes (LSS151, PES161, and PEP282) due to the unavailability of a sufficient number of real-world images for the PEP241 ceramic class. Both modeling experiments consist of an overall training set of 1800 images, divided into 600 images for each class, however, each experiment uses different image domains (synthetic/real) to analyze the impact of using synthetic images in classification models. The use of 600 images per class is grounded in several factors. Insights from similar studies where similar amounts of data are employed; transfer learning from a large-scale dataset (ImageNet) means the model's weights are pre-optimized for general visual features and as the model only requires slight adjustments for the new task; the classification task involves only three classes, making it not requiring a large amount of training data; data augmentation generates new data from the existing training set, augmenting the diversity and robustness of the training even more (Chatterjee et al., 2022; Manettas et al., 2021; Bird et al., 2022).

For the classification tasks, several neural network designs were employed, including: AlexNet (Alom et al., 2018), EfficientNetV2-S (Tan and Le, 2021), ResNet (He et al., 2016), VGG16 (Simonyan and Zisserman, 2015) and ViT-S16 (Dosovitskiy et al., 2020). These models were chosen based on a set of criteria involving synthetic data in classification tasks:

- **Architectural Variety** — To represent a range of architectures from classical convolutional neural networks (CNNs) to modern transformer-based models. This includes simpler architectures like AlexNet and VGG16, more advanced and deeper CNNs like ResNet18 and EfficientNetV2-S, and a transformer-based model ViT-S16.
- **Performance History** — These models have shown strong performance in various benchmarks, making them suitable candidates for evaluating the efficacy of synthetic data in classification tasks (Öztürk and Erçelebi, 2021; Manettas et al., 2021; Zhu et al., 2023; Naeem et al., 2024).
- **Complexity and Depth** — The chosen models vary in terms of depth and complexity, which is critical for understanding how model complexity affects performance on synthetic data.

All of the employed NN architectures were pre-trained on the ImageNet dataset (Deng et al., 2009) and transfer learning techniques were used to fine-tune pre-trained models on our specific dataset. Transfer learning is an ML technique that leverages the knowledge acquired from a pre-trained model designed for a specific task to tackle a related task, resulting in faster training times and improved performance (Anvar and Mohammadi, 2023). The process involved using these pre-trained models with their pre-trained weights, modifying the final fully connected layer, and retraining it to output classifications for our three classes.

In practice, this meant that while the majority of the model's parameters remained unchanged, the final layers were adapted to the new task. This adaptation involved a combination of freezing earlier layers, which retain the learned features from the ImageNet dataset and fine-tuning the final layers to ensure the model accurately classified our dataset classes. This approach enabled the model to retain valuable pre-learned features while becoming specialized in the new classification task (features that are relevant to the task in the real-world can also be relevant to synthetic data Tremblay et al., 2018).

Table 1
Grid search hyperparameters.

Hyperparameters	Values
Pretraining dataset	ImageNet
Input size	224 × 224
Activation function	SoftMax
Loss function	Cross-entropy Loss
Optimizer	Adam
Batch size	32, 64, 128
Learning rate	0.001, 0.0001
Training datasets augmentations	Horizontal and vertical flips, Random cropping, Color jitter, Rotation
Normalization	z-score normalization function
ROI	Applied

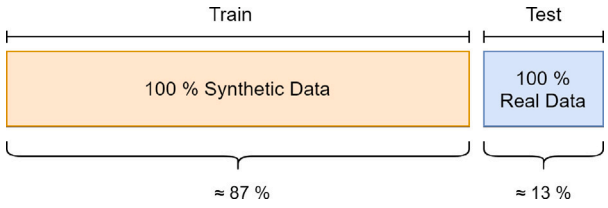


Fig. 12. Synthetic modelization experiment data distribution.

To enhance models results, we have followed a grid search technique (Liashchynskyi and Liashchynskyi, 2019) allowing us to explore multiple hyperparameters (Table 1) based on several studies (Öztürk and Erçelebi, 2021; Manettas et al., 2021; Kandel and Castelli, 2020). The SoftMax activation function was employed, and Cross-entropy Loss was used as the loss function. For optimization, the Adaptive Moment Estimation (Adam) optimizer (Kingma and Lei Ba, 2015) was employed. Although tests were also conducted with Stochastic Gradient Descent, Adam yielded better results. The values for this search ranged between predefined limits based on previous studies, specifically Batch size (BS) values of 32, 64, and 128, as well as Learning Rate (LR) values of 0.001 and 0.0001, aiming to determine optimal configurations for the model and enhance the overall quality of the data. Data augmentation was applied to the training sets, resulting in an average improvement of 1.7% in accuracy. Additionally, normalization was performed, yielding an average accuracy improvement of 0.7% based on earlier tests. Furthermore, all the images used in the experiments underwent the ROI selection method to ensure focus on the relevant features (Section 5.1.1).

The machine used in the test cases is composed of an AMD EPYC 7313P 16-core processor, an NVIDIA graphics card A100 80 GB, and 256 GB of RAM.

6.1. Synthetic modelization - Training using solely on synthetic data

In this test case, the objective is to investigate the effectiveness of using synthetic images as training data in scenarios in which no real-world data is available for the model's training (Fig. 12).

The training dataset comprises 600 synthetic images for each evaluated class. To assess the model's performance, the test dataset was formed by incorporating images from both industrial and simulation scenarios, resulting in 30 images per class, totaling 90 images. The rendering capabilities of *CeramicFlow* were finely tuned to generate images that closely resemble the testing dataset.

Table 2 presents the mean and standard deviation of results for all the employed metrics obtained by each of the utilized neural networks. As can be seen, the results reveal the effectiveness of synthetic images serving as strong evidence that precisely tuned NNs and synthetic data can effectively substitute the need to acquire real-world images in situations where time constraints or difficulties in obtaining such images arise for classification tasks.

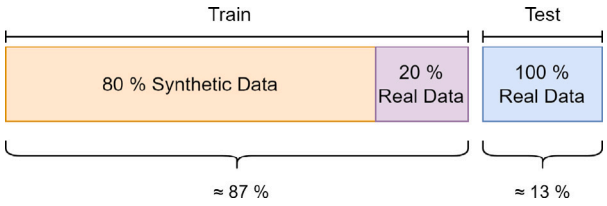


Fig. 13. Hybrid modelization experiment data distribution.

The results indicate that AlexNet, ResNet18, and VGG16 achieved remarkable results; and that ResNet18 stands out by performing excellently by correctly classifying all the test images. Additionally, EfficientNetV2-S demonstrated good performance in the classification task; ViT-S16 showed relatively lower accuracy and slightly worse overall results in comparison to the other models.

6.2. Hybrid modelization - Model training using synthetic and real-world data

Real-world images capture intricate details that are often uniform in synthetic images and can offer significant benefits when combined with synthetic data sources (Eversberg and Lambrecht, 2024; Kim et al., 2024). In synthetic datasets, for example, textures are consistent across certain parts of an object, varying only in lighting, shadows, and other environmental factors. While these synthetic images provide a valuable and scalable means of generating data, they can lack the natural variability and nuanced detail found in real-world datasets. In real-world ceramics, even pieces of the same class have different surface characteristics. Although made from the same materials, factors such as color composition, minor defects, and imperfections, as well as variations due to the manufacturing process, contribute to their unique appearance (Valtchev and Wu, 2021; Hinterstoisser et al., 2017).

With this test case, the objective is to investigate the effectiveness of using synthetic images as a data augmentation technique in scenarios where there is a limited number of real-world images, thus being insufficient for algorithm training (Fig. 13).

The determination of the ideal ratio between synthetic and real data in hybrid model training depends on several factors, including dataset size, data quality, and problem complexity. To address the challenge of limited data availability, we selected an 80–20 ratio of synthetic to real-world images aligning with the specific challenge of scarce real-world images in manufacturing, and based on multiple studies (Kim et al., 2024; Anderson et al., 2022; Khosravi et al., 2024) and tailored to the quantity of images we had available for testing.

Our goal was not to identify a universally ideal ratio, as the availability of ceramic images can vary significantly across different pieces. Instead, we aimed to leverage a larger proportion of synthetic data as the primary knowledge source while incorporating the limited amount of real-world data derived from this scarcity to enhance model robustness. This strategy ensures that the model benefits from the diversity of synthetic images while treating real-world data as another domain variation. As a result, the classification model's overall performance and generalization capabilities improve (Tobin et al., 2017b; Valtchev and Wu, 2021).

The training dataset counts with a combination of these synthetic and real-world images consisting of 600 images per class, including 480 synthetic images and 120 real-world images (distribution of 80% synthetic images and 20% real-world images). To ensure a fair comparison with the first test case (Section 6.1), the test dataset is the same as in the first test (composed of the same 90 images).

Table 3 presents the mean and standard deviation of results for all the employed metrics obtained by each of the utilized neural networks in Section 6.1. The results obtained indicate that the tested NN architectures performed exceptionally well on the classification task, where the

Table 2
Synthetic modelization test results.

NN	Acc. (%)	Test loss	Prec. (%)	Rec. (%)	F1 (%)
AlexNet	97.78 ± 4.06	0.39 ± 0.84	98.27 ± 3.05	97.16 ± 4.25	97.73 ± 4.17
EfficientNetV2-S	83.33 ± 3.14	0.41 ± 0.11	88.98 ± 1.45	88.66 ± 1.25	83.58 ± 3.14
ResNet18	100.00 ± 0.00	0.07 ± 0.05	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
VGG16	96.85 ± 1.97	0.13 ± 0.08	97.13 ± 1.74	98.00 ± 1.29	96.83 ± 1.97
ViT-S16	65.93 ± 7.22	1.52 ± 0.56	76.35 ± 12.45	76.72 ± 10.12	58.63 ± 8.18

Table 3
Hybrid modelization test results.

NN	Acc. (%)	Test loss	Prec. (%)	Rec. (%)	F1 (%)
AlexNet	100.00 ± 0.00	0.01 ± 0.01	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
EfficientNetV2-S	100.00 ± 0.00	0.07 ± 0.08	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
ResNet18	100.00 ± 0.00	0.01 ± 0.02	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
VGG16	100.00 ± 0.00	0.01 ± 0.01	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
ViT-S16	84.68 ± 8.24	0.36 ± 0.19	87.24 ± 6.89	86.39 ± 8.26	86.51 ± 10.09

combination of synthetic and real-world training data appears to have contributed to their high accuracy and effectiveness, achieving a perfect accuracy of 100% as can be seen in Table 3. Furthermore, we can see a major improvement from the tests made in the first case, and even though some models already had achieved an accurate classification for all test images (100%); by comparing the loss values, we can infer that the addition of real-world images helped to reduce the overall error between the predicted and actual values. It is important to note that the results from the test using only synthetic data were not necessarily poor; in fact, the models achieved high accuracy rates. However, the incorporation of real-world images acted as a complement, enhancing the performance of the models in situations where data are scarce and where traditional training algorithms, which rely on large quantities of data, would not yield satisfactory results.

This demonstrates that real-world data combined with synthetic data as a data augmentation technique can be an effective strategy for training DL models, leading to better generalization and enhanced accuracy in real-world tasks.

6.3. Discussion

Overall, the findings acquired from the two test cases provide compelling evidence supporting the use of synthetic images as training data and their complementary role alongside real-world data. As can be seen in Table 2, taking the example of EfficientNetV2-S, by employing a dataset built solely with basis on synthetic images a high accuracy of 83% is achieved; also, and more significantly, when employing a hybrid dataset, which mixes real and synthetic images (in a proportion of 20/80), a 100% rate was attained. This result highlights not only the potential of using synthetic images as a viable alternative in the absence of real images (which is a common scenario in the industry), but also their adequacy to augment real-world datasets that contain limited real-world data.

Although synthetic data alone has demonstrated exceptional performance in the classification task (Section 6.1), with certain NN already achieving 100% accuracy, there is still room for improvement: by integrating synthetic data with real-world images as a data augmentation technique (Section 6.2), we can enhance the results even further. This combination becomes a highly valuable resource in augmenting the models' capabilities to generalize, adapt, and exceed our expectations, as evidenced in Table 3.

In other experiments, we also tested the DL models using the same NN and hyperparameters used both in Sections 6.1 and 6.2 but only with the few real-world images employed in the hybrid approach as the training set. We found that the problem could be solved using only the original real-world images, achieving better results than training with synthetic data alone (synthetic modelization), but worse than the results achieved by using the hybrid modelization. However, as mentioned, this method would still require significant time for image acquisition and labeling, additionally, the smaller quantity of images and the potential for imbalanced datasets – due to the varying availability of ceramic images across different pieces – make the model highly susceptible to overfitting, potentially compromising its generalization capabilities.

Table 4
Test cases overall comparison.

NN	Synthetic modelization		Hybrid modelization	
	Acc. (%)	Test loss	Acc. (%)	Loss
AlexNet	97.78 ± 4.06	0.39 ± 0.84	100.00 ± 0.00	0.01 ± 0.01
EfficientNetV2-S	83.33 ± 3.14	0.41 ± 0.11	100.00 ± 0.00	0.07 ± 0.08
ResNet18	100.00 ± 0.00	0.07 ± 0.05	100.00 ± 0.00	0.01 ± 0.02
VGG16	96.85 ± 1.97	0.13 ± 0.08	100.00 ± 0.00	0.01 ± 0.01
ViT-S16	65.93 ± 7.22	1.52 ± 0.56	84.68 ± 8.24	0.36 ± 0.19

Given the rapid pace of manufacturing in the ceramics industry and the increasing demand, it is critical that all operations keep pace, especially when manufacturers introduce new designs, such as bespoke pieces, where the production line must adapt quickly. Our results indicate that while a relatively low quantity of real-world images does provide good outcomes, the synthetic data approach offers significant advantages in terms of scalability and immediate applicability in an industrial setting. This way a synthetic approach could be particularly useful in the initial phase of introducing new ceramic models and as more images were acquired, a hybrid approach can be adopted, given that it yields better outcomes than using real images alone. These outcomes demonstrate that synthetic data can be highly effective in achieving good results while also facilitating the automation and streamlining of the production process.

The observed disparities in the performance of different neural network architectures under both synthetic and hybrid modelization (Table 4) can be attributed to several key factors, including architectural differences, capacity to generalize from synthetic to real data, and the use of real-world data in training.

Firstly, architectural differences play a significant role. ResNet18 exhibited the highest performance across all metrics in both synthetic and hybrid approaches. Its residual connections facilitate effective learning and generalization by mitigating issues like vanishing gradients, making the model robust in transferring knowledge from synthetic to real images. AlexNet and VGG16, despite their simpler architectures, managed to capture relevant features effectively due to their significant number of layers, which are well-suited for extracting spatial features in image data. These models showed perfect accuracy scores in the hybrid modelization, demonstrating their robustness in handling a mix of synthetic and real data. In contrast, models like EfficientNetV2-S and ViT-S16 displayed varying levels of performance. While EfficientNetV2-S performed excellently in the hybrid modelization, it exhibited lower accuracy and higher test loss in the synthetic approach, indicating

a need for more real-world data or precise tuning. ViT-S16 showed the lowest performance in both approaches, with significant drops in accuracy and other metrics, suggesting that its complex architecture may require extensive fine-tuning or larger amounts of real-world data to perform optimally.

Apart from that, in terms of hyperparameters, we concluded that a LR of 0.0001 stood out in terms of performance. However, there were very few differences between the configurations of BS 32, 64, and 128, with no configuration significantly outperforming the others.

With *CeramicFlow*, synthetic data becomes an easy-to-get, powerful option that can effectively simulate and capture the characteristics of real-world data, enabling the creation of diverse and representative training datasets for use in a wide range of areas.

The findings align with a growing body of research emphasizing the potential of synthetic data in ML applications. Similar studies in fields such as healthcare, in which patient data privacy is a concern, have successfully employed synthetic medical images for training diagnostic models (Chen et al., 2021; McDuff et al., 2023). In robotics, using synthetic data has also been instrumental in training robots for various tasks, demonstrating the versatility and effectiveness of synthetic data in real-world applications (Lips et al., 2022; Lambrecht and Kästner, 2019).

While other industries have explored synthetic data for different applications (He and Zhou, 2024; Venkataramanan et al., 2023; Ekbatani et al., 2017), in the context of ceramics and product classification our work stands as a new approach leveraging synthetic data to address the challenges of data scarcity.

In summary, the research highlights the potential of using synthetic data in classification problems, both on its own and in conjunction with real-world data, serving as an effective data augmentation technique. This approach can significantly benefit industries such as ceramics by accelerating production processes, reducing costs, and improving classification accuracy, making it a notable contribution to the broader field of ML and its practical applications.

7. Conclusion and future work

The lack of data in industries poses a significant challenge; one that hinders dataset building and, consequently, the development and deployment of effective ML models and systems. Demand for high-quality images that enable DL classification mechanisms is particularly high, as was verified, in loco, when working with the ceramic industry: real-world images are scarce, especially at the beginning of the production process as, at this stage, no ceramic pieces exist that can be photographed and labeled.

We propose tackling this problem by strategically generating high-quality, synthetic images adapted for the ceramic industry, by means of a process that involves: the development of a computer graphics rendering pipeline designed for the creation of synthetic data; the creation of a synthetic dataset for the purpose of image classification in the ceramic industry; the validation of the proposed approach by training state-of-the-art DL models using the synthetic dataset, and assessing their performance when classifying real-world industrially manufactured ceramic images obtained in a production environment. We emphasize the significance of only requiring the ceramic pieces' CAD files used in the industry for their production — unlike alternative approaches that continue to depend heavily on a significant volume of real-world data to generate synthetic images.

The experimental studies provided compelling evidence for supporting the use of synthetic images as the basis for dataset building: models trained using solely synthetic images yielded remarkable results; only surpassed by the performance of models trained with hybrid datasets, i.e., those composed by real-world and synthetic images. These results demonstrate the adequacy of employing synthetic images, both for data augmentation and whole dataset building — lowering the dependency on real-world images and leading to more efficient production lines. Nevertheless, the integration of real data, even if in a reduced ratio, proved to be relevant for capturing more intricate characteristics.

The unavailability of real-world ceramic images due to the inherent challenges of the problem constrained the feasibility of some experiments. Future work involves collaborating with industry partners to gather more real-world ceramic images, allowing us to explore larger datasets, expand class diversity, and test different ratios to find the optimal balance for various applications and DL models. Furthermore, we aim to train models using synthetic images without transfer learning to assess their generalization capabilities when tested against real-world images. We plan to investigate few-shot classification techniques to evaluate how well models can learn from a limited number of real-world examples compared to synthetic data approaches. In addition, we also expect to obtain results of ongoing research on embedding depth information into the generated synthetic images — which, we believe, will prove to be a crucial enhancement to the dataset and result in more efficient classification models.

CRediT authorship contribution statement

Fábio Gaspar: Writing – original draft, Software, Methodology, Investigation, Conceptualization. **Daniel Carreira:** Writing – review & editing, Visualization, Software. **Nuno Rodrigues:** Writing – review & editing, Supervision, Resources, Project administration, Data curation, Conceptualization. **Rolando Miragaia:** Writing – review & editing, Supervision, Investigation, Formal analysis, Data curation, Conceptualization. **José Ribeiro:** Writing – review & editing, Validation, Project administration, Data curation, Conceptualization. **Paulo Costa:** Writing – review & editing, Validation, Supervision, Data curation. **António Pereira:** Writing – review & editing, Visualization, Supervision, Resources, Project administration, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

• This work was supported by national funds through the Portuguese Foundation for Science and Technology (FCT), I.P., under the project UIDB/04524/2020.

• STC 4.0 HP - New Generation of Stoneware Tableware in Ceramic 4.0 by High Pressure Casting Robot work cell, Referência: Nº 69654, I&DT Empresarial (Copromoção, Parcerias Internacionais).

• I DECOR - I & D sistema avançado de aplicação de decorações em tableware – Controlo de Qualidade (Refª Candidatura: C679416409-00009762, Refª Submissão: T679994395-00003392).

Data availability

Data will be made available on request.

References

Abu Alhaija, H., Mustikovela, S.K., Mescheder, L., Geiger, A., Rother, C., 2018. KITTI segmentation. Int. J. Comput. Vis. 126 (9), 961–972.

Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Essen, B.C.V., Awwal, A.A.S., Asari, V.K., 2018. The history began from AlexNet: A comprehensive survey on deep learning approaches. CoRR abs/1803.01164.

Anderson, J.W., Ziolkowski, M., Kennedy, K., Apon, A.W., 2022. Synthetic image data for deep learning. arXiv:2212.06232 URL https://arxiv.org/abs/2212.06232.

Anggoro, D.A., Supriyanti, W., 2019. Improving accuracy Bb applying Z-score normalization in linear regression and polynomial regression model for real estate data. Int. J. Emerg. Trends Eng. Res. 7 (11), 549–555.

Anvar, A.A.T., Mohammadi, H., 2023. A novel application of deep transfer learning with audio pre-trained models in pump audio fault detection. Comput. Ind. 147, 103872.

Aranha, C., Kiyoiti, F.H., Tanaka, S., Lee, W.S., Suzuki, T., 2019. Data augmentation using GANs new approaches for multi objective optimization view project application of evolutionary algorithms to reservoir history matching view project data augmentation using GANs. Proceedings of Machine Learning Research XXX, 1–16.

Barth, R., Hemming, J., Van Henten, E., 2020. Optimising realism of synthetic images using cycle generative adversarial networks for improved part segmentation. Comput. Electron. Agric. 173, 105378.

Bird, J.J., Barnes, C.M., Manso, L.J., Ekárt, A., Faria, D.R., 2022. Fruit quality and defect image classification with conditional GAN data augmentation. Sci. Hort. 293, 110684.

Chai, J., Zeng, H., Li, A., Ngai, E.W., 2021. Deep learning in computer vision: A critical review of emerging techniques and application scenarios. Mach. Learn. Appl. 6, 100134.

Chatterjee, S., Hazra, D., Byun, Y.-C., Kim, Y.-W., 2022. Enhancement of image classification using transfer learning and GAN-based synthetic data augmentation. Mathematics 10 (9).

Chen, R.J., Lu, M.Y., Chen, T.Y., Williamson, D.F., Mahmood, F., 2021. Synthetic data in machine learning for medicine and healthcare. Nat. Biomed. Eng. 5 (6), 493–497.

Chu, C., Hsu, A.L., Chou, K.H., Bandettini, P., Lin, C.P., 2012. Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. Neuroimage 60 (1), 59–70.

Dawson, H.L., Dubrule, O., John, C.M., 2023. Impact of dataset size and convolutional neural network architecture on transfer learning for carbonate rock classification. Comput. Geosci. 171, 105284.

de Melo, C.M., Torralba, A., Guibas, L., DiCarlo, J., Chellappa, R., Hodgins, J., 2022. Next-generation deep learning based on simulators and synthetic data. Trends in Cognitive Sciences 26, 174–187.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. ImageNet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. pp. 248–255. http://dx.doi.org/10.1109/CVPR.2009.5206848.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2020. An image is worth 16x16 words: Transformers for image recognition at scale. CoRR abs/2010.11929.

Ekbatani, H.K., Pujol, O., Segui, S., 2017. Synthetic data generation for deep learning in counting pedestrians.. In: ICPRAM. pp. 318–323.

Elhabian, S.Y., El-Sayed, K.M., Ahmed, S.H., 2010. Moving object detection in spatial domain using background removal techniques - state-of-art. Recent Patents Comput. Sci. 1 (1), 32–54.

Eversberg, L., Lambrecht, J., 2024. Combining synthetic images and deep active learning: Data-efficient training of an industrial object detection model. J. Imaging 10 (1).

Gaidon, A., Lopez, A., Perronnin, F., 2018. The reasonable effectiveness of synthetic visual data. Int. J. Comput. Vis. 126 (9), 899–901.

Gaidon, A., Wang, Q., Cabon, Y., Vig, E., 2016. Virtual worlds as proxy for multi-object tracking analysis. CoRR abs/1605.06457.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Vol. 2016-December, pp. 770–778.

He, Z., Zhou, W., 2024. Development of machine learning-based burst capacity models for pipelines containing dent-gouges with synthetic full-scale burst test data generated using tabular generative adversarial network. Eng. Appl. Artif. Intell. 133, 108090.

Hinterstoisser, S., Lepetit, V., Wohlhart, P., Konolige, K., 2017. On pre-trained image features and synthetic images for deep learning. arXiv:1710.10710 URL https://arxiv.org/abs/1710.10710.

Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: 32nd International Conference on Machine Learning, ICML 2015. Vol. 1, pp. 448–456.

Iwasaki, M., Yoshioka, R., 2019. Data augmentation based on 3D model data for machine learning. 2019 IEEE 4th International Conference on Computer and Communication Systems, ICCCS 2019 1–4.

Jain, S., Seth, G., Paruthi, A., Soni, U., Kumar, G., 2022. Synthetic data augmentation for surface defect detection and classification using deep learning. J. Intell. Manuf. 33 (4), 1007–1020.

Kandel, I., Castelli, M., 2020. The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. ICT Express 6 (4), 312–315.

Khosravi, B., Li, F., Dapamede, T., Rouzrokh, P., Gamble, C.U., Trivedi, H.M., Wyles, C.C., Sellergren, A.B., Purkayastha, S., Erickson, B.J., Gichoya, J.W., 2024. Synthetically enhanced: unveiling synthetic data's potential in medical imaging research. eBioMedicine 104, 105174.

Kim, J., Wang, I., Yu, J., 2024. Experimental study on using synthetic images as a portion of training dataset for object recognition in construction site. Buildings 14 (5).

Kingma, D.P., Lei Ba, J., 2015. 15Iclr-ADAM. Iclr 1–15.

Lambrecht, J., Kästner, L., 2019. Towards the usage of synthetic data for marker-less pose estimation of articulated robots in RGB images. In: 2019 19th International Conference on Advanced Robotics. ICAR, pp. 240–247. http://dx.doi.org/10.1109/ICAR46387.2019.8981600.

Liao, J., Wang, Y., Zhu, D., Zou, Y., Zhang, S., Zhou, H., 2020. Automatic segmentation of crop/background based on luminance partition correction and adaptive threshold. IEEE Access 8, 202611–202622.

Liashchynskyi, P., Liashchynskyi, P., 2019. Grid search, random search, genetic algorithm: A big comparison for NAS. CoRR abs/1912.06059.

Lips, T., Gusseme, V.-L.D., wyffels, F., 2022. Learning keypoints from synthetic data for robotic cloth folding. arXiv:2205.06714.

Mahmood, F., Chen, R., Durr, N.J., 2018. Unsupervised reverse domain adaptation for synthetic medical images via adversarial training. IEEE Trans. Med. Imaging 37 (12), 2572–2581.

Mandyartha, E.P., Anggraeny, F.T., Muttaqin, F., Akbar, F.A., 2020. Global and adaptive thresholding technique for white blood cell image segmentation. J. Phys. Conf. Ser. 1569 (2), 022054.

Manettas, C., Nikolakis, N., Alexopoulos, K., 2021. Synthetic datasets for deep learning in computer-vision assisted tasks in manufacturing. Procedia CIRP 103, 237–242, 9th CIRP Global Web Conference – Sustainable, resilient, and agile manufacturing and service operations : Lessons from COVID-19.

McDuff, D., Curran, T., Kadambi, A., 2023. Synthetic data in healthcare. arXiv:2304.03243.

Melo, C.M.D., Rothrock, B., Gurram, P., Ulutan, O., Manjunath, B.S., 2020. Vision-based gesture recognition in human-robot teams using synthetic data. In: IEEE International Conference on Intelligent Robots and Systems. Institute of Electrical and Electronics Engineers Inc., pp. 10278–10284.

Movshovitz-Attias, Y., Kanade, T., Sheikh, Y., 2016. How useful is photo-realistic rendering for visual learning? In: Computer Vision–ECCV 2016 Workshops: Amsterdam, the Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part III 14. Springer, pp. 202–217.

Müller, M., Casser, V., Lahoud, J., Smith, N., Ghanem, B., 2018. Sim4CV: A photo-realistic simulator for computer vision applications. Int. J. Comput. Vis. 126 (9), 902–919.

Naeem, S., Al-Sharawi, R., Khan, M.R., Tariq, U., Dhall, A., Al-Nashash, H., 2024. Real, fake and synthetic faces – does the coin have three sides?. arXiv:2404.01878 URL https://arxiv.org/abs/2404.01878.

Nikolenko, S.I., 2019. Synthetic data for deep learning. CoRR abs/1909.11512.

Osinski, B., Jakubowski, A., Ziecina, P., Milos, P., Galias, C., Homoceanu, S., Michalewski, H., 2020. Simulation-based reinforcement learning for real-world autonomous driving. In: Proceedings - IEEE International Conference on Robotics and Automation. pp. 6411–6418.

Öztürk, A.E., Erçelebi, E., 2021. Real uav-bird image classification using cnn with a synthetic dataset. Appl. Sci. (Switzerland) 11 (9).

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E.Z., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. PyTorch: An imperative style, high-performance deep learning library. CoRR abs/1912.01703.

Perez, L., Wang, J., 2017. The effectiveness of data augmentation in image classification using deep learning. CoRR abs/1712.04621.

Rather, I.H., Kumar, S., 2024. Generative adversarial network based synthetic data training model for lightweight convolutional neural networks. Multimedia Tools Appl. 83 (2), 6249–6271.

Ren, Z., Lee, Y.J., 2017. Cross-domain self-supervised multi-task feature learning using synthetic imagery. CoRR abs/1711.09082.

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. J. Big Data 6 (1).

Shrivastava, A., Pfister, T., Tuzel, O., Susskind, J., Wang, W., Webb, R., 2016. Learning from simulated and unsupervised images through adversarial training. CoRR abs/1612.07828.

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. pp. 1–14.

Solovyev, R., Kalinin, A.A., Gabruseva, T., 2022. 3D convolutional neural networks for stalled brain capillary detection. Comput. Biol. Med. 141 (November 2021), 105089.

Spindler, A., Geach, J.E., Smith, M.J., 2020. AstroVaDEr: astronomical variational deep embedder for unsupervised morphological classification of galaxies and synthetic image generation. Mon. Not. R. Astron. Soc. 502 (1), 985–1007.

Sun, C., Shrivastava, A., Singh, S., Gupta, A., 2017. Revisiting unreasonable effectiveness of data in deep learning era. abs/1707.02968.

Tan, M., Le, Q.V., 2021. EfficientNetV2: Smaller models and faster training. CoRR abs/2104.00298.

Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., Abbeel, P., 2017a. Domain randomization for transferring deep neural networks from simulation to the real world. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS, pp. 23–30. http://dx.doi.org/10.1109/IROS.2017.8202133.

Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., Abbeel, P., 2017b. Domain randomization for transferring deep neural networks from simulation to the real world. CoRR abs/1703.06907.

Tremblay, J., Prakash, A., Acuna, D., Brophy, M., Jampani, V., Anil, C., To, T., Cameracci, E., Boochoon, S., Birchfield, S., 2018. Training deep networks with synthetic data: Bridging the reality gap by domain randomization. CoRR abs/1804.06516.

Valtchev, S.Z., Wu, J., 2021. Domain randomization for neural network classification. J. Big Data 8 (1), 94.

Venkataramanan, A., Faure-Giovagnoli, P., Regan, C., Heudre, D., Figus, C., Usseglio-Polatera, P., Pradalier, C., Laviale, M., 2023. Usefulness of synthetic datasets for diatom automatic detection using a deep-learning approach. Eng. Appl. Artif. Intell. 117, 105594.

Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E., 2018. Deep learning for computer vision: A brief review. Comput. Intell. Neurosci. 2018.

Xu, K., Kong, X., Wang, Q., Han, B., Sun, L., 2023. Intelligent fault diagnosis of bearings under small samples: A mechanism-data fusion approach. Eng. Appl. Artif. Intell. 126, 107063.

Xu, K., Kong, X., Wang, Q., Yang, S., Huang, N., Wang, J., 2022. A bearing fault diagnosis method without fault data in new working condition combined dynamic model with deep learning. Adv. Eng. Inform. 54, 101795.

Zhu, X., Bilal, T., Mårtensson, P., Hanson, L., Björkman, M., Maki, A., 2023. Towards Sim-to-Real industrial parts classification with synthetic dataset. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. CVPRW, pp. 4454–4463. http://dx.doi.org/10.1109/CVPRW59228.2023.00468.