

# A convolutional neural network approach to classifying urban spaces using generative tools for data augmentation

International Journal of  
Architectural Computing  
2024, Vol. 0(0) 1–20  
© The Author(s) 2024



Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/14780771231225697  
[journals.sagepub.com/home/jac](https://journals.sagepub.com/home/jac)



Carlos Medel-Vera<sup>1</sup> , Pelayo Vidal-Estévez<sup>2</sup> and Thomas Mädler<sup>3</sup> 

## Abstract

This article discusses an application for classifying urban spaces using convolutional neural networks (CNNs). A seed dataset was initially generated composed of 630 photographs of urban spaces from the Adobe Stock repository. This dataset was topped up with images produced by two generative artificial intelligence (AI) engines, namely, Deep Dream Generator and Midjourney, making two additional augmented datasets, each composed of 2200 images. The training process was carried out using four well-known CNNs, namely, GoogLeNet, ResNet-18, ShuffleNet, and MobileNet-v2. The results show an increase of roughly 30% in the predicting capabilities in both augmented datasets when compared to the seed dataset. Furthermore, performance metrics are generally higher when using ResNet-18 which may suggest that this CNN architecture is more applicable to urban classification projects. Finally, although both generative AI engines have similar performance, Midjourney seems to slightly outperform Deep Dream Generator as a data augmentation engine for urban spaces.

## Keywords

Urban categories, classification, machine learning, deep learning, diffusion models, neural architecture

## Introduction

Urban spaces play a vital role in shaping the livability, functionality, and sustainability of cities. They are designed for human use and interaction and, as such, are crucial for the well-being of individuals and communities.<sup>1,2</sup> Thus, the classification of urban spaces becomes a relevant topic for policymakers and city

<sup>1</sup>School of Architecture, University of Liverpool, Liverpool, UK

<sup>2</sup>School of Civil Engineering, Faculty of Engineering and Sciences, Universidad Diego Portales, Santiago, Chile

<sup>3</sup>School of Civil Engineering and Institute of Astrophysical Studies, Faculty of Engineering and Sciences, Universidad Diego Portales, Santiago, Chile

## Corresponding author:

Carlos Medel-Vera, School of Architecture, University of Liverpool, 25 Abercromby Square, Liverpool L697ZG, UK.

Email: [carlos.medel-vera@liverpool.ac.uk](mailto:carlos.medel-vera@liverpool.ac.uk)

planners in order to make informed investment decisions to best serve the function of interaction in cities.<sup>3</sup> Several attempts to classify urban spaces have been published in the literature<sup>4-6</sup> depending on a wide range of characteristics, such as (i) design-functional perspectives, (ii) socio-cultural perspectives, and (iii) political-economic perspectives. Design-functional perspectives classify spaces according to their intended purpose, characteristics, and physical layout,<sup>7,8</sup> for example, public parks, markets, and playgrounds. Socio-cultural perspectives classify spaces according to the type of individuals who utilize them and how they interact with them,<sup>9</sup> for example, everyday places, social environments, and negative spaces. Political-economic perspectives classify spaces according to the ownership, responsibilities, or management associated to them,<sup>10,11</sup> for example, public property, secured public space, and themed public space. Even with these varied perspectives, it is widely accepted in the literature<sup>12</sup> that the increasing complexity of urban spaces makes their classification a somewhat complicated task. In this light, machine learning models are ideal candidates to facilitate such a process due to their capacity to learn patterns based on analyzing large amounts of data. This article deals with the classification of urban spaces as described in Carmona,<sup>4</sup> who defines 20 types that, to some extent, are a combination of all abovementioned perspectives, using four overarching categories that move in a highly continuous manner from clearly public spaces to clearly private spaces.

Convolutional neural networks (CNNs) are a type of deep learning model that are specifically designed for image and video analysis based on the concept of convolution, which is a mathematical operation that is used to extract features from an image.<sup>13</sup> They can learn hierarchical representations of the input data by successively applying convolutional and pooling layers to the input image, which eventually allows the model to determine whether or not the input image falls into a given category or class. CNNs are particularly relevant in data science, as they have demonstrated considerable effectiveness in a wide range of image and video analysis tasks, such as object recognition, object detection, and semantic segmentation, as well as in other fields, such as natural language processing, speech recognition, and medical image analysis.<sup>14,15</sup> The application of CNNs in architecture and urban design is a relatively recent field, though it has gained increasing attention in research and practice.<sup>16,17</sup> In architectural design, CNNs have been applied building façades for the classification of their Gestalt principles,<sup>18</sup> architectural styles,<sup>19</sup> and their components.<sup>20</sup> Other examples in urban design include the classification of land use and land cover types, such as residential, commercial, industrial, and green spaces,<sup>21-23</sup> recognition of different street-level features by means of street-view image analysis, such as urban versus non-urban settlements,<sup>24</sup> quality of street frontage,<sup>25</sup> and vegetation.<sup>26</sup> The use of CNNs in this context is accomplished by training a model on a relatively large dataset of labelled images, where the model learns to recognize different patterns and features associated with the classes specified in the dataset.

For meaningful deep learning models such as CNNs, generally large datasets must be used for training in order to allow generalization of results and avoid overfitting issues; a well-known technique to address these potential problems is data augmentation.<sup>27</sup> Some popular data augmentation techniques include (i) geometric transformations, such as flipping, cropping, rotation, and noise injection, and (ii) generative-based models, such as generative adversarial networks (GANs)<sup>28,29</sup> and diffusion models (DMs),<sup>30,31</sup> where artificial instances are generated in an attempt to emulate similar characteristics of the dataset needing augmentation.

GANs are basically composed of two sub-modules: a generator and a discriminator. The generator, given a vector of random values, generates data with the same structure as the training data; the discriminator attempts to classify data as real or fake when shown observations from both the training data and data coming from the generator.<sup>32</sup> GAN-based augmentations have been used previously to enlarge datasets for deep learning models in a wide range of fields, most notably in medical sciences research.<sup>33-35</sup> DMs, on the other hand, follow a different path: during training, they progressively destruct data by injecting noise, which allows them to then learn how to reverse this process. In this way, new instances can then be generated from just random noise as a starting point.<sup>30</sup> Text-to-image models are specific DMs that generate images from a descriptive text or prompt. They have recently stunned the architectural visual culture (see Steinfeld<sup>36</sup> for a recent

thought-provoking discussion and recount of events on the rise of text-to-image generative models in architecture). The use of DMs as a data augmentation technique is becoming a predominant method as they seem to outperform GANs due to superior image quality and diversity.<sup>37,38</sup> Under the generic umbrella term of ‘generative AI’, both GANs and DMs have garnered increasing interest in architecture and urban studies due to the possibility of generating a virtually infinite number of new designs/images.<sup>39–42</sup>

Once a suitable and meaningful dataset has been curated for classification, a CNN architecture is needed to conduct the training. The architecture can be set up from scratch or based on an architecture and corresponding weights from an already trained CNN in a related domain, the latter being commonly referred to as transfer learning.<sup>43</sup> Transfer learning techniques have been extensively applied in a wide range of applications, for example, text sentiment classification, image classification, human activity classification, software defect classification, and multi-language text classification (see Weiss et al.<sup>44</sup> for a comprehensive review on transfer learning).

This work has three main objectives: first, to validate the suitability of CNNs as an effective AI technique for classifying urban spaces; second, to assess the effectiveness of publicly available generative AI engines in producing images for data augmentation purposes; and third, to identify, if any, a specific CNN architecture among a set of well-known CNN architectures traditionally used for general image classification that consistently performs higher when classifying urban spaces. Following the urban classes described in Carmona,<sup>4</sup> this work uses three datasets for the classification of urban spaces: a seed dataset composed of 630 photographs downloaded from the Adobe Stock repository<sup>45</sup> and two augmented datasets, which are each composed of the seed dataset topped up with generative AI images to make a total of 2200 instances. The images used in the enlarged datasets were generated using the generative AI engines Deep Dream Generator<sup>46</sup> and Midjourney-v4<sup>47</sup> (these enlarged datasets are called ‘DDG-augmented dataset’ and ‘MJ4-augmented dataset’, respectively, hereafter). These generative AI engines were selected because they are two of the most popular and highly regarded image generators publicly available and accessible at the time of writing.<sup>48</sup> The generated images, although possessing high definition and detail, are sometimes disjointed and do not necessarily coherently represent the urban space specified in the text prompt when judged by a human. Nevertheless, they preserve many of the features of the particular class that make the urban space recognizable when observed by a human; therefore, they were considered suitable for the purposes of this work. Due to the varying number of photographs in the seed dataset, it is highly unbalanced; on the other hand, the augmented datasets are perfectly balanced, as every urban category or class has the same number of elements. For the training of these datasets, a transfer learning technique was used where the architectures of well-known CNNs used for image classification were transferred to the domain discussed in this article. Four CNNs widely used in data science for image classification problems were used for training in this work, namely, (i) GoogLeNet,<sup>49</sup> (ii) ResNet-18,<sup>50</sup> (iii) ShuffleNet,<sup>51</sup> and (iv) MobileNet-v2.<sup>52</sup> It was found that prediction capacity increases roughly 30% when enlarging the dataset in all CNNs, even though the GAN-enlarged dataset possesses images of urban spaces that are not always coherent. It was also observed that ResNet-18 outperforms the other three CNNs used in this work in all prediction metrics, suggesting that this architecture could be more capable of handling much larger datasets for urban classification. Finally, it was possible to observe that Midjourney-v4 slightly outperforms Deep Dream Generator when used as a data augmentation technique in the urban spaces domain, although both are decidedly capable of interpreting the main features of each class of urban space studied in this work.

This article is organized as follows: the next section describes the generation of the seed and augmented datasets for each category or class of urban space; the subsequent part covers the training process for the four CNN architectures used on the datasets described in the previous section; then, the following section analyses the results obtained and presents comparisons among the CNNs and datasets; the following section presents a thorough discussion of the results, finishing with a section covering the conclusions of this work.

## Dataset generation

This section covers how the datasets were generated for this research. It is worth mentioning that the seed dataset has been made publicly available for reproducibility purposes and can be found on Kaggle,<sup>53</sup> a widely used platform for data science resources. The augmented datasets are available upon request to the corresponding author.

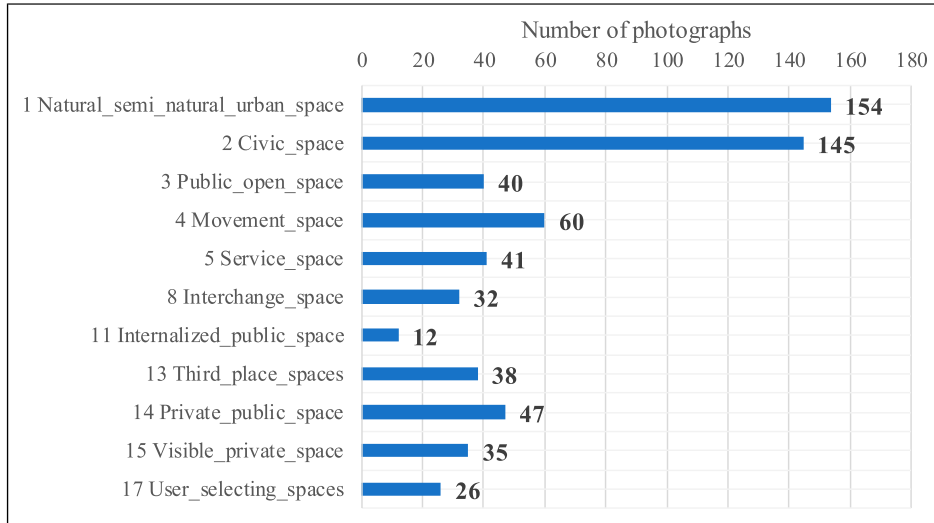
### Seed dataset

Initially, a seed dataset was generated as the main source of information using photographs from the Adobe Stock repository.<sup>45</sup> Only those belonging to the 'Free' category were used, which avoided the need to pay royalty fees. The search of photographs was conducted using keywords and descriptions as provided in the classes definition stated in Carmona<sup>4</sup> and reproduced in Table 1. The corresponding labels for each class were simply defined as the space type (e.g. 'natural/semi-natural urban space', 'civic space', 'public open space'). Regarding the image labelling process, it is acknowledged that this task can, in general, be conducted in three different ways<sup>54</sup>: (i) whole image labelling, (ii) bounding box (rectangular subregion) labelling, and (iii) polygon (complex region) labelling. In this work, a whole image labelling process was used, as it is of interest to assess the capacity of a machine learning model to interpret images in their entirety and classify them as a certain type of urban space. It is worth mentioning that no bias was introduced by the authors of this work in the process of gathering photographs for each class during the search on the Adobe Stock repository, as the photographs' authors themselves labelled their photographs as part of uploading the files to the repository. It is also relevant to state that photographs taken at night and those with a predominant presence of people were ruled out. Although the search was conducted for the 20 classes of urban spaces defined by Carmona,<sup>4</sup> some classes were left void as no meaningful photographs were found. As a consequence, a dataset of 11 classes was generated composed of 630 photographs; the 11 classes are highlighted in bold in Table 1. The distribution of photographs within each class is highly uneven (as shown in Figure 1) and, therefore, unbalanced from a data science perspective. For consistency purposes, the classes are assigned in the same order as those reported in Carmona<sup>4</sup>; therefore, labels for each class in the datasets will not appear in sequence. A sample of photographs from the seed dataset is shown in Figure 2.

From Figure 1, it is possible to see the unbalanced nature of the seed dataset; class imbalance is a common problem in real-world datasets and can have a detrimental effect in the performance of CNNs in classification problems.<sup>55</sup> Generally speaking, there are two types of class imbalance: (i) step or inter-class imbalance, where there is a minority of classes with a similar number of samples, say S1, in comparison to a majority of classes that also has a similar number of samples, say S2, but S1 and S2 differ substantially; and (ii) linear or intra-class imbalance, where the number of samples among all classes varies in such a way that the difference between consecutive classes is nearly constant (see Buda et al.<sup>56</sup> and Ali et al.<sup>57</sup> for comprehensive reviews on class imbalance problems in classification problems using CNNs). In this light, the seed dataset used in this work possesses a clear step or inter-class imbalance, as there is a minority of classes with a similar number of instances (classes 1 and 2, which have ~140 samples on average) compared to a majority of classes with a relatively similar number of instances (classes 3 to 17, which have ~36 samples on average). Certainly, the number of samples between the majority and the minority is significantly different. It has been reported that CNNs trained using datasets with step or inter-class imbalance, as is the case for the seed dataset, are more likely to perform poorly when classifying those classes with a low number of samples because they are predicted as rare occurrences.<sup>55</sup> The imbalance problem with the seed dataset was then corrected by using augmented datasets, which, from a data science perspective, can be regarded as ideal datasets, as they contain an equal number of samples within each category.

**Table 1.** Urban space types<sup>4</sup> (categories in bold were used in this work).

| Space type                                | Distinguishing characteristics   | Examples  |
|---|--|---|
| 'Positive' spaces                         |  |   |
| <b>1 Natural/semi-natural urban space</b> | Natural and semi-natural features within urban areas, typically under state ownership                        | Rivers, natural features, seafronts, canals                                   |
| <b>2 Civic space</b>                      | The traditional forms of urban space, open and available to all and catering for a wide variety of functions | Streets, squares, promenades  |
| <b>3 Public open space</b>                | Managed open space, typically green and available and open to all, even if temporarily controlled            | Parks, gardens, commons, urban forests, cemeteries                            |
| 'Negative' spaces                         |  |   |
| <b>4 Movement space</b>                   | Space dominated by movement needs, largely motorized transportation  | Main roads, motorways, railways, underpasses                                  |
| <b>5 Service space</b>                    | Space dominated by modern servicing, requirements needs  | Car parks, service yards  |
| 6 Left over space                         | Space left over after development, often designed without a function   | 'SLOAP' (space left over after planning), modernist open space                |
| 7 Undefined space                         | Undeveloped space, either abandoned or awaiting redevelopment  | Redevelopment space, abandoned space, transient space                         |
| Ambiguous spaces                          |  |   |
| <b>8 Interchange space</b>                | Transport stops and interchanges, whether internal or external   | Metros, bus interchanges, railway stations, bus/tram stops                    |
| 9 Public 'private' space                  | Seemingly public external space, in fact privately owned and to greater or lesser degrees controlled         | Privately owned 'civic' space, business parks, church grounds                 |
| 10 Conspicuous spaces                     | Public spaces designed to make strangers feel conspicuous and, potentially, unwelcome                        | Cul-de-sacs, dummy gated enclaves   |
| <b>11 Internalized 'public' space</b>     | Formally public and external uses, internalized and, often, privatized                                       | Shopping/leisure malls, introspective megastructures                          |
| 12 Retail space                           | Privately owned but publicly accessible exchange spaces  | Shops, covered markets, petrol stations                                       |
| <b>13 Third place spaces</b>              | Semi-public meeting and social places, public and private  | Cafes, restaurants, libraries, town halls, religious buildings                |
| <b>14 Private 'public' space</b>          | Publicly owned, but functionally and user determined spaces  | Institutional grounds, housing estates, university campuses                   |
| <b>15 Visible private space</b>           | Physically private, but visually public space  | Front gardens, allotments, gated squares                                      |
| 16 Interface spaces                       | Physically demarked but publicly accessible interfaces between public and private space                      | Street cafes, private pavement space  |
| <b>17 User selecting spaces</b>           | Spaces for selected groups, determined (and sometimes controlled) by age or activity                         | Skateparks, playgrounds, sports fields/grounds/courses                        |
| Private spaces                            |  |   |
| 18 Private open spaces                    | Physically private open space  | Urban agricultural remnants, private woodlands                                |
| 19 External private space                 | Physically private spaces, grounds and gardens   | Gated streets/enclaves, private gardens, private sports clubs, parking courts |
| 20 Internal private space                 | Private or business space  | Offices, houses, etc.   |



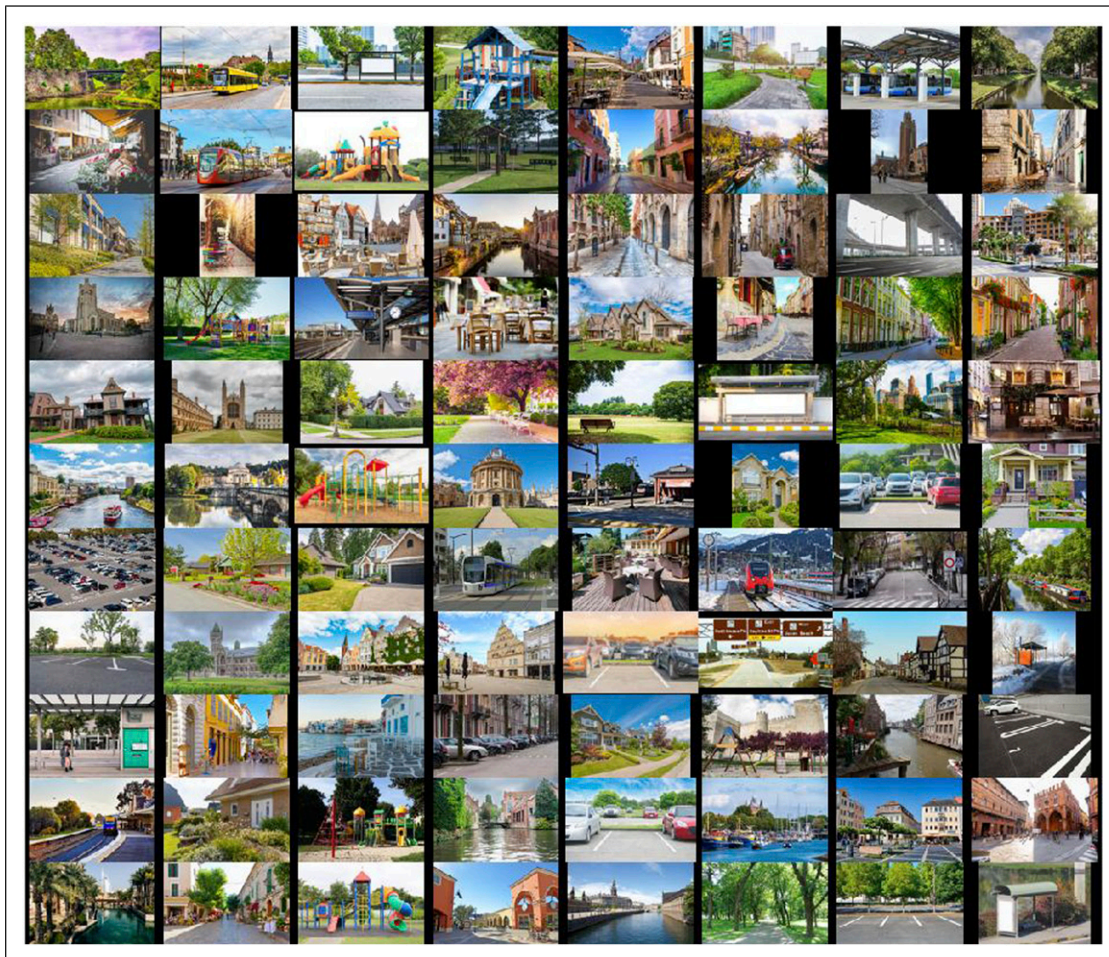
**Figure 1.** Distribution of photographs per class in the seed dataset.

### Augmented datasets

The aim of the augmented datasets was to balance the seed dataset by topping up all classes with images produced by popular text-to-image DMs until every one of the 11 classes is composed of 200 images; hence, each augmented dataset is composed of 2200 images. Images were generated using the Text 2 Dream tool from the Deep Dream Generator engine<sup>46</sup> and the command/imagine in Midjourney-v4.<sup>47</sup> Only text prompts (no base images) were inputted in the generation of the images, which, in turn were the keywords and description used for each urban class, as shown in Table 1 (e.g. ‘a car park’ was used as a prompt to generate images from the category ‘service space’). All prompts included additional descriptors to make the images more in line with those from the seed dataset, such as ‘in broad daylight’ and ‘as seen at street level’. Figures 3–5 show examples of generated images for every urban class in both enlarged datasets studied in this work. It is worth highlighting the unpredictable nature of these images, as they may sometimes be incoherent by having constituting elements that are fragmented or disorganized. Nevertheless, they do overall possess similar feature maps for the urban space for which they were created in comparison with the photographs in the seed dataset; hence, they were deemed suitable for the classification problem studied in this work. This assumption has been previously validated in the literature: complete realism is not needed when using artificial instances for data augmentation.<sup>58</sup> Figure 6 shows a small sample of instances for interchange spaces with a degree of incoherence or misplaced components that can be detected by human curation. Such misplacement is mainly related to the position of trains, which are depicted in positions other than on tracks.

### Convolutional neural network training processes

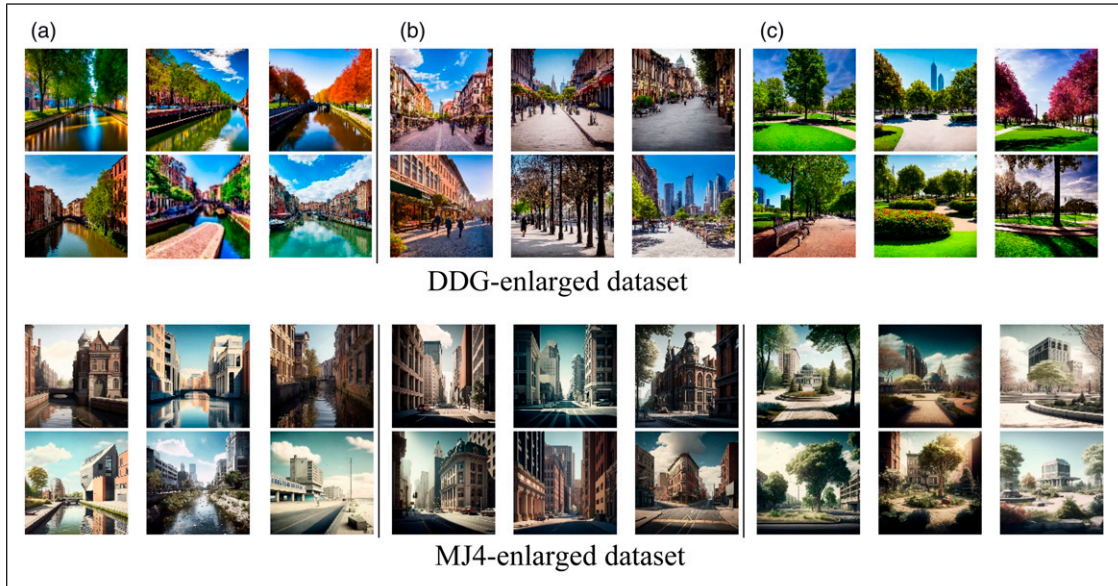
Training was carried out in MATLAB’s Deep Learning toolbox<sup>59</sup> using the three datasets described in Section 2 by making use of the architecture and weights of four CNNs widely used in data science for image classification problems, namely, (i) GoogLeNet,<sup>49</sup> (ii) ResNet-18,<sup>50</sup> (iii) ShuffleNet,<sup>51</sup> and (iv) MobileNet-v2.<sup>52</sup> They are all classed as deep neural networks with a varying number of layers, specifically 22, 18, 50, and 53 layers for GoogLeNet, ResNet-18, ShuffleNet, and MobileNet-v2, respectively. All of them have been previously used in



**Figure 2.** Sample photographs in the seed dataset.

computer vision research in a wide range of fields and applications, such as aerial image classification,<sup>60</sup> face recognition,<sup>61</sup> bruised nectarine detection,<sup>62</sup> variants of cancer identification,<sup>63</sup> handwritten character recognition,<sup>64</sup> garbage classification,<sup>65</sup> dairy cow identification,<sup>66</sup> crop diseases identification,<sup>67</sup> palmprint recognition,<sup>68</sup> and classification of retinal diseases,<sup>69</sup> among many others. It is worth adding that the four CNNs selected in this work have previously successfully performed in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), which allows researchers to test algorithms for object detection and image classification at large scale.<sup>70</sup> ILSVRC uses a dataset of more than one million images, which are grouped into 1000 categories, including animals, vehicles, fruits/vegetables, landscape elements, home appliances, and clothing items. The generic architectures of these CNNs are shown schematically in Figure 7.

Regarding the training parameters, for all datasets, the learning rate was fixed to 0.01 and the size of the subsets for training/test was defined as 80% and 20%, respectively. These were selected for the sake of simplicity, as they are within typical ranges used in classification projects in the architectural domain.<sup>18,20</sup> Additionally, the number of iterations per epoch was set to 3 (maximum of 20 epochs) for the seed dataset and 13 (maximum of 30 epochs) for the augmented datasets. These were selected after observing convergence

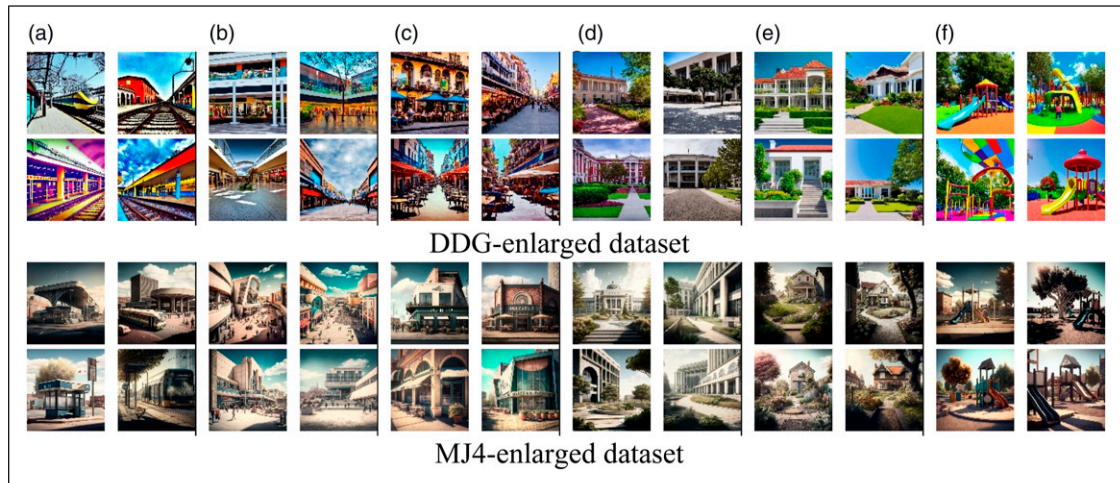


**Figure 3.** Sample of generated images for 'positive' urban spaces. (a) Natural/semi-natural urban spaces; (b) civic spaces; (c) public open spaces.

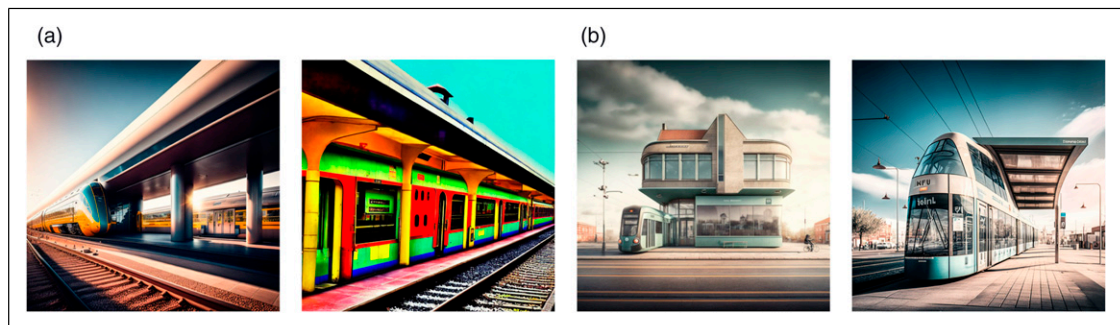


**Figure 4.** Sample of generated images for 'negative' urban spaces. (a) Movement spaces; (b) service spaces.





**Figure 5.** Sample of generated images for ‘ambiguous’ urban spaces. (a) Interchange space; (b) internalized public space; (c) third place spaces; (d) private public space; (e) visible private space; (f) user selecting spaces.

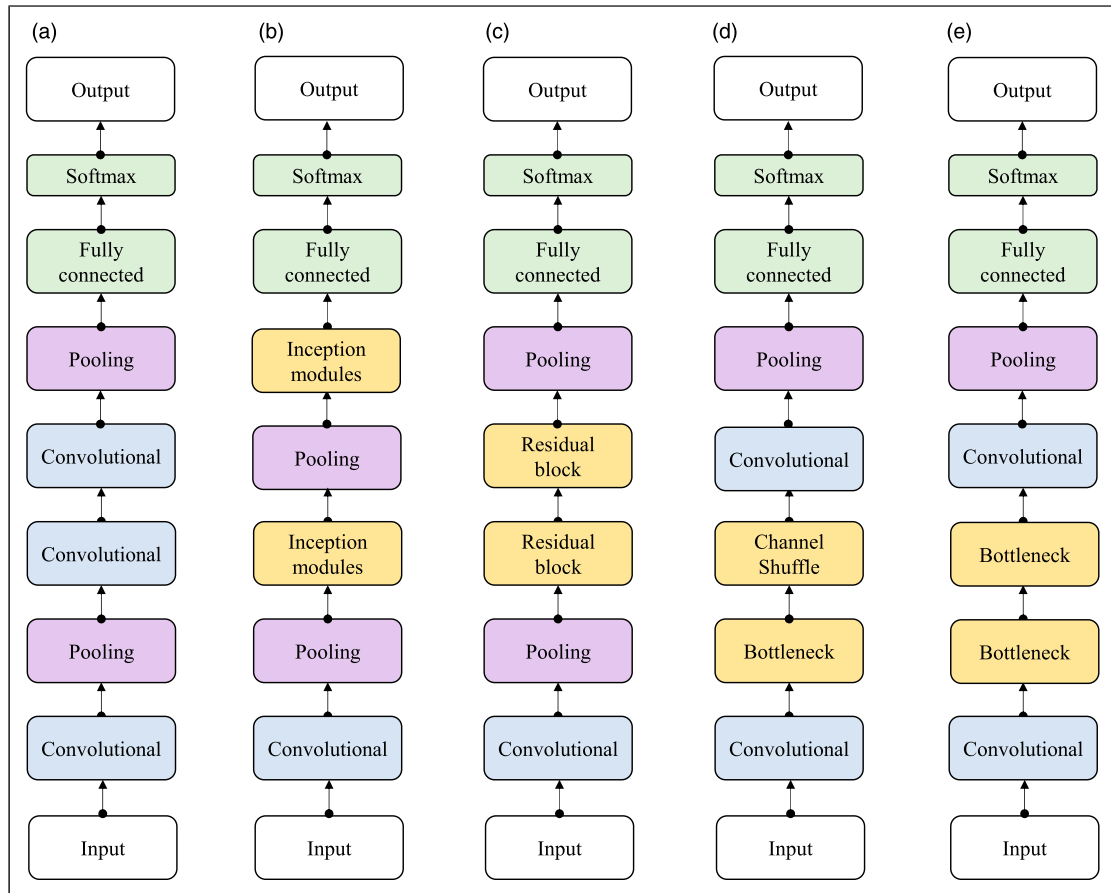


**Figure 6.** Sample instances with a degree of incoherence or misplacement in interchange spaces. (a) From DDG-enlarged dataset; (b) from MJ4-enlarged dataset.

during previous trial runs with all datasets. In this line, [Figure 8](#) summarizes the training accuracy (%) versus the number of iterations for all the training processes carried out in this work using the three datasets and four CNNs. From this figure, it is possible to see that all four CNNs are able to reach high accuracy at the end of the training process; however, the rate at which such accuracy is reached varies significantly for the three datasets, as indicated by the slope of each curve. For the augmented datasets, ResNet-18 and MobileNet-v2 seem to be the architectures that reach higher levels of accuracy during training, which is a significant characteristic, especially when it comes to datasets that are substantially larger than the ones used in this work (as is likely to be the case in a real-world application in urban studies).

## Analysis of the results

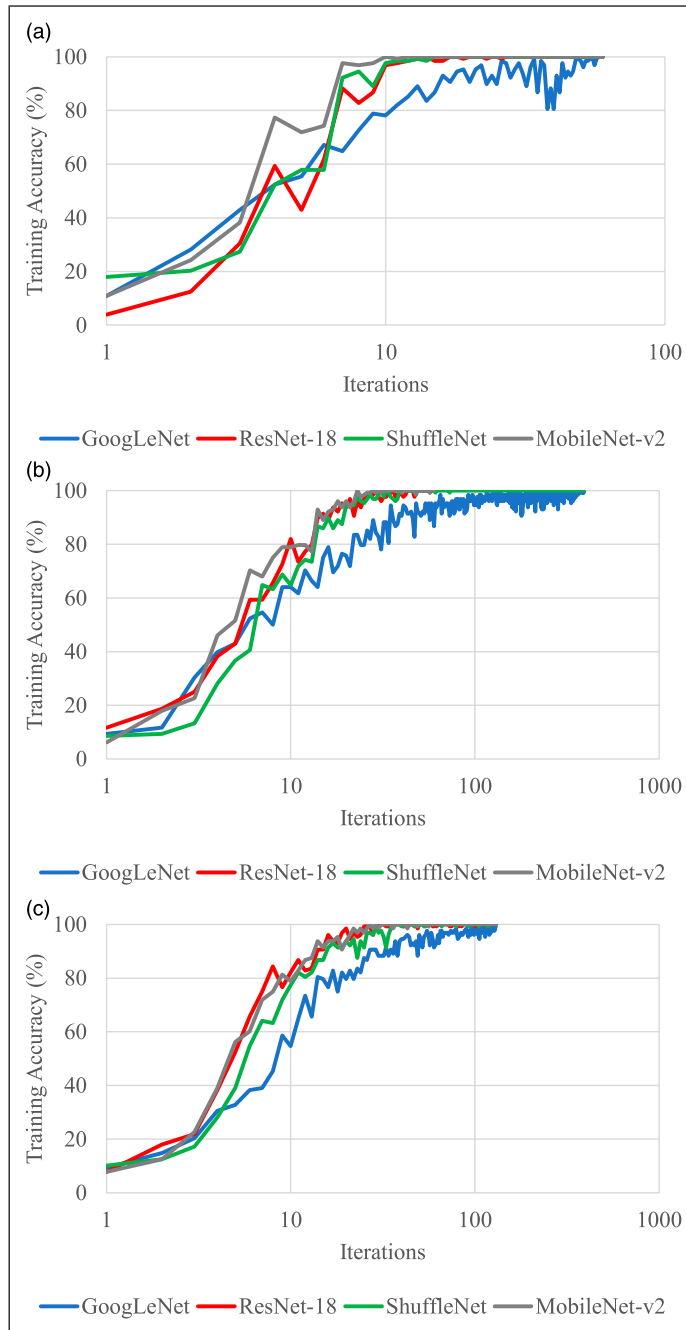
Once training was carried out, the predictive capacity of the CNNs was measured against the test datasets, which were generated by randomly choosing 20% of the images from the original datasets (i.e.



**Figure 7.** Generic architectures of the CNNs used in this work compared to a classic CNN architecture shown in (a); (b) GoogLeNet; (c) ResNet-18; (d) ShuffleNet; (e) MobileNet-v2.

126 photographs for the seed dataset and 440 images for each augmented dataset). Figures 9–11 show the confusion matrices obtained for the test datasets from the seed dataset, DDG-augmented dataset, and MJ4-augmented dataset, respectively. It is worth mentioning that the confusion matrix for the test seed dataset (Figure 9) shows a higher prediction capacity for those classes better populated (e.g. classes 1 and 2), which is an expected outcome considering the class distribution in the seed dataset, as shown in Figure 1. Using the results shown in Figures 9–11, performance evaluation was conducted using four well-known metrics for multiclass classification, namely, accuracy, precision, recall, and F1-score.

Accuracy was determined using equation (1), which measures how much the CNNs are correctly making predictions on the entire dataset. Precision was determined for each class individually using equation (2) to then take the unweighted mean (sometimes referred to as macro-precision) to measure the ability of the CNNs to predict an individual positive. Similarly, recall was determined for each class individually using equation (3) to then take the unweighted mean (sometimes referred to as macro-recall) to measure the ability of the CNNs to find all positive individuals in the datasets. Finally, the F1-score was determined using equation (4) (referred to as macro-F1-score, as it is determined using individual precision and recall for each class) by combining precision and recall into a single measure, which mathematically represents the harmonic mean of



**Figure 8.** Training accuracy for all training processes carried out in this work for the three datasets. (a) Seed dataset; (b) DDG-augmented dataset; (c) MJ4-augmented dataset.

precision and recall. Table 2 summarizes these metrics for the four CNNs and the three datasets used in this work, where the maximum value is highlighted in bold. From this table, it is possible to observe that the highest performance is obtained when using the ResNet-18 architecture. Additionally, Figure 12 shows a graphical summary of these metrics, where it is possible to see that all metrics are substantially higher for the augmented datasets compared to the seed dataset. From Figure 12, it is also possible to observe that the MJ4-augmented dataset slightly outperforms the DDG-augmented dataset for three out of the four CNN architectures considered in this work

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \quad (1)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

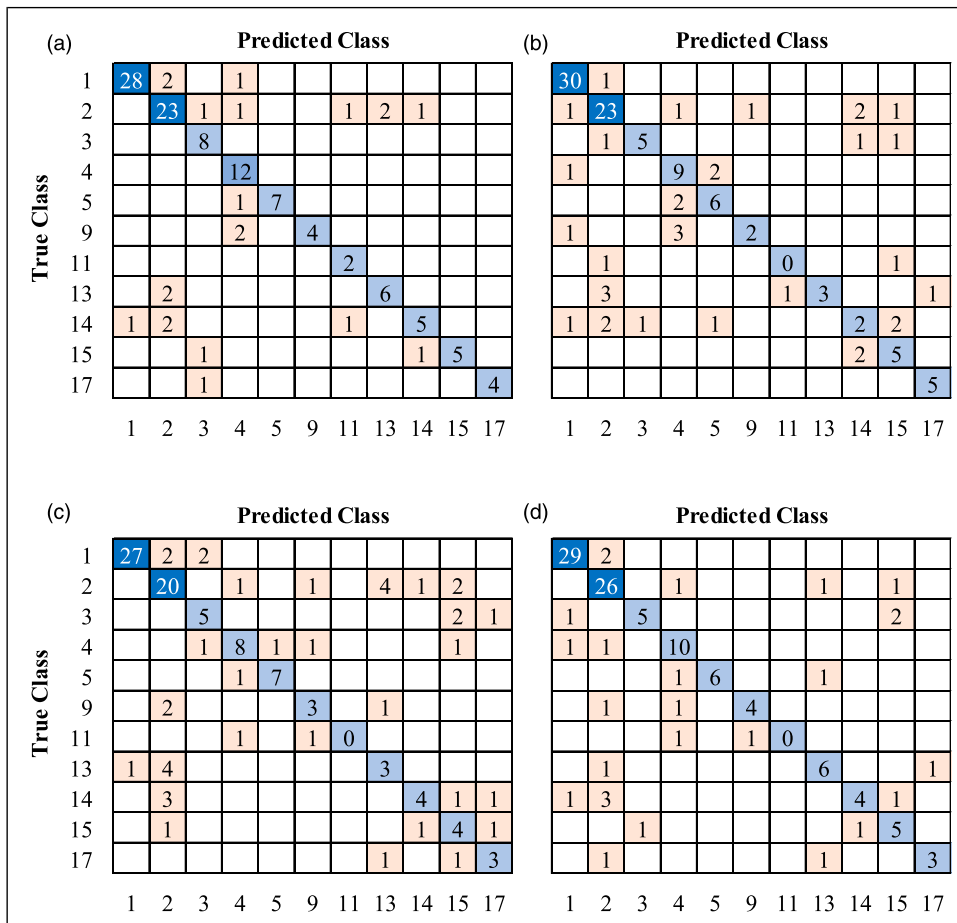
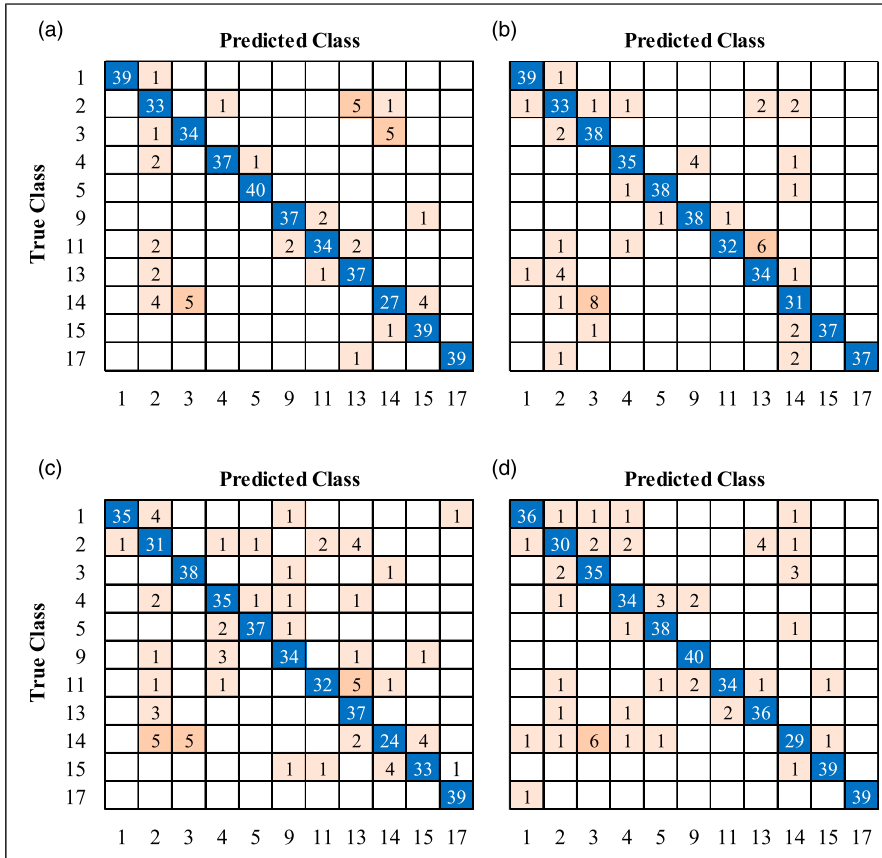


Figure 9. Confusion matrices for the test seed dataset; (a) GoogLeNet; (b) ResNet-18; (c) ShuffleNet; (d) MobileNet-v2.



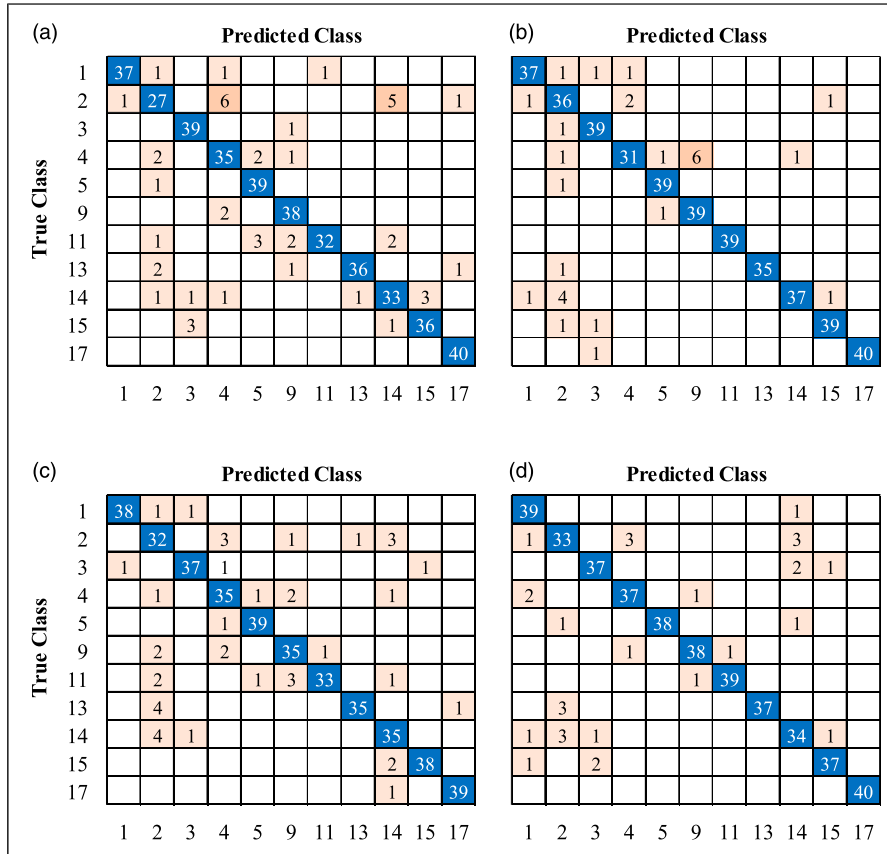
**Figure 10.** Confusion matrices for the test DDG-augmented dataset; (a) GoogLeNet; (b) ResNet-18; (c) ShuffleNet; (d) MobileNet-v2.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (3)$$

$$F1\_score = 2 \left( \frac{Precision \cdot Recall}{Precision + Recall} \right) \quad (4)$$

## Discussion

- *Seed (imbalanced) vs augmented (balanced) datasets:* when assessing the performance of the four CNN architectures and three datasets, it is possible to observe from the confusion matrices in Figure 9 that all CNNs naturally struggle to correctly classify instances that belong to classes with few instances. This is a known result when training CNNs architectures using imbalanced datasets; it has been reported in the literature<sup>55</sup> and was confirmed in this work. On the other hand, when the same CNN architectures are trained using a perfectly balanced dataset, their performance substantially increases, where all classes



**Figure 11.** Confusion matrices for the test MJ4-augmented dataset; (a) GoogLeNet; (b) ResNet-18; (c) ShuffleNet; (d) MobileNet-v2.

have equal probability of being correctly classified, as seen from the confusion matrices shown in Figures 10 and 11.

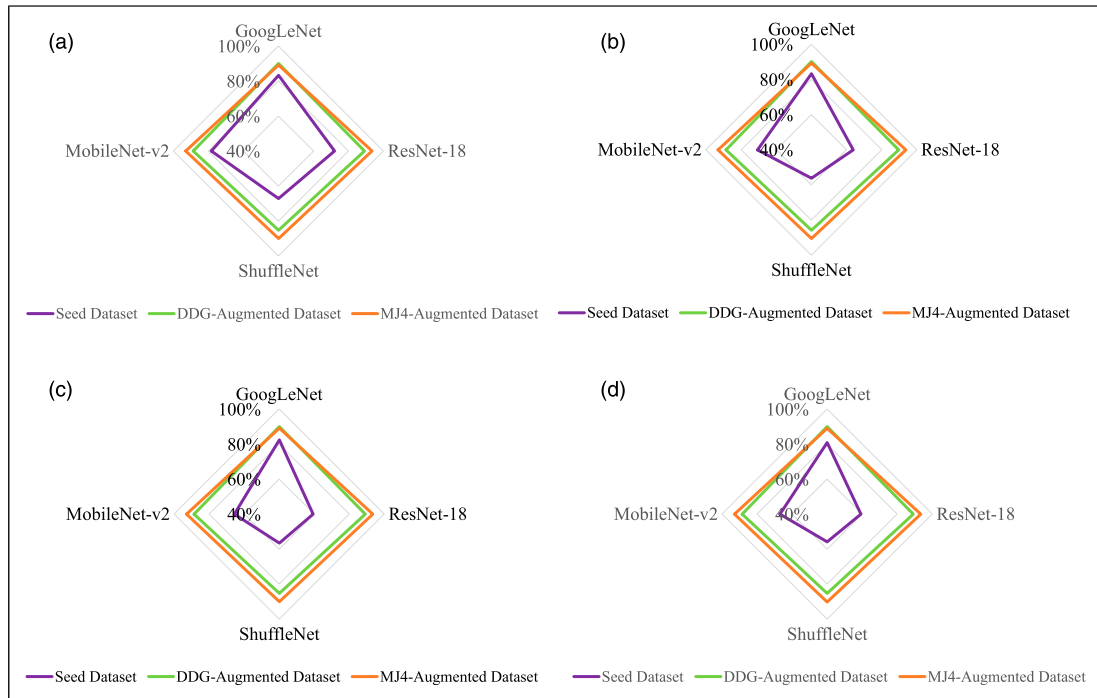
- *Labelling process:* the results reported in this work are valid when a whole image labelling process is used for urban space classification. However, it is acknowledged that the results could be different when approaching the problem as a multi-label classification task, that is, where each instance can be associated with several labels simultaneously.<sup>71</sup> Such an approach would require a different labelling process to the one used in this work, such as bounding box labelling. This would allow identification of individual components that belong to more than one urban space, for example, streets or trees, where the model could then identify an urban space by integrating its different components. In any case, the approach used in this work can be considered as a first approximation to the problem, and the use of other labelling processes is left as a matter of further research.
- *Training parameters:* the results in this work were obtained for a specific set of numerical values for the training parameters, namely, learning rate, size of subsets for training/test, number of iterations per epoch, and maximum number of epochs. Although the numerical values selected for these parameters were within traditional values typically used for this type of problem, it is acknowledged that they remained unmodified throughout the training process. Hence, the results could be improved if these

**Table 2.** Metrics for multiclass classification for the four CNNs and three test datasets.

|                  | Seed Dataset (%) | DDG-Augmented Dataset (%) | MJ4- Augmented Dataset (%) |
|------------------|------------------|---------------------------|----------------------------|
| <b>Accuracy</b>  |                  |                           |                            |
| GoogLeNet        | 83.2             | 90.0                      | 89.1                       |
| ResNet-18        | 72.0             | 89.1                      | <b>93.4</b>                |
| ShuffleNet       | 67.2             | 85.2                      | 90.0                       |
| MobileNet-v2     | 78.4             | 88.6                      | 93.0                       |
| <b>Precision</b> |                  |                           |                            |
| GoogLeNet        | 83.2             | 90.2                      | 89.3                       |
| ResNet-18        | 63.7             | 89.7                      | <b>93.8</b>                |
| ShuffleNet       | 56.2             | 85.8                      | 90.6                       |
| MobileNet-v2     | 70.6             | 88.7                      | 93.1                       |
| <b>Recall</b>    |                  |                           |                            |
| GoogLeNet        | 82.3             | 90.0                      | 89.1                       |
| ResNet-18        | 59.4             | 89.1                      | <b>93.4</b>                |
| ShuffleNet       | 56.5             | 85.2                      | 90.0                       |
| MobileNet-v2     | 65.6             | 88.6                      | 93.0                       |
| <b>F1-score</b>  |                  |                           |                            |
| GoogLeNet        | 80.9             | 90.0                      | 89.0                       |
| ResNet-18        | 59.2             | 89.2                      | <b>93.4</b>                |
| ShuffleNet       | 55.8             | 85.2                      | 90.2                       |
| MobileNet-v2     | 67.0             | 88.5                      | 93.0                       |

parameters were fine-tuned and/or enabled to dynamically change in an adaptive manner. Such an approach has been reported in the literature<sup>18,20</sup> and shown to be an effective and efficient way of training. However, further improvement to the already high accuracy achieved with static training parameters is left as a matter of future research on this topic.

- *Generative AI engines:* As seen from [Figure 12](#), both generative engines used in this work seem to have a fairly similar performance when used as an augmentation technique for classifying urban spaces. However, when curated by a human, the generated images produced by Midjourney seem to have a superior coherence, that is, it is less likely for components to be misplaced in comparison to those produced by Deep Dream Generator. This is in line with a research result already reported in the literature:<sup>58</sup> complete realism is not absolutely necessary when using generative images as a data augmentation technique; this is because the actual features that appear on the produced images, not their placement therein, will likely dictate the quality of the image for classification purposes. Certainly, this statement might not be valid for applications where realism and accuracy are required. Considering that Midjourney and Deep Dream Generator were the two generative AI engines used, the topic discussed in this paragraph is not necessarily valid when other publicly available generative AI engines are used, such as DALL-E, Stable Diffusion AI, or Runway.
- *Limitations:* This work has been conducted using a specific set of CNN architectures that were trained using three particular datasets. Therefore, the results reported in this article are limited to such a domain and are not necessarily valid outside of it. Other well-known CNN architectures used in classification problems, such as VGGNet or AlexNet, may also be well suited to perform in classification problems of urban spaces and could potentially compete against the highest-performing architecture in this work, namely, ResNet-18. In addition, and aiming at generalization, it is yet to be seen how these CNN



**Figure 12.** Metrics for multiclass classification for the four CNNs and three test datasets. (a) Accuracy; (b) precision; (c) recall; (d) F1-score.

architectures would perform when including the other urban spaces defined by Carmona<sup>4</sup> that were not included in this work. Finally, the results reported in this article are constrained by the specific urban space classification used, that is, the design-functional perspective proposed by Carmona.<sup>4</sup> In summary, how the results reported in this work could be impacted by the use of other CNN architectures, the inclusion of urban spaces left void in this work, and even a totally different system to classify urban spaces is left as a matter of exploration in future research.

## Conclusions

Classification of urban spaces is a pertinent topic for policymakers and city planners in order to make informed investment decisions to foster interaction in cities. This is an ever-growing topic due to the increasing complexity of contemporary cities, which makes the task of classification also progressively more difficult. This makes machine learning models an ideal candidate for this purpose due to their capacity to learn patterns based on analyzing large amounts of data. Among the several attempts made in the literature to establish categories or classes of urban spaces, one such approach is based on a design-functional perspective. This standpoint focuses on the intended use, character, and shape of the public space. Following a design-functional perspective widely used in the literature, this work used three datasets to classify urban spaces using CNNs. A seed dataset composed of 630 photographs, which was highly imbalanced, was initially used in the training process and was later topped up with GAN-generated images to make two augmented datasets composed of 2200 images each, which were perfectly balanced. Four CNNs widely known for their data science and AI applications were used in the training processes, namely, GoogLeNet, ResNet-18, ShuffleNet,



and MobileNetv2, which are 22, 18, 50, and 53 layers deep, respectively. These CNN architectures demonstrated their suitability as effective AI techniques for classifying urban spaces, as their metrics for multiclass classification averaged 75% in accuracy, 68% in precision, and 66% for both recall and F1-score. These results validate the first research objective of this work. When using the generative AI engines for data augmentation purposes, results showed that prediction performance for the test datasets (images not used during training) increases substantially when CNNs were trained with the enlarged datasets in comparison with those trained with the seed dataset. Although both generative AI engines demonstrated their suitability for this purpose, Midjourney seems to slightly outperform Deep Dream Generator: when using the former, metrics for multiclass classification improved by 21% in accuracy, 35% in precision, and 38% for both recall and F1-score, averaging roughly a 30% improvement. These results validate the second research objective of this work. Additionally, ResNet-18 seems to perform better in comparison with the other three CNN architectures studied in this work; therefore, it might be more capable of handling larger training datasets, such as those associated with big data projects in urban studies. This result validates the third research objective of this work. This article's expected contributions to the public space design domain include facilitating the analysis of public space utilization through automation and improving understanding of public spaces' roles in fostering community engagement for governance and policymaking. Some real-life examples of projects for which the results reported in this article could be useful are (i) urban revitalization projects for the redevelopment of neglected public spaces to enhance aesthetic appeal, accessibility, and functionality; (ii) placemaking initiatives to transform public spaces into inclusive and people-centric areas; and (iii) temporary urban interventions to test and improve public spaces, such as pop-up parks or street closures. Further improvement in prediction performance by using other labelling processes, training parameters, generative AI engines, and CNN architectures is left as a matter of further research in the field of urban classification projects.

## Acknowledgments

The authors thank two anonymous reviewers for their valuable comments that substantially improved the quality of the manuscript.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iDs

Carlos Medel-Vera  <https://orcid.org/0000-0003-0343-4202>

Thomas Mädler  <https://orcid.org/0000-0001-5076-3362>

## References

1. Martinelli L, Battisti A and Matzarakis A. Multicriteria analysis model for urban open space renovation: an application for Rome. *Sustain Cities Soc* 2015; 14: e10–e20, DOI: [10.1016/j.scs.2014.07.002](https://doi.org/10.1016/j.scs.2014.07.002).
2. Villanueva K, Badland H, Hooper P, et al. Developing indicators of public open space to promote health and wellbeing in communities. *Appl Geogr* 2015; 57: 112–119, DOI: [10.1016/j.apgeog.2014.12.003](https://doi.org/10.1016/j.apgeog.2014.12.003).

3. Alzahrani A. Classification of urban spaces: an attempt to classify Al-baha city urban spaces using carmona's classification. *Sage Open* 2022; 12(2): 21582440221097892. DOI: [10.1177/21582440221097892](https://doi.org/10.1177/21582440221097892).
4. Carmona M. Contemporary public space, Part Two: classification. *J Urban Des* 2010; 15(2): 157–173. DOI: [10.1080/13574801003638111](https://doi.org/10.1080/13574801003638111).
5. Nochian A, Tahir OM, Maulan S, et al. A comprehensive public open space categorization using classification system for sustainable development of public open spaces. *Alam Cipta* 2015; 8(1): 20–40.
6. Stanley BW, Stark BL, Johnston KL, et al. Urban open spaces in historical perspective: a transdisciplinary typology and analysis. *Urban Geogr* 2012; 33(8): 1089–1117. DOI: [10.2747/0272-3638.33.8.1089](https://doi.org/10.2747/0272-3638.33.8.1089).
7. Sandalack BA and Uribe FGA. Open space typology as a framework for design of the public realm. *The Faces of Urbanized Space* 2010; 5: 35–75.
8. Gehl J and Gemozoe L. *New city spaces*. Copenhagen: The Danish Architectural Press, 2001.
9. Dines N and Cattell V. *Public spaces, social relations and well-being in east london*. Bristol: The Policy Press, 2006.
10. Carmona M. Contemporary public space: critique and classification, Part One: critique. *J Urban Des* 2010; 15(1): 123–148. DOI: [10.1080/13574800903435651](https://doi.org/10.1080/13574800903435651).
11. Carmona M and Wunderlich F. *Capital spaces: the multiple complex public spaces of a global city*. London, UK: Routledge, 2013.
12. Hanna S. *Urban complexity. Machine learning and the City 2022*. Oxford, UK: John Wiley & Sons, pp. 1–13.
13. Elgendy M. *Deep learning for vision systems*. New York, USA: Manning Publications Co, 2020.
14. Coulibaly S, Kamsu-Foguem B, Kamissoko D, et al. Deep convolution neural network sharing for the multi-label images classification. *Machine Learning with Applications* 2022; 10: 100422, DOI: [10.1016/j.mlwa.2022.100422](https://doi.org/10.1016/j.mlwa.2022.100422).
15. Rahman T and Islam MS. MRI brain tumor detection and classification using parallel deep convolutional neural networks. *Measurement: Sensors* 2023; 26: 100694, DOI: [10.1016/j.measen.2023.100694](https://doi.org/10.1016/j.measen.2023.100694).
16. Castro Pena ML, Carballal A, Rodríguez-Fernández N, et al. Artificial intelligence applied to conceptual design. A review of its use in architecture. *Autom ConStruct* 2021; 124: 103550, DOI: [10.1016/j.autcon.2021.103550](https://doi.org/10.1016/j.autcon.2021.103550).
17. del Campo M, Carlson A and Manninger S. Towards hallucinating machines - designing with computational vision. *Int J Architect Comput* 2021; 19(1): 88–103. DOI: [10.1177/1478077120963366](https://doi.org/10.1177/1478077120963366).
18. Demir G, Çekmiş A, Yeşilkaynak VB, et al. Detecting visual design principles in art and architecture through deep convolutional neural networks. *Autom ConStruct* 2021; 130: 103826, DOI: [10.1016/j.autcon.2021.103826](https://doi.org/10.1016/j.autcon.2021.103826).
19. Yi YK, Zhang Y and Myung J. House style recognition using deep convolutional neural network. *Autom ConStruct* 2020; 118: 103307, DOI: [10.1016/j.autcon.2020.103307](https://doi.org/10.1016/j.autcon.2020.103307).
20. Zhang G, Pan Y and Zhang L. Deep learning for detecting building façade elements from images considering prior knowledge. *Autom ConStruct* 2022; 133: 104016, DOI: [10.1016/j.autcon.2021.104016](https://doi.org/10.1016/j.autcon.2021.104016).
21. Carranza-García M, García-Gutiérrez J and Riquelme JC. A framework for evaluating land use and land cover classification using convolutional neural networks. *Remote Sensing [Internet]* 2019; 11(3).
22. Liu S, Qi Z, Li X, et al. Integration of convolutional neural networks and object-based post-classification refinement for land use and land cover mapping with optical and SAR data. *Remote Sensing [Internet]* 2019; 11(6).
23. Xu Y, Zhou B, Jin S, et al. A framework for urban land use classification by integrating the spatial context of points of interest and graph convolutional neural network method. *Comput Environ Urban Syst* 2022; 95: 101807, DOI: [10.1016/j.compenvurbsys.2022.101807](https://doi.org/10.1016/j.compenvurbsys.2022.101807).
24. Chen B, Feng Q, Niu B, et al. Multi-modal fusion of satellite and street-view images for urban village classification based on a dual-branch deep neural network. *Int J Appl Earth Obs Geoinf* 2022; 109: 102794, DOI: [10.1016/j.jag.2022.102794](https://doi.org/10.1016/j.jag.2022.102794).
25. Law S, Seresinhe CI, Shen Y, et al. Street-Frontage-Net: urban image classification using deep convolutional neural networks. *Int J Geogr Inf Sci* 2020; 34(4): 681–707. DOI: [10.1080/13658816.2018.1555832](https://doi.org/10.1080/13658816.2018.1555832).
26. Yang M, Mou Y, Liu S, et al. Detecting and mapping tree crowns based on convolutional neural network and Google Earth images. *Int J Appl Earth Obs Geoinf* 2022; 108: 102764, DOI: [10.1016/j.jag.2022.102764](https://doi.org/10.1016/j.jag.2022.102764).

27. Shorten C and Khoshgoftaar TM. A survey on image data augmentation for deep learning. *Journal of Big Data* 2019; 6(1): 60. DOI: [10.1186/s40537-019-0197-0](https://doi.org/10.1186/s40537-019-0197-0).
28. Aggarwal A, Mittal M and Battineni G. Generative adversarial network: an overview of theory and applications. *International Journal of Information Management Data Insights* 2021; 1(1): 100004, DOI: [10.1016/j.ijime.2020.100004](https://doi.org/10.1016/j.ijime.2020.100004).
29. Gui J, Sun Z, Wen Y, et al. A review on generative adversarial networks: algorithms, theory, and applications. *IEEE Trans Knowl Data Eng* 2023; 35(4): 3313–3322. DOI: [10.1109/TKDE.2021.3130191](https://doi.org/10.1109/TKDE.2021.3130191).
30. Yang L, Zhang Z, Song Y, et al. Diffusion models: a comprehensive survey of methods and applications. *ACM Comput Surv* 2023; 56(4): 1–39.
31. Chang Z, Koulteris GA and Shum HPH. *On the design fundamentals of diffusion models: a survey*. arXiv:2306.04542, 2023.
32. Sajeeda A and Hossain BMM. Exploring generative adversarial networks and adversarial training. *International Journal of Cognitive Computing in Engineering* 2022; 3: 78–89, DOI: [10.1016/j.ijcce.2022.03.002](https://doi.org/10.1016/j.ijcce.2022.03.002).
33. Frid-Adar M, Diamant I, Klang E, et al. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 2018; 321: 321–331, DOI: [10.1016/j.neucom.2018.09.013](https://doi.org/10.1016/j.neucom.2018.09.013).
34. Han C, Hayashi H, Rundo L, et al., (eds). GAN-based synthetic brain MR image generation. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Washington, DC, USA. 2018 4-7 April 2018.
35. Bowles C, Chen L, Guerrero R, et al. *GAN augmentation: augmenting training data using generative adversarial networks*. arXiv:1810.10863, 2018.
36. Steinfeld K. Clever little tricks: a socio-technical history of text-to-image generative models. *Int J Architect Comput* 2023; 21(2): 211–241. DOI: [10.1177/14780771231168230](https://doi.org/10.1177/14780771231168230).
37. Burg MF, Wenzel F, Zietlow D, et al. *A data augmentation perspective on diffusion models and retrieval*. arXiv:2304.10253, 2023.
38. Trabucco B, Doherty K, Gurinas M, et al. *Effective data augmentation with diffusion models*. arXiv:2302.07944, 2023.
39. Newton D. Generative deep learning in architectural design. *Technology|Architecture + Design* 2019; 3(2): 176–189. DOI: [10.1080/24751448.2019.1640536](https://doi.org/10.1080/24751448.2019.1640536).
40. Huang J, Johanes M, Kim FC, et al. On GANs, NLP and architecture: combining human and machine intelligences for the generation and evaluation of meaningful designs. *Technology|Architecture + Design* 2021; 5(2): 207–224. DOI: [10.1080/24751448.2021.1967060](https://doi.org/10.1080/24751448.2021.1967060).
41. Ennemoser B and Mayrhofer-Hufnagl I. Design across multi-scale datasets by developing a novel approach to 3DGANs. *Int J Architect Comput*; 21(2): 14780771231168231. DOI: [10.1177/14780771231168231](https://doi.org/10.1177/14780771231168231).
42. del Campo M. *Diffusions in architecture: artificial intelligence and image generators*. Hoboken, NJ: Wiley, 2024.
43. Pan SJ and Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng* 2010; 22(10): 1345–1359. DOI: [10.1109/TKDE.2009.191](https://doi.org/10.1109/TKDE.2009.191).
44. Weiss K, Khoshgoftaar TM and Wang D. A survey of transfer learning. *Journal of Big Data* 2016; 3(1): 9. DOI: [10.1186/s40537-016-0043-6](https://doi.org/10.1186/s40537-016-0043-6).
45. Adobe\_Stock. <https://stock.adobe.com/2022> [Accessed December 2022].
46. Deep\_Dream\_Generator. <https://deepdreamgenerator.com/2023> [Accessed January 2023].
47. Midjourney. <https://www.midjourney.com> [Accessed January 2023].
48. Johnson A. Here are the best AI image generators, 2023, <https://www.forbes.com/sites/ariannajohnson/2023/04/28/here-are-the-best-ai-image-generators/?sh=2d2bb23d3f3c>. *Forbes magazine* (Accessed December 2023).
49. Szegedy C, Wei L, Yangqing J, et al., (eds). *Going deeper with convolutions*. 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. arXiv:1409.4842. 2015 7-12 June 2015.
50. He K, Zhang X, Ren S, et al., (eds). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA. arXiv:1512.03385. 2016 27-30 June 2016.

51. Zhang X, Zhou X, Lin M, et al. *ShuffleNet: an extremely efficient convolutional neural network for mobile devices*. arXiv preprint arXiv:1707.01083vol. 22017.
52. Sandler M, Howard A, Zhu M, et al. *MobileNetV2: inverted residuals and linear bottlenecks*. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. arXiv:1801.04381.
53. Medel-Vera C. <https://www.kaggle.com/datasets/carlosmedelvera/urban-spaces.www.kaggle.com2023>. [Accessed January 2023]
54. Lu S, Wagstaff KL, Alanis R, et al. Chapter 5 - tutorial: how to access, process, and label PDS image data for machine learning☆☆Contribution prepared by the Contributor on behalf of JPL/Caltech. In: J Helbert, M D'Amore, M Aye, et al. (eds). *Machine Learning for Planetary Science*. Amsterdam: Elsevier, 2022, pp. 91–110.
55. Sampath V, Murtua I, Aguilar Martín JJ, et al. A survey on generative adversarial networks for imbalance problems in computer vision tasks. *Journal of Big Data* 2021; 8(1): 27. DOI: [10.1186/s40537-021-00414-0](https://doi.org/10.1186/s40537-021-00414-0).
56. Buda M, Maki A and Mazurowski MA. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Network* 2018; 106: 249–259, DOI: [10.1016/j.neunet.2018.07.011](https://doi.org/10.1016/j.neunet.2018.07.011).
57. Ali A, Shamsuddin SM and Ralescu AL. Classification with class imbalance problem: a review. *Int J Adv Soft Comput Its Appl* 2013; 5(3).
58. Dosovitskiy A, Fischer P, Ilg E, et al. FlowNet: learning optical flow with convolutional networks. *International conference on computer Vision2015*. Piscataway, NJ: IEEE: 2758–2766.
59. Mathworks. MATLAB. <https://la.mathworks.com/products/matlab.html2021>
60. Minu MS and Canessane RA. Deep learning-based aerial image classification model using inception with residual network and multilayer perceptron. *Microprocess Microsyst* 2022; 95: 104652, DOI: [10.1016/j.micpro.2022.104652](https://doi.org/10.1016/j.micpro.2022.104652).
61. Peng S, Huang H, Chen W, et al. More trainable inception-ResNet for face recognition. *Neurocomputing* 2020; 411: 9–19, DOI: [10.1016/j.neucom.2020.05.022](https://doi.org/10.1016/j.neucom.2020.05.022).
62. Yang Y, Wang L, Huang M, et al. Polarization imaging based bruise detection of nectarine by using ResNet-18 and ghost bottleneck. *Postharvest Biol Technol* 2022; 189: 111916, DOI: [10.1016/j.postharvbio.2022.111916](https://doi.org/10.1016/j.postharvbio.2022.111916).
63. Sarwinda D, Paradisa RH, Bustamam A, et al. Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. *Proc Comput Sci* 2021; 179: 423–431, DOI: [10.1016/j.procs.2021.01.025](https://doi.org/10.1016/j.procs.2021.01.025).
64. Abu Al-Haija Q. Leveraging ShuffleNet transfer learning to enhance handwritten character recognition. *Gene Expr Patterns* 2022; 45: 119263, DOI: [10.1016/j.gep.2022.119263](https://doi.org/10.1016/j.gep.2022.119263).
65. Chen Z, Yang J, Chen L, et al. Garbage classification system based on improved ShuffleNet v2. *Resour Conserv Recycl* 2022; 178: 106090, DOI: [10.1016/j.resconrec.2021.106090](https://doi.org/10.1016/j.resconrec.2021.106090).
66. Wang Y, Xu X, Wang Z, et al. ShuffleNet-Triplet: a lightweight RE-identification network for dairy cows in natural scenes. *Comput Electron Agric* 2023; 205: 107632, DOI: [10.1016/j.compag.2023.107632](https://doi.org/10.1016/j.compag.2023.107632).
67. Chen J, Zhang D, Suzauddola M, et al. Identifying crop diseases using attention embedded MobileNet-V2 model. *Appl Soft Comput* 2021; 113: 107901, DOI: [10.1016/j.asoc.2021.107901](https://doi.org/10.1016/j.asoc.2021.107901).
68. Michele A, Colin V and Santika DD. MobileNet convolutional neural networks and support vector machines for palmprint recognition. *Proc Comput Sci* 2019; 157: 110–117, DOI: [10.1016/j.procs.2019.08.147](https://doi.org/10.1016/j.procs.2019.08.147).
69. Lu Z, Miao J, Dong J, et al. Automatic classification of retinal diseases with transfer learning-based lightweight convolutional neural network. *Biomed Signal Process Control* 2023; 81: 104365, DOI: [10.1016/j.bspc.2022.104365](https://doi.org/10.1016/j.bspc.2022.104365).
70. Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition Challenge. *Int J Comput Vis* 2015; 115(3): 211–252. DOI: [10.1007/s11263-015-0816-y](https://doi.org/10.1007/s11263-015-0816-y).
71. Tarekegn AN, Giacobini M and Michalak K. A review of methods for imbalanced multi-label classification. *Pattern Recogn* 2021; 118: 107965, DOI: [10.1016/j.patcog.2021.107965](https://doi.org/10.1016/j.patcog.2021.107965).