buildings

MDPI

Article

BIM Product Style Classification and Retrieval Based on Long-Range Style Dependencies

Jia Cui [1], Mengwei Zang [2], Zhen Liu [1], Meng Qi [2], Rong Luo [1], Zhenyu Gu [3] and Hongju Lu [4,*]

1 School of Design, South China University of Technology, Guangzhou 510006, China; cuijia1247@scut.edu.cn (J.C.); liuzjames@scut.edu.cn (Z.L.); 202130100388@mail.scut.edu.cn (R.L.)
2 School of Information Science and Engineering, Shandong Normal University, Jinan 250001, China; 2021317153@stu.sdnu.edu.cn (M.Z.); qimeng@sdnu.edu.cn (M.Q.)
3 School of Design, Shanghai Jiaotong University, Shanghai 205530, China; zygu@sjtu.edu.cn
4 School of Management, Guangzhou City University of Technology, Guangzhou 510800, China
* Correspondence: luhj@gcu.edu.cn

Abstract: The rapid increase in building components on the building information model (BIM) object database has created new demand for BIM product recommendations to improve design efficiency. Current efforts mainly focus on the shape and contents of the products, instead of stylistic consistency, which is a crucial factor during the practical design process. To tackle such a problem, this paper proposes a novel framework to capture stylistic features based on long-range design dependencies with structural preservation, of which the snapshots of BIM products have been used to extract the stylistic features; core patches with strong style, generated by the pre-trained saliency model, are the root nodes; stylistic correlations are calculated as the hyperedges by tree-based operations; deep features and design features are proposed to represent the low-level and style distribution based on the study of design theory; and an ensemble learning strategy is introduced to solve the unbalanced classifier performance. An ablation study is conducted to validate the effectiveness of the proposed framework, in which comparative experiments with state-of-the-art baselines demonstrate the advantages of the proposed method.

Keywords: building information model; style classification; BIM product retrieval; deep learning; design feature extraction

check for updates

Citation: Cui, J.; Zang, M.; Liu, Z.; Qi, M.; Luo, R.; Gu, Z.; Lu, H. BIM Product Style Classification and Retrieval Based on Long-Range Style Dependencies. *Buildings* **2023**, *13*, 2280. https://doi.org/10.3390/buildings13092280

Academic Editor: Lucio Soibelman

Received: 14 July 2023
Revised: 13 August 2023
Accepted: 6 September 2023
Published: 7 September 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Building information modelling (BIM), representing a paradigm shift in the architecture, engineering, and construction (AEC) industry, is capable of sharing data during facilities' life cycles by establishing a digital model and rendering different departments in a project for the identification of early errors and cost overruns in time. In the post-pandemic era, BIM has been extensively adopted as a promising technology to optimize efficiency in various fields [1], such as the design [2], construction operation [3], and maintenance phases of construction [4]. Artificial Intelligence (AI) technologies can catalyze the digital generation and automation of the BIM process. Several BIM product libraries, e.g., BIMobject, SmartBIM, and AutodeskSeek, have been created to improve the efficiency of AEC projects.

Faced with the huge number of BIM products, designers manually select satisfactory ones from BIM libraries to create a virtual indoor and outdoor environment, which is a time- and labor-consuming process. Therefore, the BIM recommendation system has recently drawn increasing attention. The most pertinent models can assist users in the areas of e-commerce [5], software engineering [6], smart cities [7], etc. Some studies [8,9] explore the text-based BIM retrieval scheme and case-based reasoning (CBR) [10] strategy to stimulate manual BIM product selection. Semantic information relating to shape and apparencies is used to retrieve BIM models. However, stylistic consistency is one of the

essential principles in the practical design process. Few studies consider BIM product recommendation from the perspective of design styles, which means that the research field is still challenging and interesting.

This paper proposes a novel framework to classify and retrieve BIM models according to both style and shape. To reduce the computational cost, snapshots of 3D models are used to extract visual and stylistic features. Because the 3D BIM models have multiple snapshots with different spatial positions and gestures, the proposed long-range style dependencies are helpful in capturing core stylistic features based on the strong–weak style phenomenon [11] and preserve the shape of the structures by calculating optimal global correlations. The patches with strong–weak stylistic features are the nodes, and the style correlations are used to compute the similarities as edges. The proposed tree operations can generate the optimal path according to the guided minimum spanning tree (MST) algorithms to capture stylistic features globally. The deep features from the pre-trained VGG-16 model are capable of representing the local shape features, and the design distribution features can enhance the robustness of classifiers and retrieval models.

Compared with existing methods, the contributions of the proposed work can be summarized as follows:

1. This study proposes long-range style dependencies based on patches with strong style for global structural preservation. In the visual representation of a 3D BIM product, there are some salient regions with strong stylistic features and other regions with weak ones. The patches with strong style are sampled to rebuild the tree map path based on the proposed long-range similarity computation algorithm. According to the similarities, the proposed stylistic dependencies can cover the patches even with long distances to the root node by the linked hyperedges. Therefore, for BIM product retrieval, 3D models with different spatial gestures and positions but similar styles can be recommended efficiently by the proposed framework.
2. This study introduces the compositional features of both visual and design perspectives by the pre-trained deep model and the design distribution representations. The deep features extracted by deep learning algorithms can represent the low- and semantic-level information from the visual aspect but with data-sensitive problems because of the content noises. The image contents cause matching errors during the style classification. To solve the issue, this paper combines the 69-dimensional design distribution features to improve the robustness, inspired by the design study of style theories.
3. This study presents a novel framework for retrieval of BIM products based on stylistic consistency, instead of the shape contents. Stylistic consistency is one of the crucial principles in the design process. However, few studies explore the BIM product recommendation according to the style. In this paper, the intelligent BIM product retrieval framework can release the manual selection pressure to improve design efficiency.

The rest of this paper is organized as follows: Section 2 presents related works, while the proposed methods are described in Section 3. A discussion on BIM product retrieval application in a small-scale dataset is presented in Section 4. The experimental setup and several comparative experiments are analyzed in Section 5. Finally, the conclusions and future work are discussed in Section 6.

2. Related Works

2.1. BIM Product Recommendation

The concept of BIM, introduced by Eastman et al. [12] in 1974 as a descriptive system for buildings, gained widespread recognition in 2002 after it was studied by Autodesk [13]. With the rapid development in recent years, many BIM product libraries have been released to facilitate digital building and interior design, which has created new demand for intelligent BIM product recommendations for improving the efficiency of the design process [14]. There are two main categories to retrieve BIM products based on text index and designers' labels. For the former one, keyword retrieval technologies are used to select the

potential BIM product due to ease of use. In 2016, Lin et al. [15] proposed an NLP (natural language processing)-based approach to retrieve data from cloud BIM datasets. Li et al. [9] proposed a BIM product retrieval algorithm by calculating the similarity between attributes in 2020. The semantic search engine is established based on BIM-oriented ontology for the contextual meaning of terms and a local context analysis technique for query expansion. Gao et al. [14] proposed a BIM prototype semantic search engine by combing the usability of a keyword-based interface with automatic query expansion techniques. Despite the convenient usage, there were irrelevant results, so a large amount of time had to be spent on searching useful and professional text keywords that meet the design requirements for users [15–17]. For better retrieval performance, the stylistic consistency of BIM products is a pivotal indicator of design quality and an essential supplement during the BIM design process.

For the latter one, the BIM product is retrieved using data on designers' behavior, such as labels. Machine learning and deep learning techniques are widely used based on the designers' labels. Guerrero et al. [10] introduced a decision support system based on case-based reasoning in healthcare building design. Fazeli et al. [18] illustrated a BIM-integrated TOPSIS-Fuzzy framework to optimize the selection of sustainable building components. Lee et al. [19] presented a dynamic BIM component recommendation approach based on probabilistic matrix factorization and a gray model. The designers' behaviors focus on explicit features, such as shape and appearance, with less-implicit ones, i.e., style. These schemes require extra efforts from designers to collect the labels, which will limit their practical application. The proposed framework in this paper is designed to select BIM products automatically based on intelligent algorithms to reduce designers' labor-intensive workloads.

2.2. Style-Based Image Classification

The study of image classification is widely applied in various research areas. The style-related image-classification techniques are less studied than their general counterparts. Only a few study areas, i.e., style image generation [20] and style transfer [21], consider stylistic factors in their algorithms, because of the difficulties of capturing the consensus labels for stylistic factors.

For image style classification, machine learning algorithms are used because the current style-based datasets are in medium and small sizes. However, with the recent advances in deep learning, the pre-trained deep models have accelerated feature extraction in the majority of studies. Chen et al. [22] extracted the different features from various layers in the CNN structure, and the adaptive cross-layer correlation was used to classify the image styles. Joshi et al. [23] extracted the features through an auto-encoder with random self-disturbance and trained the classifier in four machine-learning models for style recognition. Liu et al. [24] calculated the color histogram and color semantic distance by SVM for traditional Chinese painting classification. The pre-trained VGG-16 model is helpful for extracting local and semantic features, combining multi-scale strategy, image augmentation, and random transformations [25,26]. The deep correlation features [27] were studied to enlarge the texture and lightness factors for better style recognition. Nguyen et al. [28] proposed an extended OCR model to recognize font style in visually rich documents. Zhang et al. [29] enhanced the stylistic features via a contrastive learning scheme. Despite promising results, the deep representation learning schemes extract the image features based on limited receptive fields, which leads to deficiencies in the long-range and structured ones. This paper proposes a stylistic feature extraction based on a tree operational layer which can capture long-range dependencies with structural preservation.

3. Methods

3.1. Research Framework

For BIM product recommendation, similarity computation is costly in 3D space. As a result, Zhou et al. [30] proposed the DrStyle model to use style learning algorithms on 2D

image snapshots with the most style richness information. Inspired by this work, the stylistic features extracted from snapshots of 3D models are used to classify the different styles and make stylistic recommendations. There are three steps in the proposed framework shown in Figure 1: the calculation of long-range style dependencies, the compositional features with deep and design perspectives, and the ensemble learning strategy.

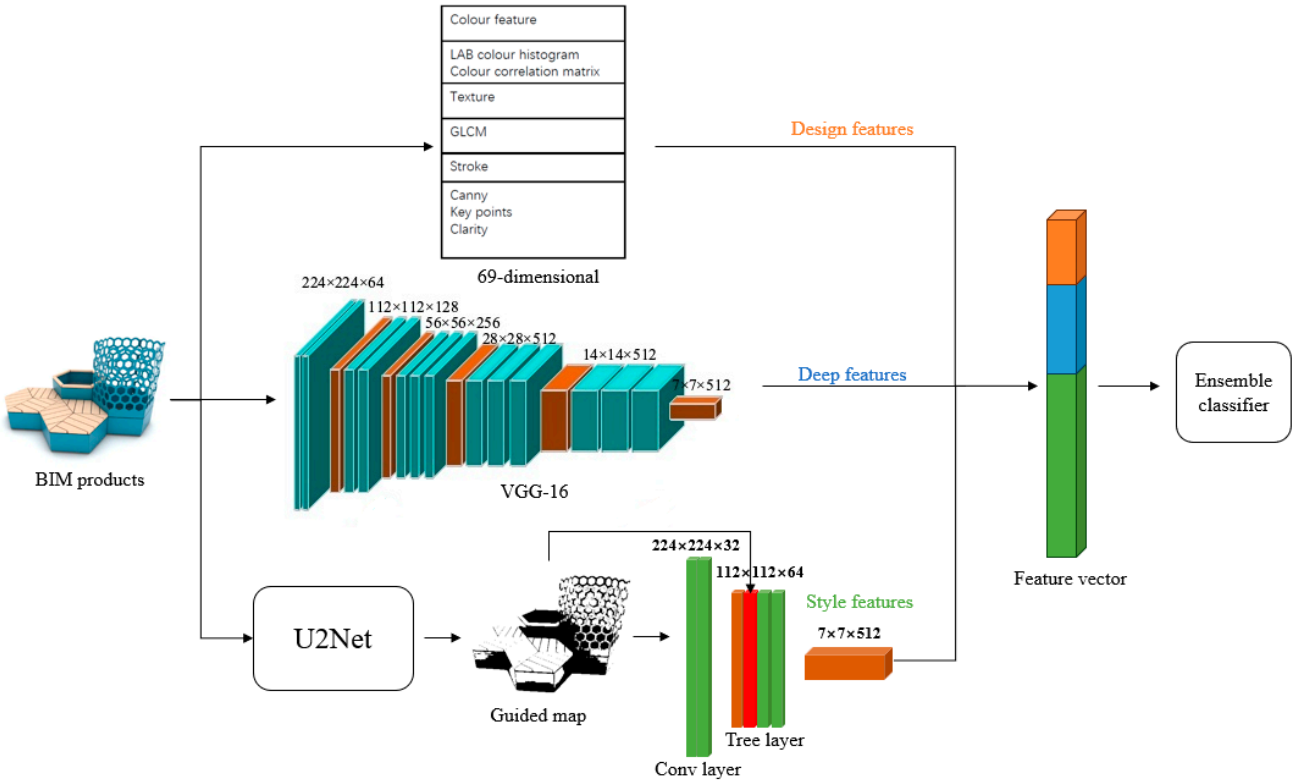


Figure 1. The proposed framework.

Visual saliency computation has been widely applied in many studies to locate and recognize the areas that draw the most prominent visual attention [31–33]. Based on the strong–weak style phenomenon [11], there are different regions in the target images with different stylistic strengths, which is the reason why people always have various impressions about the stylistic categories of the same image. This paper uses the style saliency map generated by the pre-trained saliency model as the guided map. The proposed long-range style dependencies can be calculated by the tree-based operations to determine the optimal structural relationship between strong and weak patches. In other words, the tree layer, shown in Figure 1, works as a filter to calculate the patch-based stylistic features along the hyperedges instead of the reception fields used in common VGG. The guided map outputted from the pre-trained U2Net model is used to build trees with stylistic similarities. Only the patches along the tree hyperedges are considered for stylistic feature computation in the long-range distance. In this way, the strong stylistic features can be used to improve the style-based classification and recommendations, and details of this can be found in Section 3.2.

Deep learning models make solid contributions to image feature representation. Many studies [34–36] adopt deep learning, e.g., VGG-16 [37], as the feature extractor in image classification and regression tasks. However, the mainstream deep models are designed for content classification instead of style classification. Influenced by the content-based features, the performance of style classification in existing deep classification models is not stable enough because of data-sensitive issues [38]. To solve this problem, the design distribution features are proposed based on reference [39], such as texture, color, and fundamental lines.

There are 69-dimensional design features proposed based on the distribution of model characteristics to improve the robustness of deep features.

Combining the deep, design, and stylistic features, the selectKBest algorithm is used to decrease the dimensionalities based on the loss of information entropies to reduce the computational complexity for the small-sized style database. Considering the unbalanced classifier performance, the ensemble learning strategy has been applied. Different classifiers have different performances on different types of data distribution, i.e., the DT works well on the linear dataset, SVM has difficulties when the dataset is large, and RF performs better in high-dimensional datasets. To reduce the various influences of different classifiers, the ensemble strategy is applied. There are 8 classic machine learning classifiers, namely, random forest, SVM, linear analysis, linear regression, KNN, Adaboost, decision tree, and GBDT. When the feature vectors are generated, the 8 classifiers make predictions by themselves. The final decision is made by the 8 classifiers based on the mode principle. As all the classifiers are lightweight, the ensemble learning approach has a close computational efficiency to the single one.

3.2. Long-Range Style Dependencies

To reduce the style recognition conflicts, only the image patches with strong stylistic features are considered for the BIM product style recognition. In three-dimensional space, various spatial gestures and positions lead to quite different 2D snapshots. DrStyle [30] introduces the image evaluation module to select the one with rich stylistic features. Different from DrStyle, it is found that slight spatial differences will not cause different style saliencies at the patch level, as shown in Figure 2. There are four different spatial snapshots of the BIM model "Ekko 1 Seater Sofa". The patch regions with the most significant saliency (the red circle), generated by the pre-trained U2Net [40] model, are not changed too much. This means that regions with strong stylistic features will not be changed when the spatial positions and gestures are slightly different. Based on this, this paper proposes to use the salient patches as guided clues to calculate the style saliency map.



Figure 2. The salient patches by U2Net [33].

The current existing deep models adopt stacks of local convolutions to capture the context feature by the convolution kernels in 3×3, 5×5, and 7×7. These operations ignore the objects' details when enlarging the receptive field. Therefore, this paper proposes tree-based operations to capture the style dependencies along the optimal path based on the saliency-guided image.

This paper represents the style saliency dependencies as an undirected graph G = (V, E), with the dissimilarity weight ω for edges. The vertices V are the salient patch features, and their interconnections can be denoted as E. The saliency dependencies feature map is adopted as the guidance for the 4-connected planar graph construction illustrated in Figure 3. In Figure 3, for the feature map generated by the convolutional layers, the saliency map is the guided map which can be used to calculate the similarity value to organize the spanning tree. Thus, a spanning tree can be generated from the graph by performing a pruning algorithm to remove the edge with a substantial dissimilarity. The root node of

the spanning tree is the sample patch from the convolutional results of the style saliency map generated by a pre-trained U2Net [40]. From this perspective, graph G becomes the minimum spanning tree (MST) whose sum of dissimilarity weights is the minimum of all spanning trees. In the guided MST algorithm, the edge weight depends on two factors, the patch pixel feature embedding and the similarity value with its 4-connected neighbors. Different from the conventional MST algorithm, we sequentially index the edges based not only on the weights but also on the saliency value guided by the outputs from the pre-trained U2Net [40] model. In this way, only the pixels in the salient regions are considered to be the root, and the MST tree represents the relationships between the target objects and their background for image structural preservation. We use the improved MST algorithm proposed by Zhong et al. [41] to calculate the whole graph of the salient pixels.

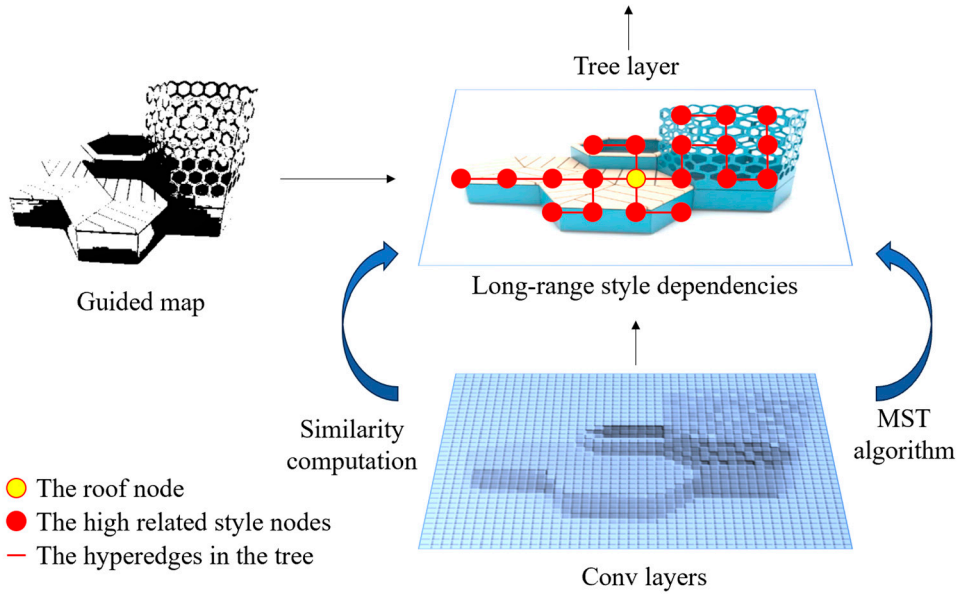


Figure 3. The long-range style dependency.

Motivated by the tree algorithm, the generic tree operation module by deep learning algorithm can be formulated as

$$y_i = \frac{1}{z_i}\sum_{\Omega} S(E_{i,j}) f(x_j) \tag{1}$$

where i and j indicate the index of patches, Ω denotes the set of all patches in the tree G, x represents the target patch feature vector, and y is the output stylistic features related to the long-range dependencies with x. $E_{i,j}$ is the hyperedge between i and j, which includes the intermediate nodes in the optimal path from i to j. f represents the feature embedding transformation to generate vectors based on image patch i. z is the summation of similarity calculated by function $S(E_{i,j})$, in Equation (2):

$$S(E_{i,j}) = \exp(-D(i,j)), where\ D(i,j) = D(j,i) = \sum_{k,m \in E_{i,j}} \omega_{k,m} \tag{2}$$

According to Equations (1) and (2), the tree operation can be treated as a weighted-average filter. The $\omega_{k,m}$ indicates the dissimilarity between adjacent patches that can be approximately computed as the Euclidean distance for simplicity. The target feature vectors from the style saliencies guided map x and the dissimilarity $\omega_{k,m}$ are responsible for the output stylistic feature y. Therefore, the derivative of the output concerning the input variables can be derived as Equations (3) and (4).

$$\frac{\partial y_i}{\partial x_i} = \frac{S(E_{i,j})}{z_i} \frac{\partial f(x_j)}{\partial x_i} \tag{3}$$

$$\frac{\partial y_i}{\partial \omega_{k,m}} = \frac{S(E_{i,j})}{z_i} \frac{\partial \sum_{k,m \in E_{i,j}} \omega_{k,m}}{\partial \omega_{k,m}} \tag{4}$$

where Equations (4) and (5) are defined with the children of vertex m in the tree operations whose root node is the vertex i. In this way, the proposed tree operations can be formulated as a differentiable module. The tree layer is used as a style relation accumulator to calculate the stylistic features along the high-related path, which is the replacement of the common local region calculation. Therefore, the long-range style dependencies can be calculated based on the sampled salient patches for better structural preservation.

3.3. Deep and Design Feature Computation

Pre-trained deep networks [42,43] in image classification have been widely adopted; the deep features can steadily represent the image details and semantics. In this paper, the VGG-16 model is used, in which 13 convolutional layers are chosen as the feature extractor for the BIM product snapshots' deep features. Besides that, inspired by the design style study, the designers make style decisions based on the distribution analysis, for example, the scientific style is full of light effects and the color red represents the classic Chinese style. Therefore, we proposed 69-dimensional design features with mode characteristics in the form of color, texture, and fundamental strokes. All the design features are capable of improving the robustness of deep features to avoid any data-sensitive issues.

For the color features, the color correlation matrix and color histogram are used to represent the colors at the pixel level and their neighboring areas. The value of mean, variance, skewness, and kurtosis can be calculated by Equations (5)–(8). For the color histogram, in the LAB color space, the mean, variance, and the first, second, and third quantile and entropy are calculated.

$$Mean = \frac{1}{N} \sum_{i=0}^{N} x_i \tag{5}$$

$$Variance = \frac{1}{N} \sum_{i=1}^{N} (x_i - Mean)^2 \tag{6}$$

$$Skewness = \frac{\frac{1}{N} \sum_{i=1}^{N} (x_i - Mean)^3}{Variance^3} \tag{7}$$

$$Kurtosis = \frac{\frac{1}{N} \sum_{i=1}^{N} (x_i - Mean)^4}{Variance^4} - 3 \tag{8}$$

For the texture feature, a gray-level co-occurrence matrix (GLCM) is used to depict the local texture and gray-degree distribution, i.e., the ASM, CON, ENG, and IDM, as shown in Equations (9)–(13):

$$P(i,j) = \frac{\sum_{x=1}^{n} \sum_{y=1}^{m} I(x,y) \times I(x + d_x,\ y + d_y)}{\sum_{x=1}^{n} \sum_{y=1}^{m} I(x,y)} \tag{9}$$

$$ASM = \sum_{i=0}^{gray_1} \sum_{j=0}^{gray_1} P(i,j)^2 \tag{10}$$

$$CON = \sum_{i=0}^{gray_1} \sum_{j=0}^{gray_1} (i-j)^2 P(i,j) \tag{11}$$

$$ENG = \sum_{i=0}^{gray_1} \sum_{j=0}^{gray_1} P(i,j) \log P(i,j) \tag{12}$$

$$IDM = \sum_{i=0}^{gray_1} \sum_{j=0}^{gray_1} \frac{P(i,j)}{1 + (i-j)^2} \tag{13}$$

where *I(x,y)* represents the pixel at position (x, y) within the image with width n and height m. The offsets d_x and d_y represent the number of pixels that the image moves horizontally and vertically, which illustrates the gray-level co-occurrence matrix in different directions. The ASM represents the summation of the GLCM; the CON represents the neighbor's gray relationship at the pixel level; the ENG represents the distribution status of the gray value; and the IDM depicts the changing balance in the local patch.

For the stroke features, the Canny operator, FAST algorithm, and Sobel operator are used to depict the contour, key points, and clarity features [44]. The 69-dimensional design features are listed in Table 1.

Table 1. The 69-dimensional design features.

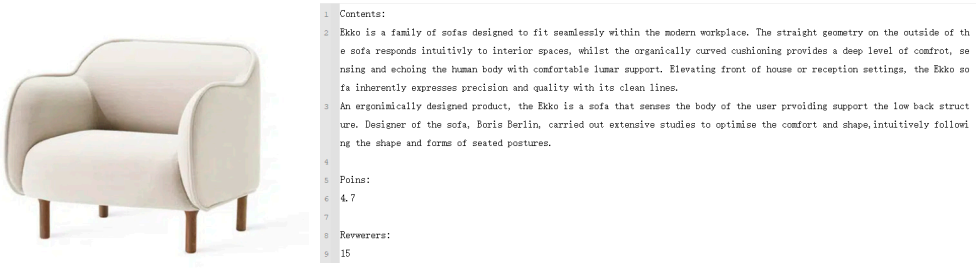
Design Features	Feature Name	Dimensions
Lab color features	Lab_mean, Lab_std, Lab_quantile25, Lab_quantile50, Lab_ quantile75, Lab_ mode	6 dimensions per channel for L, a, b; 18 dimensions total
Color histogram	Lab_L_hist_entropy, Lab_a_hist_entropy, Lab_b_hist_entropy	3 dimensions total
Texture (GLCM)	CoOccurrence_Con, CoOccurrence_Eng, CoOccurence_Asm, CoOccurrence_Idm	4 dimensions in 8 connectivity directions; 32 dimensions total
Stroke—keypoints	keypoints_with_nonmaxSuppression, keypoints_without_nonmaxSuppression, keypoints_difference	3 dimensions total
Stroke—outline	long_line_count, long_line_length, long_line_maxLength, long_line_minLength, Long_line_meanLength, long_line_angle_entropy, short_line_count, short_line_length, short_line_maxLength, short_line_minLength, short_line_meanLength, short_line_meanLength, short_line_angle_entropy, short_entropy_difference	13 dimensions total
Total dimensions		69 dimensions

4. BIM Retrieval Results

The BIM products from the BimObject library are collected. There are two categories of models downloaded from the website, the "chairs" and the "desks & tables". There are 988 chair models and 803 table models; only the official images are used as snapshots of the 3D BIM products for fair comparison. The style of the BIM products is specified by designers according to the descriptions in the library. For example, the snapshot and descriptions of the "Ekko 1 Seater Sofa" are shown in Figure 4a; the style is "modern" and "clean lines" are labelled because of the reviewers' description, as shown in Figure 4b. Excluding the models without review descriptions and highly similar images, there are 679 chair models and 538 table models used to build the style retrieval dataset in this study.

Three retrieval approaches, such as ZS-SBIR [45], DrStyle [30], and the proposed one, are chosen to compare the BIM product retrieval, as shown in Figure 5. In Figure 5, the ZS-SBIR approach retrieves the models based mainly on the shape contents. As a result, models with similar colors and shapes can be recommended. DrStyle can recognize the structures of models based on the optimal image features, such as the leg structure of chairs. However, only features with similar spatial positions can be found using the local feature-matching strategy. There are several objects with similar spatial positions but they are recommended in different styles. For the proposed work, because of the long-range style dependencies, the bar chairs with "comfortable" and "formal" styles can be retrieved,

and the tables with "simply robust" and "natural beauty" styles can be recommended, regardless of color differences and slightly different spatial gestures.



(**a**) (**b**)

Figure 4. The BIM style retrieval dataset from BimObject. (**a**) Snapshot; (**b**) the reviewers' description collected from the BimObject library.



Figure 5. The BIM product retrieval experiments. (**a**) Chair retrieval with the style of "comfortable" and "formal"; (**b**) table retrieval with the style of "simply robust" and "natural beauty".

5. Experiments

5.1. Experimental Setup

The experimental environment is Windows 10 system, Intel(R)Xeon(R) Gold 6226R CPU 2.9 GHz, 256G DDR3 and 16T hard drives, the graphic card is RTX V100. To validate the performance of the proposed framework in terms of style classification, there are two public drawing art datasets, Pandora [46] and Painting91 [47], and one architecture style dataset, Arch [48]. The proposed method is capable of calculating the style of 3D BIM objects based on 2D snapshots. Therefore, it can be used for 2D image datasets. Because most current style classification algorithms are studied on 2D image datasets, this study tests the performance of the proposed framework in 2D datasets as well for fairness of comparison. The Pandora and Painting91 datasets cover various styles of paintings, such

as "Impressionism", "Cubism", "Renaissance", and "Romanticism". For better comparison, we randomly select 100 images per style class from the datasets for training and 50 images for testing. In the Arch dataset, there are 423 architectural images from four styles: Gothic, Islamic, Korean, and Georgian. The training to testing ratio is 8:2.

For the experiment details, the 13 convolutional layers in VGG-16 are used to finetune the deep features in datasets. The SGD optimizer is used, the learning rate is 0.01, the batch size is 200, and after 50 epochs, the deep feature extraction model is trained. The necessary image-enhanced techniques are used for model generality, such as crop, flip, and rotate. Because there are 69-dimensional design features, too many dimensions of deep features will cause an imbalance. Therefore, it uses the feature selection algorithm—selectKBest [49]—to select the top 200 informative deep features based on the information entropy. The long-range dependency stylistic features are also calculated by convolutional operations. The same feature selection strategy also works, and there is a total of 300 informative stylistic features that are used.

The ablation study is used to show the role of deep features, design features, and long-range style dependencies, described in Section 5.2. The quantitative comparisons of the accuracy of style classification are described in Section 5.3 with several new style classification studies.

5.2. Ablation Study

In this paper, the deep features, design features, and stylistic features are the three main components of the proposed framework. The first ablation study is designed to test the effectiveness of deep and design features in the Pandrola dataset for style classification. There are 20 groups of data with 1 K images per group, roughly. For test 1, only the deep features are extracted by the VGG-16 model and the composition features of both deep and design features are used in test 2. The model produced the results shown in Figure 6. For the deep features, the accuracy fluctuates greatly between 0.73 and 0.83 for the data-sensitive issues. The different contents with various backgrounds will reduce the accuracy of deep features. For the compositional features, the average accuracy is stable when kept at a higher stage, i.e., (0.80–0.85). Therefore, the combination of deep and design features is helpful for improving the robustness of the proposed framework.

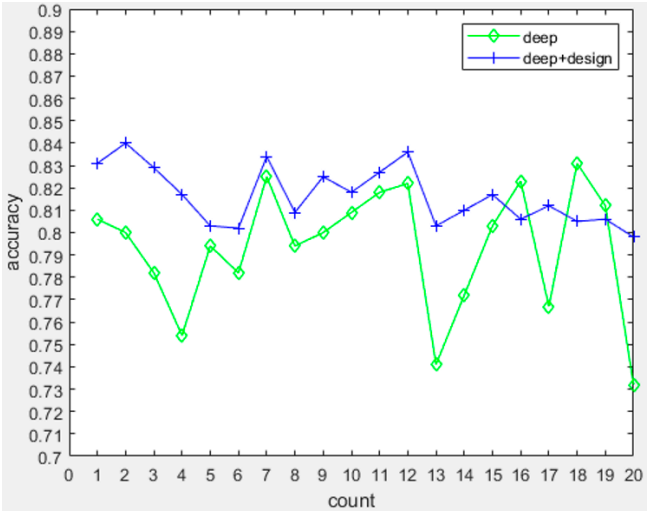


Figure 6. The ablation study of deep and design features.

Then, the deep, design, and long-range stylistic features verify their performance in the three datasets; the results are shown in Table 2. For the deep and design features, the accuracy is not higher than 0.79. The performance and accuracy stability is improved somewhat when combining the two, around 0.83. With the long-range stylistic features, the performance improved obviously (the average accuracy increased by 4%, 7%, and 4% in

Pandora, Painting-91, and Arch datasets, respectively), which shows the effectiveness of the proposed feature extraction strategy.

Table 2. Ablation study in three datasets.

	Pandora	**Painting-91**	**Arch**
Deep feature	0.79	0.73	0.65
Design feature	0.71	0.68	0.66
Deep + Design	0.79	0.78	0.75
Proposed	**0.83**	**0.85**	**0.79**

5.3. Quantitative Experiments

The results of the quantitative experiments on the three datasets with four new style classification studies are shown in Table 3. Gairola et al. [42] proposed an auto-encoder structure to represent the image, and the adversarial loss was applied to distinguish the different styles. Nag et al. [43] introduced a data augmentation algorithm to enhance the styles, and a transfer learning scheme was proposed. Milani et al. [50] combined the VGG network with the SVM classifier to perform the style classification at the vector output from the final convolutional layer. Ghosal et al. [51] extracted visual geometric features by CNN to classify the styles. For a fair comparison, all the experiments conducted on the codes and pre-trained model were published by the authors or can be downloaded from their websites. The evaluation metrics are F1 and ACC values. From the three datasets, the proposed framework achieves the top two positions in all three datasets with the SOTA effects.

Table 3. The style classification results in three datasets.

	Reference [34]		**Reference [35]**		**Reference [40]**		**Reference [41]**		**Proposed**	
	F1	**Acc**	**F1**	**Acc**	**F1**	**Acc**	**F1**	**Acc**	**F1**	**Acc**
Pandora	0.72	0.71	0.78	0.77	0.80	**0.79**	0.78	0.78	0.81	**0.83**
Painting-91	0.74	0.74	0.79	**0.80**	0.77	0.76	0.75	0.77	0.84	**0.85**
Arch	0.67	0.69	0.72	**0.73**	0.69	0.69	0.67	0.68	0.75	**0.79**

6. Conclusions

In this paper, a novel BIM product style classification and retrieval framework is proposed. A 2D snapshot of 3D models is used to extract stylistic features in order to reduce the computational cost. The introduced long-range style dependencies are implemented to capture strong patch-based stylistic features and to preserve the model structures through the use of a tree layer similarity calculation. The compositional features combined with the deep and design features help enhance the robustness of the model. Compared with other machine learning-based algorithms, the style classification on the 2D image dataset and the BIM product recommendation on 3D models have shown the effectiveness and advantages of the proposed method.

However, the proposed method cannot handle two aspects of style-based retrieval tasks, that is, snapshots with complex backgrounds and BIM models with complex fine structures. The former drawback is due to the possibility that the complex background could influence the saliency estimation based on U2Net, which could reduce the guided MST algorithm's ability to find the optimal long-range style dependency. In addition, complex fine structures, i.e., fractal structures, cannot be calculated because of the distribution features in the proposed design's stylistic features. Further, the obvious spatial differences of the snapshots of 3D BIM objects could reduce the retrieval accuracy for occlusive perspectives.

For future studies, the point-cloud representation of 3D models will offer a potential solution to the comprehensive retrieval of 3D objects. The study of disentangled representation learning of style and content in language processing models presents an inspiring way

to extract the stylistic features without the entanglement of content factors. As such, the above two directions will be the next step to improving the style-based BIM object retrieval study in the future.

Author Contributions: Supervision, H.L.; Conceptualization, Z.L. and Z.G.; Methodology, J.C. and M.Q.; Implementation and Validation, M.Z. and R.L.; Writing—original draft preparation, J.C. and H.L.; Writing—review and editing, Z.L. and H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fundamental Research Funds for the Central Universities (2022ZYGXZR020), and Joint Funds of Natural Science Foundation of Shandong Province (ZR2021LZL011).

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The authors thank all the people who have supported this research, including anonymous reviewers for their constructive comments and suggestions to improve this paper, and the School of Design, South China University of Technology, for providing a rich and diverse academic environment to support this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shishehgarkhaneh, M.B.; Keivani, A.; Moehler, R.C.; Jelodari, N.; Laleh, S.R. Internet of Things (IoT), Building Information Modeling (BIM), and Digital Twin (DT) in Construction Industry: A Review, Bibliometric, and Network Analysis. *Buildings* **2022**, *12*, 1503. [CrossRef]
2. Atencio, E.; Araya, P.; Oyarce, F.; Herrera, R.F.; Rivera, F.M.-L.; Lozano-Galant, F. Towards the Integration and Automation of the Design Process for Domestic Drinking-Water and Sewerage Systems with BIM. *Appl. Sci.* **2022**, *12*, 9063. [CrossRef]
3. Shishehgarkhaneh, M.B.; Moradinia, S.F.; Keivani, A.; Azizi, M. Application of Classic and Novel Metaheuristic Algorithms in a BIM-Based Resource Tradeoff in Dam Projects. *Smart Cities* **2022**, *5*, 1441–1464. [CrossRef]
4. Zhou, X.; Wang, M.; Liu, Y.S.; Wang, Q.; Guo, M.; Zhao, J. Heterogeneous network modeling and segmentation of building information modeling data for parallel triangulation and visualization. *Autom. Constr.* **2021**, *131*, 103897. [CrossRef]
5. Karthik, R.; Ganapathy, S. A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce. *Appl. Soft Comput.* **2021**, *108*, 107396. [CrossRef]
6. Almonte, L.; Guerra, E.; Cantador, I.; De Lara, J. Recommender systems in model-driven engineering: A systematic mapping review. *Softw. Syst. Model.* **2022**, *21*, 249–280. [CrossRef]
7. Abakarim, S.; Qassimi, S.; Rakrak, S. Review on Recent Trends in Recommender Systems for Smart Cities. In Proceedings of the 3rd International Conference on Artificial Intelligence and Computer Vision (AICV2023), Marrakesh, Morocco, 5–7 March 2023; Springer: Cham, Switzerland, 2023; pp. 317–326.
8. Wang, H.; Meng, X.; Zhu, X. Improving knowledge capture and retrieval in the BIM environment: Combining case-based reasoning and natural language processing. *Autom. Constr.* **2022**, *139*, 104317. [CrossRef]
9. Li, N.; Li, Q.; Liu, Y.S.; Lu, W.; Wang, W. BIMSeek++: Retrieving BIM components using similarity measurement of attributes. *Comput. Ind.* **2020**, *116*, 103186. [CrossRef]
10. Guerrero, J.; Miró-Amarante, G.; Martín, A. Decision support system in health care building design based on case-based reasoning and reinforcement learning. *Expert Syst. Appl.* **2022**, *187*, 116037. [CrossRef]
11. Jiang, S.; Shao, M.; Jia, C.; Fu, Y. Learning Consensus Representation for Weak Style Classification. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 2906–2919. [CrossRef] [PubMed]
12. Eastman, C. *An Outline of the Building Description System; Research Report, No. 50*; Education Resources Information Center: Online, 1974.
13. Schönfelder, P.; Aziz, A.; Faltin, B.; König, M. Automating the retrospective generation of As-is BIM models using machine learning. *Autom. Constr.* **2023**, *152*, 104937. [CrossRef]
14. Gao, X.; Pishdad-Bozorgi, P. BIM-enabled facilities operation and maintenance: A review. *Adv. Eng. Informatics* **2019**, *39*, 227–247. [CrossRef]
15. Lin, J.-R.; Hu, Z.-Z.; Zhang, J.-P.; Yu, F.-Q. A Natural-Language-Based Approach to Intelligent Data Retrieval and Representation for Cloud BIM. *Comput. Civ. Infrastruct. Eng.* **2016**, *31*, 18–33. [CrossRef]
16. Ismagilova, E.; Hughes, L.; Dwivedi, Y.K.; Raman, K.R. Smart cities: Advances in research—An information systems perspective. *Int. J. Inf. Manag.* **2019**, *47*, 88–100. [CrossRef]
17. Shin, S.; Issa, R.R.A. BIMASR: Framework for Voice-Based BIM Information Retrieval. *J. Constr. Eng. Manag.* **2021**, *147*, 4021124. [CrossRef]

18. Fazeli, A.; Jalaei, F.; Khanzadi, M.; Banihashemi, S. BIM-integrated TOPSIS-Fuzzy framework to optimize selection of sustainable building components. *Int. J. Constr. Manag.* **2022**, *22*, 1240–1259. [CrossRef]
19. Lee, P.-C.; Long, D.; Ye, B.; Lo, T.-P. Dynamic BIM component recommendation method based on probabilistic matrix factorization and grey model. *Adv. Eng. Informatics* **2020**, *43*, 101024. [CrossRef]
20. Kwon, G.; Ye, J.C. Diffusion-based image translation using disentangled style and content representation. *arXiv* **2022**, arXiv:2209.15264.
21. Jing, Y.; Mao, Y.; Yang, Y.; Zhan, Y.; Song, M.; Wang, X.; Tao, D. Learning graph neural networks for image style transfer. In *European Conference on Computer Vision*; Springer: Cham, Switzerland, 2022; pp. 111–128.
22. Chen, L.; Yang, J. Recognizing the style of visual arts via adaptive cross-layer correlation. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 2459–2467.
23. Joshi, A.; Agrawal, A.; Nair, S. Art Style Classification with Self-Trained Ensemble of AutoEncoding Transformations. *arXiv* **2020**, arXiv:2012.03377.
24. Liu, C.; Jiang, H. Classification of traditional Chinese paintings based on supervised learning methods. In Proceedings of the 2014 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Guilin, China, 5–8 August 2014; pp. 641–644.
25. Cetinic, E.; Lipic, T.; Grgic, S. Learning the Principles of Art History with convolutional neural networks. *Pattern Recognit. Lett.* **2020**, *129*, 56–62. [CrossRef]
26. Cetinic, E.; Lipic, T.; Grgic, S. A deep learning perspective on beauty, sentiment, and remembrance of art. *IEEE Access* **2019**, *7*, 73694–73710. [CrossRef]
27. Chu, W.-T.; Wu, Y.-L. Deep Correlation Features for Image Style Classification. In *Book Deep Correlation Features for Image Style Classification, Series Deep Correlation Features for Image Style Classification*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 402–406.
28. Nguyen, V.D.; Dong, N.T.; Tuan, D.M.; Ngoc, N.T.A.; Lam, N.V.S.; Anh, N.V.; Dang, N.H. Extending OCR Model for Font and Style Classification. In Proceedings of the 17th Conference on Information Technology and its Applications, Turin, Italy, 20–22 October 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 193–204.
29. Zhang, Y.; Tang, F.; Dong, W.; Huang, H.; Ma, C.; Lee, T.-Y.; Xu, C. Domain enhanced arbitrary image style transfer via contrastive learning. In Proceedings of the ACM SIGGRAPH 2022 Conference Proceedings, Vancouver, BC, Canada, 7–11 August 2022; pp. 1–8.
30. Zhou, X.; Ma, C.; Wang, M.; Guo, M.; Guo, Z.; Liang, X.; Han, J. BIM product recommendation for intelligent design using style learning. *J. Build. Eng.* **2023**, *73*, 106701. [CrossRef]
31. Qian, X.; Zeng, Y.; Wang, W.; Zhang, Q. Co-Saliency Detection Guided by Group Weakly Supervised Learning. *IEEE Trans. Multimedia* **2022**, *25*, 1810–1818. [CrossRef]
32. Yun, H.; Lee, S.; Kim, G. Panoramic Vision Transformer for Saliency Detection in 360° Videos. In Proceedings of the 17th European Conference on Computer Vision, Tel Aviv, Israel, 23 October 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 422–439.
33. Cui, J.; Zhao, Y.; Dong, X.; Tang, M. Unsupervised segmentation of multiview feature semantics by hashing model. *Signal Process.* **2019**, *160*, 106–112. [CrossRef]
34. Debbagh, M. Learning Structured Output Representations from Attributes using Deep Conditional Generative Models. *arXiv* **2023**, arXiv:2305.00980.
35. Chen, W.; Liu, Y.; Wang, W.; Bakker, E.M.; Georgiou, T.; Fieguth, P.; Liu, L.; Lew, M.S. Deep learning for instance retrieval: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 7270–7292. [CrossRef]
36. Li, X.; Cui, J.; Song, J.; Jia, M.; Zou, Z.; Ding, G.; Zheng, Y. Contextual Features and Information Bottleneck-Based Multi-Input Network for Breast Cancer Classification from Contrast-Enhanced Spectral Mammography. *Diagnostics* **2022**, *12*, 3133. [CrossRef]
37. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, 7–9 May 2015.
38. Lu, H.; Yue, P.; Zhao, Y.; Liu, R.; Fu, Y.; Zheng, Y.; Cui, J. Contour sensitive saliency and depth application in image retargeting. In Proceedings of the 9th International Conference on Graphic and Image Processing (ICGIP 2017), International Society for Optics and Photonics, Qingdao, China, 14–16 October 2017; p. 106151R.
39. De Perthuis, K. Fashion's image: The complex world of the fashion photograph. In *A Companion to Photography*; Wiley: Hoboken, NJ, USA, 2019; pp. 253–274.
40. Jiang, G.; Zhang, W.; Wang, W.; Sun, X. Saliency Detection of Logistics Packages Based on Deep Learning. In Proceedings of the 2022 IEEE 2nd International Conference on Computer Systems (ICCS), Qingdao, China, 23–25 September 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 85–92.
41. Zhong, C.; Malinen, M.; Miao, D.; Fränti, P. A fast minimum spanning tree algorithm based on K-means. *Inf. Sci.* **2015**, *295*, 1–17. [CrossRef]
42. Gairola, S.; Shah, R.; Narayanan, P. Unsupervised Image Style Embeddings for Retrieval and Recognition Tasks. In Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass Village, CO, USA, 1–5 March 2020; pp. 3270–3278.

43. Nag, P.; Sangskriti, S.; Jannat, M.E. A Closer Look into Paintings' Style Using Convolutional Neural Network with Transfer Learning. In Proceedings of the Proceedings of International Joint Conference on Computational Intelligence, Budapest, Hungary, 2–4 November 2020; Springer: Singapore, 2020; pp. 317–328.
44. Zokhidov, A.; Rikhsivoev, M. Review of Edge Detection Methods in Images: Benefits, Limitations and Development Prospects. *Proc. Int. Conf. Mod. Sci. Sci. Stud.* **2023**, *2*, 284–294.
45. Dutta, T.; Biswas, S. Style-Guided Zero-Shot Sketch-based Image Retrieval. *Proc. BMVC* **2019**, *2*, 9.
46. Florea, C.; Condorovici, R.; Vertan, C.; Butnaru, R.; Florea, L.; Vranceanu, R. Pandora: Description of a painting database for art movement recognition with baselines and perspectives. In Proceedings of the 2016 24th European Signal Processing Conference (EUSIPCO), Budapest, Hungary, 28 August–2 September 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 918–922.
47. Khan, F.S.; Beigpour, S.; van de Weijer, J.; Felsberg, M. Painting-91: A large scale database for computational painting categorization. *Mach. Vis. Appl.* **2014**, *25*, 1385–1397. [CrossRef]
48. Chu, W.-T.; Tsai, M.-H. Visual pattern discovery for architecture image classification and product image search. In Proceedings of the 2nd ACM International Conference on Multimedia Retrieval, Hong Kong, China, 5–8 June 2012; pp. 1–8.
49. Desyani, T.; Saifudin, A.; Yulianti, Y. Feature selection based on naive bayes for caesarean section prediction. *IOP Conf. Series Mater. Sci. Eng.* **2020**, *879*, 12091. [CrossRef]
50. Milani, F.; Fraternali, P. A Dataset and a Convolutional Model for Iconography Classification in Paintings. *J. Comput. Cult. Herit.* **2021**, *14*, 1–18. [CrossRef]
51. Ghosal, K.; Prasad, M.; Smolic, A. A geometry-sensitive approach for photographic style classification. *arXiv* **2019**, arXiv:1909.01040.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.