

Article

# Recovery of Natural Scenery Image by Content Using Wiener-Granger Causality: A Self-Organizing Methodology

Cesar Benavides-Alvarez <sup>†</sup>, Carlos Aviles-Cruz <sup>\*,†</sup>, Eduardo Rodriguez-Martinez <sup>†</sup>,  
Andrés Ferreyra-Ramírez <sup>†</sup> and Arturo Zúñiga-López <sup>†</sup>

Electronics Department, Metropolitan Autonomous University, Av. San Pablo 180, Col. Reynosa, Mexico City C.P. 02200, Mexico; cesarbenavides@azc.uam.mx (C.B.-A.); erm@azc.uam.mx (E.R.-M.); fra@azc.uam.mx (A.F.-R.); azl@azc.uam.mx (A.Z.-L.)

\* Correspondence: caviles@azc.uam.mx; Tel.: +52-55-53-18-90-30

† These authors contributed equally to this work.

**Abstract:** One of the most important applications of data science and data mining is organizing, classifying, and retrieving digital images on Internet. The current focus of the researchers is to develop methods for the content based exploration of natural scenery images. In this research paper, a self-organizing method of natural scenes images using Wiener-Granger Causality theory is proposed. It is achieved by carrying out Wiener-Granger causality for organizing the features in the time series form and introducing a characteristics extraction stage at random points within the image. Once the causal relationships are obtained, the k-means algorithm is applied to achieve the self-organizing of these attributes. Regarding classification, the  $k - NN$  distance classification algorithm is used to find the most similar images that share the causal relationships between the elements of the scenes. The proposed methodology is validated on three public image databases, obtaining 100% recovery results.

**Keywords:** image classification; image retrieval; Wiener-Granger causality



**Citation:** Benavides-Álvarez, C.; Avilés-Cruz, C.; Rodríguez-Martínez, E.; Ferreyra-Ramírez, A.; Zúñiga-López, A. Recovery of Natural Scenery Image by Content Using Wiener-Granger Causality: A Self-Organizing Methodology. *Appl. Sci.* **2021**, *11*, 8795. <https://doi.org/10.3390/app11198795>

Academic Editor: Byung-Gyu Kim

Received: 5 August 2021

Accepted: 17 September 2021

Published: 22 September 2021

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



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

## 1. Introduction

With the increasing usage of internet and digital gadgets, content-based image retrieval (CBIR) has grown and been applied in fields such as artificial vision and artificial intelligence [1]. Currently, improvements have been reported in new CBIR approaches and several effective algorithms have been established that allow searching and retrieving images (by content) from an input image [2–6]. The application areas include: fashion, people identification, e-commerce recovery products, remote sensing recovery images, brand images recovery, natural scenes recovery, among others [7,8].

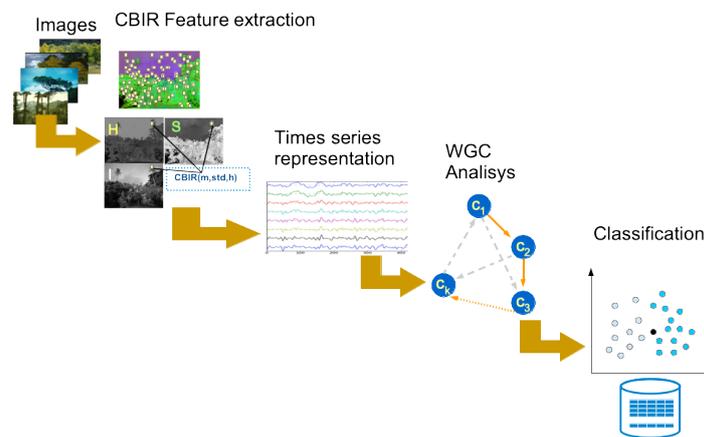
In the computer vision or artificial vision research areas, currently, the aim is to emulate the human visual system in the best way possible. Therefore, the objective pursued by artificial vision is to provide computers with three-dimensional human visual capabilities, generally starting from images in 2D [9].

Since any object or scene cannot be recognized efficiently by a single algorithm, this leaves the door open for computer vision applications where areas such as digital image processing, pattern recognition, machine learning, etc. are combined.

Due to the exponential increase in the natural scenes images in the web, one of the automated image recognition systems tasks is to successfully classify and identify natural scenes images (A natural scene is said to be natural if the image has no human intervention or alteration). It is estimated that more than half of the information on the Internet are images, of which 85% were taken through mobile devices, and a final estimate of five billion images was considered until 2018 [10]. To use images efficiently, a CBIR system is necessary, which would help the users to find relevant images having self-contained features based on our visual perception.

In this work, a natural scene retrieval system is developed by applying the Wiener-Granger Causality theory (WGC) [11] as a tool to analyze images through self-organized information. The causal relationships between the local textures contained in an image were identified, which helps in characterizing a descriptive pattern of a set of natural scenes within an image data set.

The main stages (shown in Figure 1) involved in the developed system are as follows:



**Figure 1.** Proposed general methodology.

1. Image reading: The images in the dataset are read, and then a color space change is applied from Red-Green-Blue (RGB) space to Hue-Saturation-Intensity (HSI) space.
2. Feature Extraction: CBIR statistical feature extraction is randomly generated based on 300 image points.
3. Time series conformation: Texture features are organized as a time series for each image.
4. Causality analysis: WGC analysis is applied to calculate the causal relationship matrix between different textures.
5. Classification application: The  $k - NN$  k-next neighbors classification algorithm is used to find the features closest to or similar to the searched one.

The work proposes a causality analysis of the natural scene classes based on an auto-generated texture dictionary and the WGC analysis of the CBIR features [12,13] to provide the characterization of the data set. It is also possible to search for a type of constitutive element of the scenes, i.e., water, clouds, forests, etc. The proposed system recovers all the scenes that contain a particular element, even when they are classified into different natural scenes.

The proposed methodology was tested on three natural scenes databases: Vogel and Shiele (V\_S) [14], Oliva and Torralba (O\_T) [15], and Shullani (Sh) [16]. The proposed method in this research work can be implemented in an autonomous natural scene recognition system mounted on a car or drone with 100% certainty.

The rest of the paper is organized as follows. In Section 3, the theoretical support of the WGC model to be applied is presented. In Section 4, the methodology is presented. The results are presented in Section 5. Finally, the conclusions and future works are presented in Section 6.

## 2. State of the Art

Since the beginning of the Internet, the need to search and retrieve information has increased. Initially, the information available on Internet was mostly in the form of text [17], but with the technological progress, the information circulating on the Internet has advanced to multimedia data, i.e., text, voice, image, video, and graphics, which require search and retrieval. For each media, there is a need for search and retrieval. In this

research work, search and retrieval of images of natural scenes has been carried out using Wiener-Granger causality.

Nowadays, people need to search for complete scenes or some elements within the scenes, all this within a context of large and complex image bases. Thus, many important search and retrieval methodologies and algorithms have been presented [18]. Since the existing algorithms have been proven limited in performance, the problem of image search and retrieval is still open for researchers.

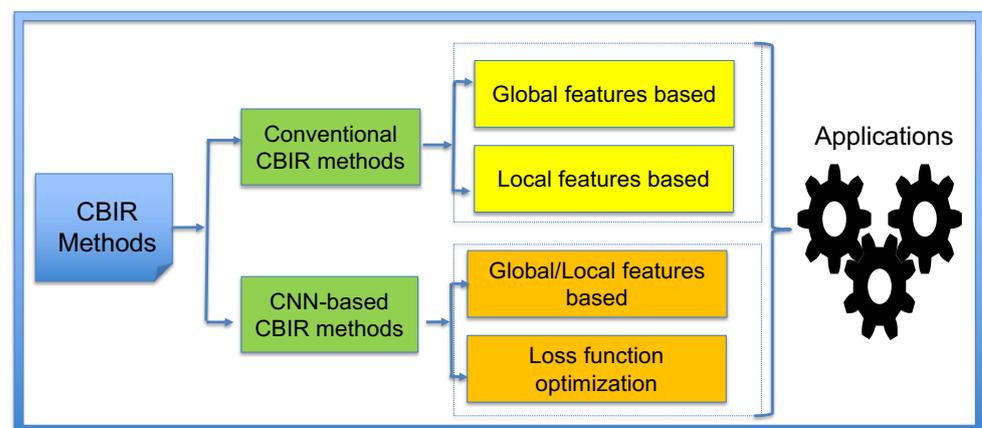
There are different methods of content-based image retrieval, globally, we can classify them into two (See Figure 2):

- *Conventional CBIR methods*: These are the methods based on conventional machine learning algorithms, and based on global and local attributes, or the combination of both [7,12,17–23].
- *CNN-based CBIR methods*: In the recent years, CBIR methods based on neural networks have increased, particularly encoders and convolutional neural networks (CNNs), originating in the use of local and global attributes [24–30].

The classification of CBIR system is shown in Figure 2. Table 1 shows a few CBIR methodologies application fields i.e., medical, biology, academic, design, video/images, Covid-19, among others.

Marinov et al. [31] presented a comparative applied CBIR systems review: AKIWI, FIRE, TINEYE, FRAUNHOFER IOSB, VIRAL, LIRE, YANDEX, PASTEC, and MIFILE. The authors present each of the previous applications, specifying the kind of methodology, image, and descriptor used (local, global, neural).

This research work adopts conventional methodologies using attributes of the Wiener-Granger Causality theory, framed in a self-organizing system of natural settings. The proposed methodology comprises of a stage of feature extraction at random points within the image, after which these features are organized in time-series form, subsequently, the estimation of Wiener-Granger causality is carried out. Once the causal relationships are obtained, a clustering algorithm “k-means” is applied to achieve the self-organization of attributes. Regarding classification, the  $k - NN$  distance classification algorithm is used to find the most similar images that share the causal relationships between the elements of the scenes. Our methodology is validated on three public image databases. The advantage of the proposed system is that it is able to search for a particular element in a scenery image i.e., clouds, forest, water, among others. Thus, the search and recovery of the proposed system maybe by means of a component of a scene.



**Figure 2.** Classification of content-based image retrieval methodologies.

**Table 1.** Examples of some CBIR application fields.

Area	Application/Reference
Medical	<ul style="list-style-type: none"> <li>★ Computer-aided diagnosis of mammography masses based on a supervised content-based image retrieval approach [4].</li> <li>★ Relevance feedback for enhancing content based image retrieval and automatic prediction of semantic image features: Application to bone tumor radiography [3].</li> </ul>
Vegetable	Vegetable Image Retrieval with Fine-tuning VGG Model and Image Hash [32].
COVID-19	Deep metric learning-based image retrieval system for chest radiography and its clinical applications in COVID-19 [33].
Academic/Educational	Design and implementation of CBIR system for academic/educational purposes [5].
Video/Image processing	Comparative analysis on different degrees of jpeg compression used in CBIR systems [34].
Design	Diagram Image Retrieval and Analysis: Challenges and Opportunities [35].
Biology	Indexing and Image Search by the Content according to the Biological Base of the Cognitive Processing of Information using a Neural Sensor [36].
Cloud Repositories	<ul style="list-style-type: none"> <li>★ Practical Privacy-Preserving Content-Based Retrieval in Cloud Image Repositories [37].</li> <li>★ Towards Privacy-Preserving Content-Based Image Retrieval in Cloud Computing [38].</li> </ul>

### 3. Wiener-Granger Causality Analysis Theoretical Fundamentals

The Wiener-Granger theory of causality is used in different areas of knowledge, for example, in neurology the WGC theory is used [39] to examine brain areas and the causal relationships between them. The WGC analysis was performed using sensors [40,41], and also, in magnetic resonance images [42,43], the WGC theory is being used for the study of causal relationships between areas of the brain that perform activities. Other fields of science where the WGC theory has been applied are video processing for massive identification of people and vehicles [44–46] and the analysis of complex scenarios [47]. In this proposal, for the first time, the WGC theory is applied for the recovery of natural elements and scenes.

For the sake of brevity, the theory is presented only for two random processes, being extensible to  $n$  – processes, and to avoid extensive mathematical models. In our approach, a random process corresponds to a signal reading associated with a type of texture within a natural scene. Thus, for the presented analysis, each texture reading corresponds to a stochastic process represented by  $C_i$ ,  $i$  being the  $i$ -th texture that has a stochastic behavior within a scene.

#### Stochastic Autoregressive Model

We consider carrying out analysis with two signals,  $C_1$  and  $C_2$ , being easily extensible to  $n$  signals-textures. Each stationary process represents a texture of the scene and can be represented by an auto-regressive model as follows:

$$C_1(t) = \sum_{k=1}^{\infty} K_{C_1}^1(k) \cdot C_1(t-k) + \eta_{C_1}^1, \quad \text{with} \quad \Sigma_{C_1}^1 = \text{var}(\eta_{C_1}^1), \quad (1)$$

$$C_2(t) = \sum_{k=1}^{\infty} K_{C_2}^1(k) \cdot C_2(t-k) + \eta_{C_2}^1, \quad \text{with} \quad \Sigma_{C_2}^1 = \text{var}(\eta_{C_2}^1), \quad (2)$$

where  $\eta_{C1}^1$  and  $\eta_{C2}^1$  are the Gaussian random noise with zero mean value and unitary standard deviation;  $K_{C1}^1(k)$  and  $K_{C2}^1(k)$  are the coefficients of the regression model to the textures C1 and C2, respectively.

The common auto-regressive model for the two textures is defined by the Equations (3) and (4)

$$C1(t) = \sum_{k=1}^{\infty} K_{C1}^{1,1}(k) \cdot C1(t-k) + \sum_{k=1}^{\infty} K_{C2}^{1,2}(k) \cdot C2(t-k) + \eta_{C1}^2, \quad \text{with } \sum_{C1}^2 = \text{var}(\eta_{C1}^2) \quad (3)$$

$$C2(t) = \sum_{k=1}^{\infty} K_{C1}^{2,1}(k) \cdot C1(t-k) + \sum_{k=1}^{\infty} K_{C2}^{2,2}(k) \cdot C2(t-k) + \eta_{C2}^2, \quad \text{with } \sum_{C2}^2 = \text{var}(\eta_{C2}^2) \quad (4)$$

where  $\sum_{C1}^2$  and  $\sum_{C2}^2$  are the variances of the residual terms  $\eta_{C1}^2$  and  $\eta_{C2}^2$ , respectively.

Now we analyze the variances/covariances residual terms  $\eta_{Ci}^2$  using the equation in matrix form  $\Sigma$  (5):

$$\Sigma = \begin{pmatrix} \sum_{C1}^2 & Y_{1,2} \\ Y_{2,1} & \sum_{C2}^2 \end{pmatrix} \quad (5)$$

where  $Y_{1,2}$  is the covariance between  $\eta_{C1}^2$  and  $\eta_{C2}^2$  (defined as  $Y_{1,2} = \text{cov}(\eta_{C1}^2, \eta_{C2}^2)$ ).

Starting from the previous conditions, and using the concept of statistical independence between two random processes at the same time (in pairs), causality can be defined over time. An example of the causality between C1 and C2 is the expression given by Equation (6).

$$F_{C2,C1} = \ln \left[ \frac{\sum_{C1}^1 \times \sum_{C2}^1}{\sum_{C1}^2 \times \sum_{C2}^2} \right] \quad (6)$$

Equation (6) is commonly known as time domain causality. From this equation, if the random processes  $C1(t)$  and  $C2(t)$  are statistically independent, then  $F_{C1,C2} = 0$ ; otherwise there will be causality from one to the other.

In the Equation (1),  $\sum_{C1}^1$  measures the precision of the auto-regressive model to predict  $C1(t)$ , established in the passed samples. In turn,  $\sum_{C1}^2$  in the expression (3) measures the precision to predict  $C1(t)$  based on the previous values of  $C1(t)$  and  $C2(t)$  at the same time. Returning to the case of taking only 2 textures at the same time  $C1(t)$  and  $C2(t)$  and according to [11], if  $\sum_{C2}^2 < \sum_{C1}^1$  then  $C2(t)$  is said to have a causal influence on  $C1(t)$ . Causality is defined by Equation (7) as follows:

$$F_{C2 \rightarrow C1} = \ln \left[ \frac{\sum_{C1}^1}{\sum_{C1}^2} \right] \quad (7)$$

If  $F_{C2 \rightarrow C1} = 0$  then there is no causal influence of  $C2(t)$  towards  $C1(t)$ , at any other value, the result will be nonzero. Also, the causal influence of  $C1(t)$  towards  $C2(t)$  is established by Equation (8):

$$F_{C1 \rightarrow C2} = \ln \left[ \frac{\sum_{C2}^1}{\sum_{C2}^2} \right] \quad (8)$$

#### 4. Methodology

The training stage in the proposed methodology is presented in Figure 3. It includes the image database reading, the change of color space, the random points selection, the

attributes extraction, the auto-grouping, the time series generation, the causality calculation, the classification-recovery of the  $k$ -next images, and finally, the most similar natural scenes are shown.

The main hypothesis of the proposed methodology is the automatic texture dictionary generation, which represents their own elements contained in the scenes such as water, foliage, clouds, rocks, sand, etc. (see Figure 4). It depends on the self-organization of the information through the  $k$ -means clustering algorithm. Each block in the methodology is described in detail below.

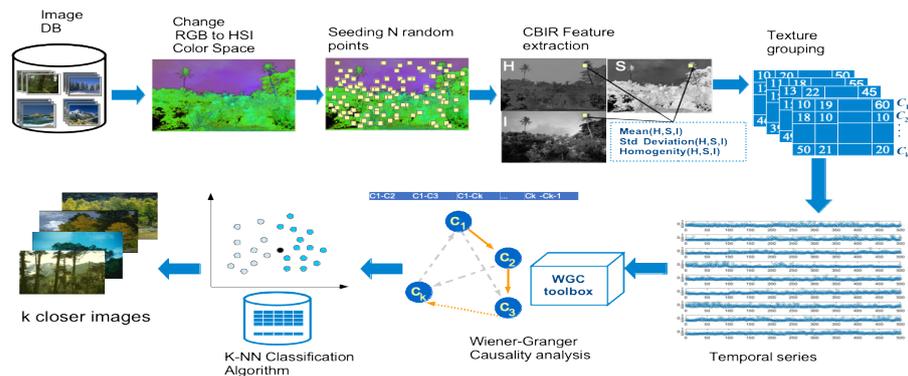


Figure 3. Training and test classification system architecture.

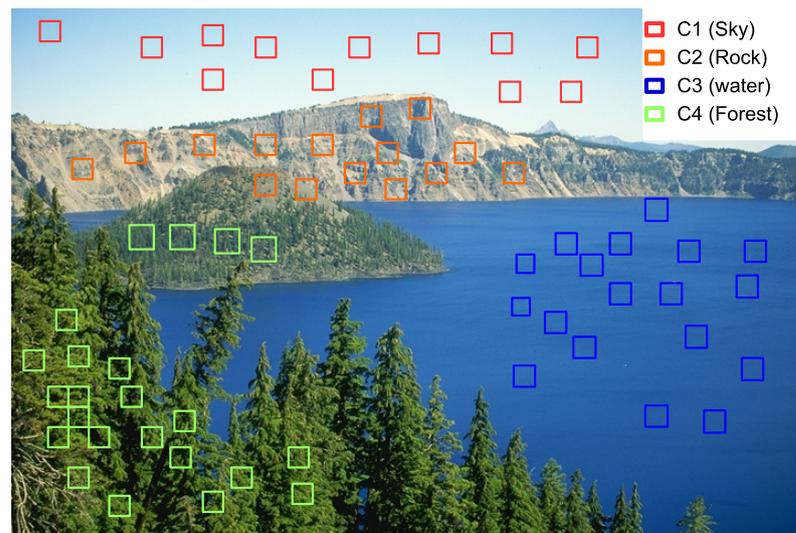


Figure 4. Textured zones example in a natural scene for  $k = 4$  value in the  $k$ -means algorithm.

1. **DB:** Contains the set of natural scenes images.
2. **Preprocessing:** Load the image from the DB, equalize the histogram in the three layers and finally change the RGB color space to the HSI (Hue, Saturation, Intensity) color space, which provides the related textured information.
3. **Random points seeding:** A uniform 300 random points seeding is done inside the image.
4. **Feature extraction.** This step is carried out in two main parts and to have several samples of the image, the process is repeated  $r$  number of times.
  - **Neighborhood generation:** At each random point, a window of size  $p \times p$  pixels is created starting with the interest point in the upper right corner, as shown in Figure 5, where  $p < image\_rows$  and  $p < image\_columns$ .
  - **CBIR feature extraction:** In each neighborhood, three CBIR features are extracted, that are mean, standard deviation, and homogeneity in the three layers of the HSI color space. This appears in a  $1 \times 9$  CBIR features vector, for an image

a matrix  $N_p \times 9$  size is created. Where  $N_p$  is the number of random points and  $F_{CBIR}$  are the extracted CBIR features.

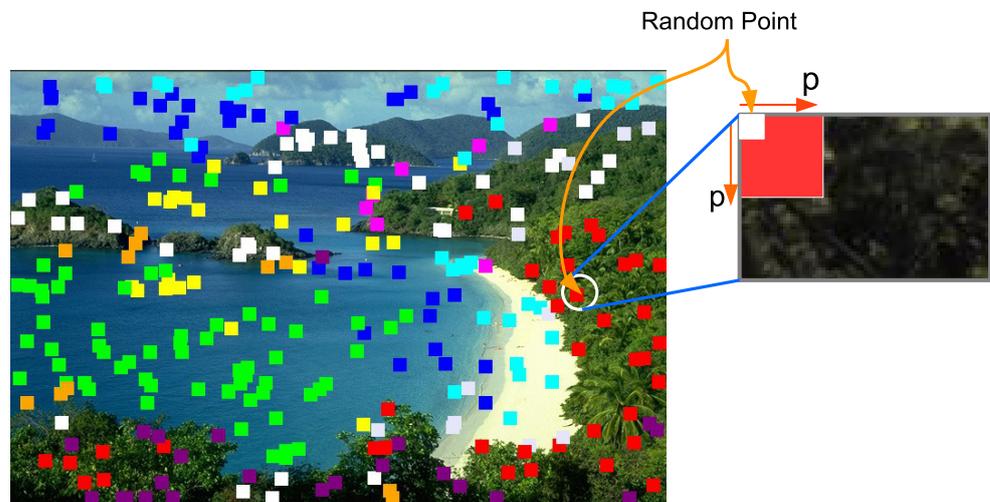


Figure 5. Image with random points, each one generating a neighborhood  $10 \times 10$  size pixels.

5. **Grouping CBIR textures:** Once the CBIR features of each image in the database are extracted, an  $F_{CBIR}$  matrix of size  $(r \times N_I \times N_p) \times 9$  is generated, where  $r$  is the number of repetitions,  $N_I$  is the number of images in the database,  $N_p$  is the number of random points and,  $F_{CBIR}$  the CBIR features at each point for each HSI color layer. By means of the k-means algorithm, the CBIR features are grouped into  $k$  clusters which constitute the  $k$  most representative database textures, generating a  $CM_k$  matrix.
6. **Time series generation:** From the  $F_{CBIR}$  matrix of the previous step, each item of the matrix is compared with the entries of the  $MC_k$  matrix of the automatic dictionary to construct a discrete signal as a time series  $T_S$  having a size of  $k \times r \times N_I$ , where  $k$  is the number of automatic textures of the k-means algorithm,  $r$  is the number of repetitions of experiment, and  $N_I$  is the number of image in the database.
7. **Wiener-Granger Causality analysis:** An entry in the  $T_S$  matrix has a size of  $k \times r$  for each image in the database. This input will be the one that feeds the CWG analysis, as shown in Figure 6. The causality analysis was calculated with the MVGC causality toolbox [48].

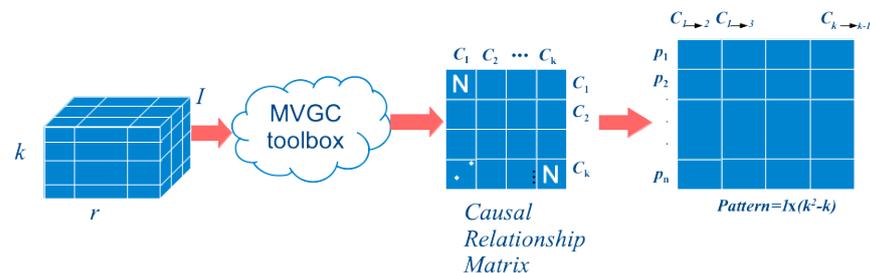


Figure 6. Causal relationships matrix generation  $\Lambda_I$  using CWG analysis.

When causality analysis has been performed for each image in the database, we obtain a causality relationships matrix  $\Lambda_I$  of size  $k \times k$ . The element  $F_{C_i, C_j}$  represents the causal relationships between the scene components  $C_i$  towards  $C_j \forall i, j \in [1...k]$  (see Equation (9)). If  $F_{C_i, C_j} = 0$ , then there is no causal relationship between the image  $i \rightarrow j$ , otherwise, there is a strong causal relationship between them.

$$\Lambda_I = \begin{bmatrix} F_{C_1,C_1} & F_{C_1,C_2} & \cdots & F_{C_1,C_k} \\ F_{C_2,C_1} & F_{C_2,C_2} & \cdots & F_{C_2,C_k} \\ \vdots & \vdots & \vdots & \vdots \\ F_{C_k,C_1} & F_{C_k,C_2} & \cdots & F_{C_k,C_k} \end{bmatrix} \times \frac{1}{ALL_I}, \quad (9)$$

where the constant  $ALL_I$  is defined as  $ALL_I = \sum_{i=1}^k \sum_{j=1}^k F_{C_i,C_j}$

$$\Theta = \cup_{I=1}^{C_s} \Lambda_I = \{\Lambda_1, \Lambda_2 \dots \Lambda_{C_s}\} \quad (10)$$

The causality matrices  $\Lambda_I$  are converted into a vector by concatenating their rows, in the same step the elements of the main diagonal are deleted because there is no causal relationship of a variable with itself. Each element converted into a vector is now an image representative pattern in the database within an array named  $\Theta$ , of size  $N_I \times (k \times k) - k$ , where  $N_I$  is the number of images, and  $k$  the number of automatic textures in the dictionary.

8. Finally, a new  $\Theta$  matrix grouping is carried out using the k-means algorithm to create a set of classes to which each pattern of causality can be seen represented employing an average value. The  $k$  value to create the number of classes is  $k = N_c$ . This is done to obtain  $N_c$  classes within the patterns in the  $\Theta$  array. Therefore, the generated class array has a size of  $N_c \times (r \times r) - r$  elements.

## 5. Experiments and Results

The evaluation of the proposal was carried out using a 19 dual-core processors workstation. The processor used is an Intel copyright Xeon copyright CPU E5-2670 v3 2.30GHz, and 128 GB of RAM.

The current methodology was tested on three natural scenes databases. The metric used was Euclidean.

- Vogel and Shiele (V\_S) [14] comprised of 700 classified scenes: 144 coast, 103 forest, 179 mountain, 131 field, 111 river/lake, and 32 sky/cloud.
- Oliva and Torralba (O\_T) [15] 1472 Images classified as: 360 coast, 328 forest, 374 mountain, and 410 field.
- Shullani et al. (Sh) [16] with 35,000 classified videos and images captured using 35 different portable cellphones. We have selected only 3000 natural scenes: 500 coast, 500 forest, 500 mountain, 500 field, 500 river/lake, and 500 sky/cloud.

### 5.1. Recovery Results

The first two previously mentioned image databases were concatenated, obtaining a single broad image base; the second image database is Shullani (Sh). The natural scenes are: forests, skies, coasts, mountains, fields, and rivers. The results were taken based on two performance evaluation methods: (i) in resubstitution (see Figure 7) and (ii) in cross-validation at 70% and 30% (see Figure 7b). Regarding the number of centers in the k-means algorithm, a value of  $K = 9$  was used, since it provided the best results.

As can be seen in Figure 7a,b, the search or query image is recovered along with the 5 more similar images. The proposed methodology gives 100% result if the the recovery of the most similar image is taken into account in the resubstitution method. In contrast, for the cross-validation method, the five recovered images belong to the same type of natural scene.

Finally, the quantification of the performance of the proposed system is given in Table 2 via confusion matrix showing 100% recovery of each of the natural scene.

	Test k-NN	k=1	k=2	k=3	k=4	k=5
Forest						
Sky						
Coast						
Mountain						
Prairie						
River						

(a) Resubstitution test

	Test k-NN	k=1	k=2	k=3	k=4	k=5
Forest						
Sky						
Coast						
Mountain						
Prairie						
River						

(b) Cross validation test 70/30%

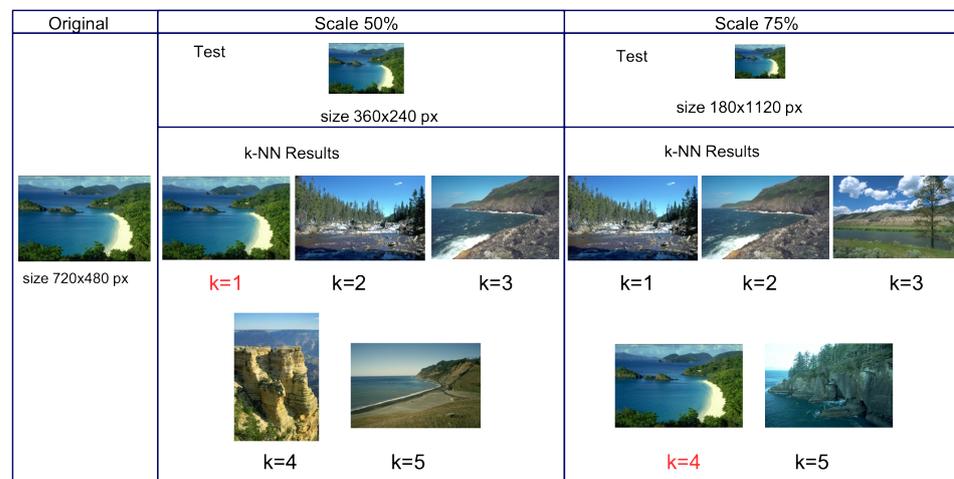
Figure 7. Image retrieval by scene and validation type.

**Table 2.** Confusion matrix using the full image base.

<i>Scene<sub>i</sub>/Scene<sub>j</sub></i>	Forest	Sky	Coast	Mount	Prier	River
Forest	100	0	0	0	0	0
Sky	0	100	0	0	0	0
Coast	0	0	100	0	0	0
Mount	0	0	0	100	0	0
Prier	0	0	0	0	100	0
River	0	0	0	0	0	100

**5.2. Scale Recovery Results**

Figure 8 shows the queries result of an image with a percentage reduction of 50% and 75%. In the confusion matrix, the 100% recovery in the query image is achieved within the 5 closest images given in Table 3.



**Figure 8.** Test with scaled images down to 50% and 75%.

**Table 3.** Confusion matrix for scale, rotation, and noised images (the same results).

<i>Scene<sub>i</sub>/Scene<sub>j</sub></i>	Forest	Sky	Coast	Mount	Prier	River
Forest	100	0	0	0	0	0
Sky	0	100	0	0	0	0
Coast	0	0	100	0	0	0
Mount	0	0	0	100	0	0
Prier	0	0	0	0	100	0
River	0	0	0	0	0	100

**5.3. Rotation Recovery Results**

An important test in this proposal is the rotational invariance of the natural scenes. Since the images can be taken from a drone, light aircraft, helicopter, satellite, etc., the images can be acquired rotated at a given angle. Thus, using the scenes from the proposed databases, In Figure 9, it can be seen that the query image is within the most similar image. Although the five most similar images are sought, the image firstly recovered is the image sought. The confusion matrix result for the rotation tests is given in Table 3.

	Test	k=1	k=2	k=3	k=4	k=5
Rotated by 0°						
Rotated by 90°						
Rotated by 180°						
Rotated by 270°						

Figure 9. Images test rotated by 90, 180, and 270 degrees.

#### 5.4. Noise Recovery Results

The most challenging tests for the proposed methodology are the images having noise, because the CBIR methodology works directly with the texture, the noise directly affects the CBIR performance. Figure 10 shows the recovery results for three types of images contaminated with salt and pepper noise at 0.1%, 0.3%, and 0.5%, the searched image again appears within the 5 most similar images. The result of the confusion matrix for the noise tests is given in Table 3.

	Test	k=1	k=2	k=3	k=4	k=5
Salt & Pepper 0.1%						
Salt & Pepper 0.3%						
Salt & Pepper 0.5%						

Figure 10. Test with images contaminated with salt and pepper noise at 0.1%, 0.3%, and 0.5%.

#### 5.5. Recovery Results for Vision Database

The second set of images of natural scenes (Shullani [16]) taken with cellphone devices was used to test the same methodology as previously defined. Figure 11 shows some natural scenarios such as forest, sky, coast, mountain, prairie, and river. Precision measure was used to determine performance of the proposal, precision measure is defined in Equation (11).

$$P = \frac{TP}{TP + FP} \tag{11}$$

where  $P$  = precision,  $TP$  = True Positive,  $FP$  = False Positive.



Figure 11. Natural scenes examples from Shullani database [16].

As shown in Figures 12 and 13, the search or query image is recovered along with the 5 more similar images by taking into account the recovery of the most similar image in the replacement method. The proposed methodology gives 100% result if the the recovery of the most similar image is taken into account in the resubstitution method. Figure 13 concatenates all the tests carried out on the referenced database such as rotated, scaled, and noised (salt and pepper noise). It can be observed that the proposed method works 100% for the rotated and scaled images, whereas the performance falls to 50% for images having noise.

	VISION DB Test	k-NN k=1	k=2	k=3	k=4	k=5
Forest						
Sky						
Coast						
Mountain						
Prairie						
River						

Figure 12. Image retrieval by scene on Shullani database [16].

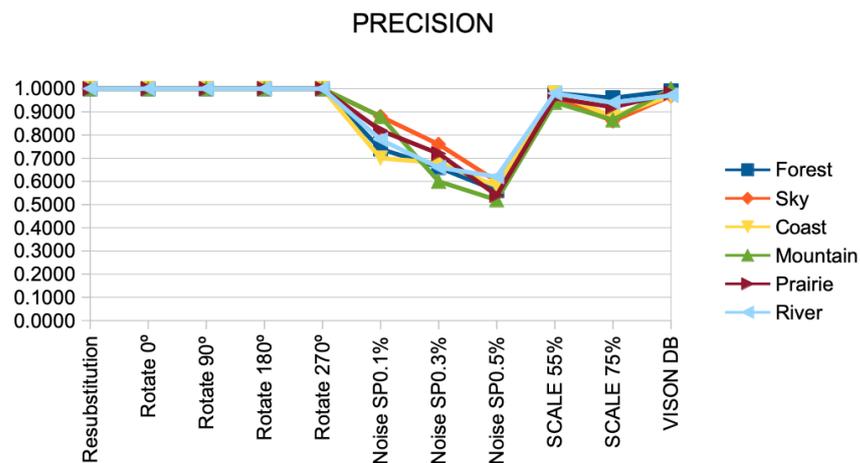


Figure 13. Shullani database [16] performance.

### 5.6. Recovery Results by Element

The proposed methodology also allows the searching of natural scenes by their constituent elements, that is, from among all the scenes in the image base, retrieving all the scenes that contain a particular element of the scene (or common element). Figure 14 shows some examples of natural scenes i.e., fields, rivers, forests, beaches, and mountains. In the lower part of the Figure 14, some semantic concepts that constitute the scenes are shown i.e., clouds, sky, water, grass, mountains, etc.

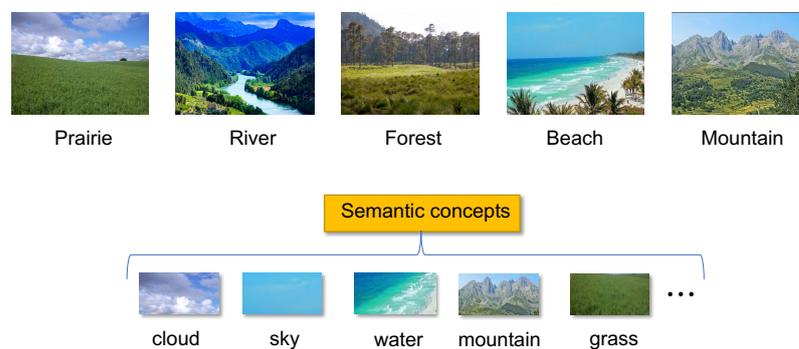


Figure 14. Natural scenes constituent elements.

The protocol experimentation to quantitatively evaluate the scene recovery by element was as follows: one of the six elements to recover in the scenes was defined, each base element being cloud, sky, water, mountain, grass, and river. Subsequently, the recovery of 100 scenes containing the searched item was carried out. It was quantified within the 100 recovered how many, indeed, contained the searched element.

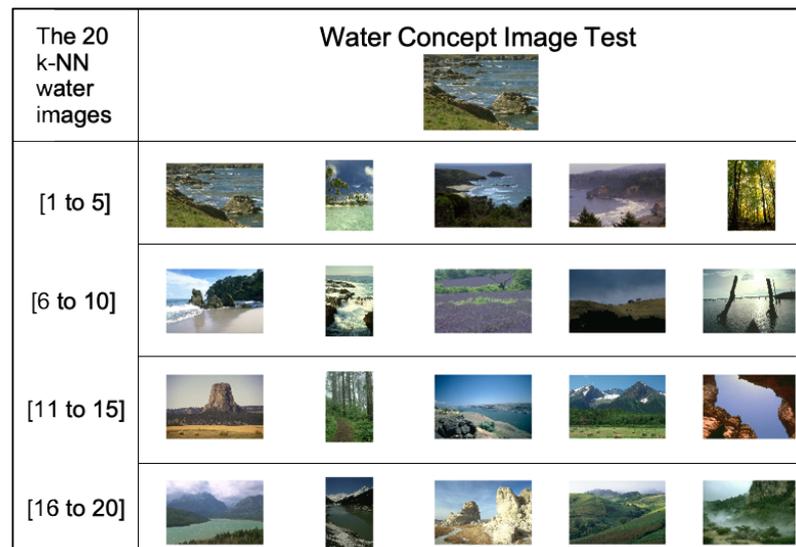
Table 4 gives the confusion matrix for the recovery of the four elements searched in the scenes. It can be seen that a good recovery is obtained since the proposal is ideal for searching one item at a time—for example water, thus having natural scenes that contain water.

- **Water Element:** Figure 15 shows an example of recovering natural scenes containing the element *water*. For the proposed methodology, the type of scene is not important, and only items containing water (sea water or river water) are recovered. Among all the 20 images, only 14 images containing water were recovered.
- **Cloud Element:** Regarding the *cloud* element, Figure 16 shows the recovery results, for the sake of clarity, first 20 mages recovered are presented. It can be seen from the example that 15 out of 20 images contained *cloud*.

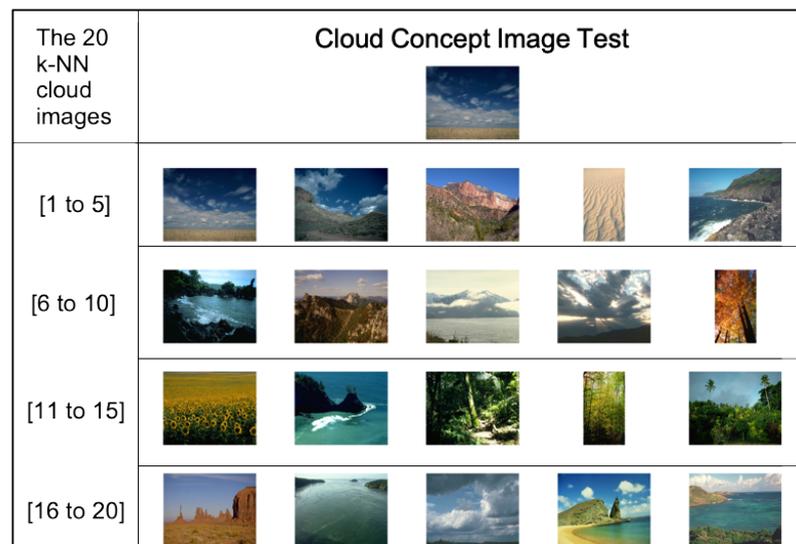
- Follage element: To recover the *follage* element, Figure 17 shows 15 out of 20 scenes recovered.
- Rock Element: Finally, for the *rock* element, Figure 18 provides an example of recovery of 20 scenes recovered; from the figure, it can be seen that 17 out of 20 images.

**Table 4.** Confusion matrix for elements constitutive of natural scenes.

<i>Scene<sub>i</sub>/Scene<sub>j</sub></i>	Clud	Follage	Water	Rock
Clud	75	0	0	0
Follage	0	75	0	0
water	0	0	70	0
Rock	0	0	0	85



**Figure 15.** Retrieval results related to the concept of water.



**Figure 16.** Retrieval results related to the concept of cloud.

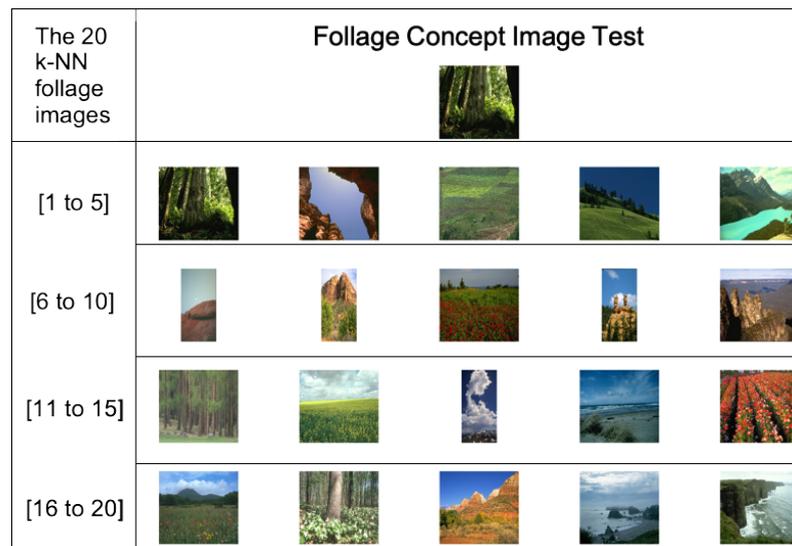


Figure 17. Retrieval results related to the concept of foliage.

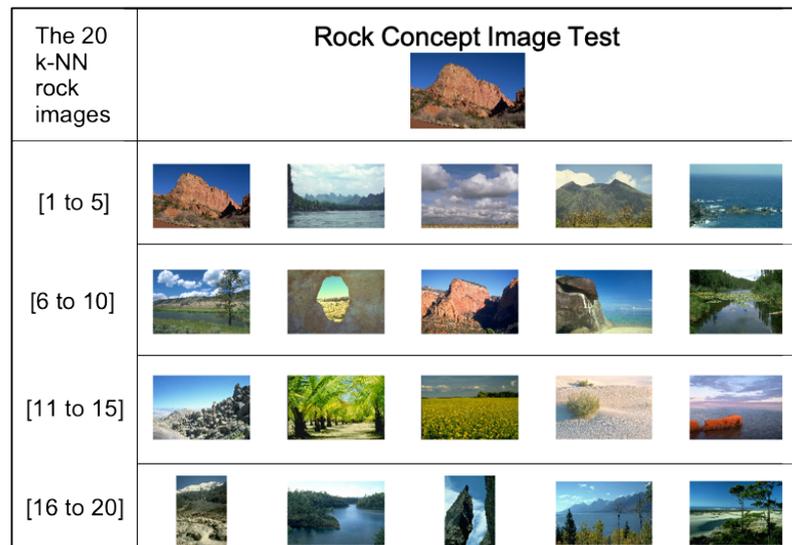


Figure 18. Retrieval results related to the concept of rock.

### 5.7. Comparison with Related Works

The comparison between the proposed method in this research paper and other competitive methods reported in the literature is given in Table 5. The classification performance for the six natural scenes, as well as the mean average performance, are presented. The comparison also includes the best competitive method using Convolutional Neural Networks [12,28–30]. It can be seen that our proposal accurately recognizes (100%) each of the natural images under consideration. As shown in Table 5, using the same database found in other research works, our proposed method shows 100% recognition.

Table 5. Detection performance of the most competitive methods in natural image recognition using the same databases Vogel-Schiele.

	Coast	Forest	Mountain	Follage	Rever/Lake	Sky/Cloud	Mean Average
<b>Our Proposal</b>	<b>100%</b>						
Damodaran [28]	86%	92%	94%	86%	86%	86%	88.33%
Sharma [29]	84.64%	84.64%	84.64%	84.64%	84.64%	84.64%	84.64%
Damodaran [30]	93%	91%	98%	88%	88%	88%	91%
Serrano-Talamantes [12]	96%	100%	88%	88%	96%	86%	92.33%

## 6. Conclusions and Future Work

This research paper proposes the utilization of the Wiener-Granger causality theory, together with the CBIR self-organization analysis. The novel proposal is applied to image retrieval of 6 natural scenes that are: coast, forest, mountain, field, river/lake, and sky/cloud. Taking into account the proposed methodology, from the causality matrix, it was fruitful to find a set of descriptors that represent a type of natural scene. Texture patterns could be defined from an automatic set of reference textures. From the self-organizing attributes, it is now possible to classify any unknown natural scene.

With the proposed methodology, 100% image retrieval is achieved for the three data sets. The proposal is advantageous since no prior labeling or knowledge of the natural scenes content is required.

The proposed methodology gives 100% of recovery on vision image database. Thus, the proposal works at 100% for rotated images and scaled images; performance falls down to 50% for noised images.

Another important contribution of our proposal is the ability to recover natural scenes by some contained element, i.e., water, cloud, foliage, and rock. The percentage of recovery of these natural elements is above 70% using the proposal presented in this research paper.

The experimental results show that our proposal outperforms the most competitive methods reported by Damodaran [28], Sharma [29], Damodaran [30], Serrano-Talamantes [12]; with an average recognition of 100% for the same image datasets.

Future work will seek the entire methodology implementation in parallel computing; using CPUs and GPU technology, which could perform the scene recovery task efficiently in the image feature extraction stage. The parallel algorithms might also help to jointly analyze the textures of the image seeking to characterize the image and its associations with the paradigm of visual understanding.

**Author Contributions:** Writing—review and editing, C.A.-C., C.B.-A.; investigation, C.B.-A., C.A.-C. and E.R.-M.; resources, A.F.-R.; writing—original draft preparation, C.A.-C. and C.B.-A.; validation, C.B.-A.; conceptualization, C.A.-C. and C.B.-A.; formal analysis, C.B.-A., E.R.-M. and A.Z.-L.; methodology, C.B.-A. and C.A.-C.; C.A.-C. supervised the overall research work. All authors contributed to the discussion and conclusion of this research. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

1. Sampath, V.; Murtua, I.; Aguilar Martín, J.; Gutierrez, A. A survey on generative adversarial networks for imbalance problems in computer vision tasks. *J. Big Data* **2021**, *8*, 1–59. [[CrossRef](#)]
2. Zafar, B.; Ashraf, R.; Ali, N.; Iqbal, M.K.; Sajid, M.; Dar, S.H.; Ratyal, N.I. A Novel Discriminating and Relative Global Spatial Image Representation with Applications in CBIR. *Appl. Sci.* **2018**, *8*, 2242. [[CrossRef](#)]
3. Banerjee, I.; Kurtz, C.; Devorah, A.E.; Do, B.; Rubin, D.L.; Beaulieu, C.F. Relevance feedback for enhancing content based image retrieval and automatic prediction of semantic image features: Application to bone tumor radiographs. *J. Biomed. Inform.* **2018**, *84*, 123–135. [[CrossRef](#)]
4. Tsochatzidis, L.; Zagoris, K.; Arikidis, N.; Karahaliou, A.; Costaridou, L.; Pratikakis, I. Computer-aided diagnosis of mammographic masses based on a supervised content-based image retrieval approach. *Pattern Recognit.* **2017**, *71*, 106–117. [[CrossRef](#)]
5. Marinov, M.; Valova, I.; Kalmukov, Y. Design and implementation of CBIR system for academic/educational purposes. In Proceedings of the 2020 International Conference Automatics and Informatics (ICAI), Varna, Bulgaria, 1–3 October 2020; pp. 1–4. [[CrossRef](#)]
6. Jiang, D.; Kim, J. Image Retrieval Method Based on Image Feature Fusion and Discrete Cosine Transform. *Appl. Sci.* **2021**, *11*, 5701. [[CrossRef](#)]
7. Li, X.; Yang, J.; Ma, J. Recent developments of content-based image retrieval (CBIR). *Neurocomputing* **2021**. [[CrossRef](#)]

8. Jena, B.; Nayak, G.; Saxena, S. Survey and Analysis of Content-Based Image Retrieval Systems. *Lect. Notes Electr. Eng.* **2021**, *710*, 427–433.
9. Ansari, M.; Singh, D. Human detection techniques for real time surveillance: A comprehensive survey. *Multimed. Tools Appl.* **2021**, *80*, 8759–8808. [[CrossRef](#)]
10. Tyagi, V. *Content-Based Image Retrieval-Ideas, Influences, and Current Trends*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 1–378.
11. AGranger, C.W.J. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* **1969**, *37*, 424–438. [[CrossRef](#)]
12. Serrano-Talamantes, J.; Avilés-Cruz, C.; Villegas-Cortez, J.; Sossa-Azuela, J. Self organizing natural scene image retrieval. *Expert Syst. Appl.* **2013**, *40*, 2398–2409. [[CrossRef](#)]
13. Villegas Cortez, J.; Benavides-Alvarez, C.; Román-Alonso, G.; Cruz, C. Reconocimiento de rostros a partir de la propia imagen usando técnica CBIR. In Proceedings of the X Congreso Español sobre Metaheurísticas, Algoritmos Evolutivos y Bioinspirados (MAEB 2015), Merida Extremadura, Spain, 4–6 February 2015; [[CrossRef](#)]
14. Vogel, J.; Schiele, B. Performance evaluation and optimization for content-based image retrieval. *Pattern Recognit.* **2006**, *39*, 897–909. [[CrossRef](#)]
15. Oliva, A.; Torralba, A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *Int. J. Comput. Vis.* **2001**, *42*, 145–175. [[CrossRef](#)]
16. Shullani, D.; Fontani, M.; Iuliani, M.; Shaya, O.A.; Piva, A. VISION: A video and image dataset for source identification. *EURASIP J. Inf. Secur.* **2017**, *2017*, 15. [[CrossRef](#)]
17. Li, Y.; Ma, J.; Zhang, Y. Image retrieval from remote sensing big data: A survey. *Inf. Fusion* **2021**, *67*, 94–115. [[CrossRef](#)]
18. Traina, A.J.; Brinis, S.; Pedrosa, G.V.; Avalhais, L.P.; Traina, C. Querying on large and complex databases by content: Challenges on variety and veracity regarding real applications. *Inf. Syst.* **2019**, *86*, 10–27. [[CrossRef](#)]
19. Salazar, A.; Igual, J.; Safont, G.; Vergara, L.; Vidal, A. Image Applications of Agglomerative Clustering Using Mixtures of Non-Gaussian Distributions. In Proceedings of the 2015 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 7–9 December 2015; pp. 459–463. [[CrossRef](#)]
20. Irtaza, A.; Adnan, S.M.; Ahmed, K.T.; Jaffar, A.; Khan, A.; Javed, A.; Mahmood, M.T. An Ensemble Based Evolutionary Approach to the Class Imbalance Problem with Applications in CBIR. *Appl. Sci.* **2018**, *8*, 495. [[CrossRef](#)]
21. Rehman Malik, N.U.; Airij, A.G.; Memon, S.A.; Panhwar, Y.N.; Abu-Bakar, S.A.; El-Khoreby, M.A. Performance Comparison Between SURF and SIFT for Content-Based Image Retrieval. In Proceedings of the 2019 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia, 17–19 September 2019; pp. 214–218. [[CrossRef](#)]
22. Feng, Q.; Wei, Y.; Yi, Y.; Hao, Q.; Dai, J. Local Ternary Cross Structure Pattern: A Color LBP Feature Extraction with Applications in CBIR. *Appl. Sci.* **2019**, *9*, 2211. [[CrossRef](#)]
23. Paolanti, M.; Frontoni, E. Multidisciplinary Pattern Recognition applications: A review. *Comput. Sci. Rev.* **2020**, *37*, 100276. [[CrossRef](#)]
24. Liu, G.H.; Yang, J.Y. Deep-seated features histogram: A novel image retrieval method. *Pattern Recognit.* **2021**, *116*, 107926. [[CrossRef](#)]
25. Hassan, A.; Liu, F.; Wang, F.; Wang, Y. Secure content based image retrieval for mobile users with deep neural networks in the cloud. *J. Syst. Archit.* **2021**, *116*, 102043. [[CrossRef](#)]
26. Gkelios, S.; Sophokleous, A.; Plakias, S.; Boutalis, Y.; Chatzichristofis, S.A. Deep convolutional features for image retrieval. *Expert Syst. Appl.* **2021**, *177*, 114940. doi: 10.1016/j.eswa.2021.114940. [[CrossRef](#)]
27. Pradhan, J.; Pal, A.K.; Banka, H.; Dansena, P. Fusion of region based extracted features for instance- and class-based CBIR applications. *Appl. Soft Comput.* **2021**, *102*, 107063. [[CrossRef](#)]
28. Damodaran, N.; Sowmya, V.; Govind, D.; Soman, K. Single-plane scene classification using deep convolution features. *Adv. Intell. Syst. Comput.* **2019**, *900*, 743–752. [[CrossRef](#)]
29. Sharma, K.; Gupta, S.; Dileep, A.; Rameshan, R. Scene Image Classification Using Reduced Virtual Feature Representation in Sparse Framework. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 2701–2705. [[CrossRef](#)]
30. Damodaran, N.; Sowmya, V.; Govind, D.; Soman, K. Effect of decolorized images in scene classification using deep convolution features. *Procedia Comput. Sci.* **2018**, *143*, 954–961. [[CrossRef](#)]
31. Marinov, M.; Valova, I.; Kalmukov, Y. Comparative Analysis of Content-Based Image Retrieval Systems. In Proceedings of the 2019 16th Conference on Electrical Machines, Drives and Power Systems (ELMA), Varna, Bulgaria, 6–8 June 2019; pp. 1–5. [[CrossRef](#)]
32. Yang, Z.; Yue, J.; Li, Z.; Zhu, L. Vegetable Image Retrieval with Fine-tuning VGG Model and Image Hash. *IFAC-PapersOnLine* **2018**, *51*, 280–285. [[CrossRef](#)]
33. Zhong, A.; Li, X.; Wu, D.; Ren, H.; Kim, K.; Kim, Y.; Buch, V.; Neumark, N.; Bizzo, B.; Tak, W.Y.; et al. Deep metric learning-based image retrieval system for chest radiograph and its clinical applications in COVID-19. *Med. Image Anal.* **2021**, *70*, 101993. [[CrossRef](#)] [[PubMed](#)]
34. Marinov, M. Comparative Analysis on Different Degrees of JPEG Compression Used in CBIR Systems. In Proceedings of the 2020 XI National Conference with International Participation (ELECTRONICA), Sofia, Bulgaria, 23–24 July 2020; pp. 1–4. [[CrossRef](#)]

35. Yang, L.; Gong, M.; Asari, V.K. Diagram Image Retrieval and Analysis: Challenges and Opportunities. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 14–19 June 2020; pp. 685–698. [[CrossRef](#)]
36. Kouddad, F.Z.; Kohili, M.; Lamari, A.C.; Amiri, A. Indexing and Image Search by the Content According to the Biological Base of the Cognitive Processing of Information using a Neural Sensor. In Proceedings of the 2020 2nd International Conference on Mathematics and Information Technology (ICMIT), Adrar, Algeria, 18–19 February 2020; pp. 169–174. [[CrossRef](#)]
37. Ferreira, B.; Rodrigues, J.; Leitão, J.; Domingos, H. Practical Privacy-Preserving Content-Based Retrieval in Cloud Image Repositories. *IEEE Trans. Cloud Comput.* **2019**, *7*, 784–798. [[CrossRef](#)]
38. Xia, Z.; Zhu, Y.; Sun, X.; Qin, Z.; Ren, K. Towards Privacy-Preserving Content-Based Image Retrieval in Cloud Computing. *IEEE Trans. Cloud Comput.* **2018**, *6*, 276–286. [[CrossRef](#)]
39. Bressler, S.; Seth, A. Wiener-Granger Causality: A well established methodology. *NeuroImage* **2011**, *58*, 323–329. [[CrossRef](#)] [[PubMed](#)]
40. Matias, F.; Gollo, L.; Carelli, P.; Bressler, S.; Copelli, M.; Mirasso, C. Modeling positive Granger causality and negative phase lag between cortical areas. *NeuroImage* **2014**, *99*, 411–418. [[CrossRef](#)]
41. Mannino, M.; Bressler, S. Foundational perspectives on causality in large-scale brain networks. *Phys. Life Rev.* **2015**, *15*, 107–123. [[CrossRef](#)] [[PubMed](#)]
42. Wang, M.Y.; Yuan, Z. EEG Decoding of Dynamic Facial Expressions of Emotion: Evidence from SSVEP and Causal Cortical Network Dynamics. *Neuroscience* **2021**, *459*, 50–58. [[CrossRef](#)] [[PubMed](#)]
43. DSouza, A.; Abidin, A.; Leistriz, L.; Wismüller, A. Exploring connectivity with large-scale Granger causality on resting-state functional MRI. *J. Neurosci. Methods* **2017**, *287*, 68–79. [[CrossRef](#)]
44. Kular, D.; Ribeiro, E. Analyzing activities in videos using latent Dirichlet allocation and granger causality. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Cham, Switzerland, 2015; Volume 9474, pp. 647–656.
45. Huang, S.N.; Huang, D.J.; Khuhro, M. High-level codewords based on granger causality for video event detection. *Adv. Multimed.* **2015**, *2015*. [[CrossRef](#)]
46. Zhang, C.; Yang, X.; Lin, W.; Zhu, J. Recognizing Human Group Behaviors with Multi-group Causalities. In Proceedings of the 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, Macau, China, 4–7 December 2012; pp. 44–48. [[CrossRef](#)]
47. Fan, Y.; Yang, H.; Zheng, S.; Su, H.; Wu, S. Video sensor-based complex scene analysis with Granger causality. *Sensors (Switzerland)* **2013**, *13*, 13685–13707. [[CrossRef](#)] [[PubMed](#)]
48. Barnett, L.; Seth, A. The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *J. Neurosci. Methods* **2014**, *223*, 50–68. [[CrossRef](#)] [[PubMed](#)]