Applied Data Science and Analysis
Vol.2024, **pp**. 17–31
DOI: https://doi.org/10.58496/ADSA/2024/003; ISSN: 3005-317X
https://mesopotamian.press/journals/index.php/ADSA

Research Article

Semantic Image Retrieval Analysis Based on Deep Learning and Singular Value Decomposition

M.H. Hadid[1, *], Z.T. Al-Qaysi [1], Qasim Mohammed Hussein[1], Rasha A. Aljanabi[1], Israa Rafaa Abdulqader[1], M. S Suzani[2], WL Shir[2]

[1] Tikrit University, College of Computer Science and Mathematics, Computer Science Department, Iraq

[2] University Pendidikan Sultan Idris, faculty of computing and meta-technology, Tanjong Malim, Perak, Malaysia

ARTICLEINFO

Article History

Received 02 Jan 2024
Accepted 04 Mar 2024
Published 25 Mar 2024

Keywords

Image retrieval

Content-Based Image Retrieval

Transfer learning

Deep Learning

ABSTRACT

The exponential growth in the total quantity of digital images has necessitated the development of systems that are capable of retrieving these images. Content-based image retrieval is a technique used to get images from a database. The user provides a query image, and the system retrieves those photos from the database that are most similar to the query image. The image retrieval problem pertains to the task of locating digital photographs inside extensive datasets. Image retrieval researchers are transitioning from the use of keywords to the utilization of low-level characteristics and semantic features. The push for semantic features arises from the issue of subjective and time-consuming keywords, as well as the limitation of low-level characteristics in capturing high-level concepts that users have in mind. The main goal of this study is to examine how convolutional neural networks can be used to acquire advanced visual features. These high-level feature descriptors have the potential to be the most effective compared to the handcrafted feature descriptors in terms of image representation, which would result in improved image retrieval performance. The (CBIR-VGGSVD) model is an ideal solution for content-based image retrieval that is based on the VGG-16 algorithm and uses the Singular Value Decomposition (SVD) technique. The suggested model incorporates the VGG-16 model for the purpose of extracting features from both the query images and the images kept in the database. Afterwards, the dimensionality of the features retrieved from the VGG-16 model is reduced using SVD. Then, we compare the query photographs to the dataset images using the cosine metric to see how similar they are. When all is said and done, images that share a high degree of similarity will be successfully extracted from the dataset. A validation of the retrieval performance of the CBIR-VGGSVD model is performed using the Corel-1K dataset. When the VGG-16 standard model is the sole one used, the implementation will produce an average precision of 0.864. On the other hand, when the CBIR-VGGSVD model is utilized, this average precision is revealed to be (0.948). The findings of the retrieval ensured that the CBIR-VGGSVD model provided an improvement in performance on the test pictures that were utilized, surpassing the performance of the most recent approaches.

1. INTRODUCTION

In recent times, a significant proportion of the global population use diverse social media platforms to express their emotions through visual content. Due to technological advancements, the internet is experiencing a significant influx of digital photographs, leading to a daily increase in the size of digital repositories. Finding a certain type of image in such a massive archive, However, it is like to searching for a small object in a large and chaotic collection. Image retrieval has become a highly difficult problem in today's era due to the immense number of photographs available on the internet [1, 2]. Image Retrieval (IR) refers to the process of searching a database to find photos that are more similar to the query being

*Corresponding author. Email: marwa.h.hadid@tu.edu.iq

made [3, 4]. Text-based image retrieval (TBIR) and content-based image retrieval (CBIR) are the two fundamental classifications of image retrieval [5]. Text-based image retrieval (TBIR) systems necessitate specific information like image tags, image name, image path, image location, keywords, and any other pertinent descriptive details pertaining to the image. Annotating huge datasets manually is the most laborious and challenging aspect of this procedure. This approach has two primary constraints. First, describing an image with a few terms is wasteful since it does not represent the entire visual information of the image. Furthermore, inaccurate human visual perception that arises during the annotation process of datasets can lead to incorrect results [6]. The downsides of this technique are as follows: 1) In TBIR, humans must personally characterize each image in the database, therefore if there are a huge number of photographs, this technique demands too much manual effort. 2) TBIR methods require a significant number of metadata to ensure that the search results are relevant and the output results do not contain an excessive number of records. 3) The description of an image's content is a person's subjective perception, which means that various people can describe the same image's content differently. 4) Because queries rely mostly on text information, their execution is heavily influenced by the degree of correlation between images and their text descriptions. This issue was resolved by employing content-based picture retrieval, which enabled the acquisition of more precise image similarity results while maintaining a high level of search performance [7, 8]. CBIR is a system that pulls many images from a vast database based on a query image, employing a range of image attributes such as texture, color, and shape information [9, 10]. The latter methodology is constructed upon the sequential execution of feature extraction and distance measurement. The first stage involves feature extraction, also known as an image representation, and the second as a similarity search, which compares the query image's representation to all of the images in the database [11, 12]. CBIR primarily involves feature extraction, wherein the significant characteristics of an image are identified and subsequently transformed into a fixed-size vector known as a feature vector [13]. The objective of CBIR is to address the problem of accessing and locating images of user interest from extensive image collections, using a query image as a basis [14]. Some examples of users of criminal background investigation (CBIR) systems include design engineers who need to find similar design projects for their clients or law enforcement agencies that want to compare a suspect's face to a database of notorious criminals. Furthermore, the commercial department must verify that no identical or confusingly similar trademark has ever been used before approving its use. In the final instance, certain hospital conditions may necessitate the search for and evaluation of X-rays in order to be diagnosed [15, 16] Prior to providing a remedy, it is necessary to obtain either comparable photographs or scanned images of the patient [17, 18]. Due to the rapid expansion of deep learning, features obtained from pre-trained CNN models have demonstrated superior adaptability and effectiveness compared to conventional Descriptors are used in tasks related to retrieving images. This feature information consists of a variety of images that contain different semantic information. This is essential for improving the precision of images retrieving [19, 20]. The primary goal of the approach that has been developed is to include CNN's precision into the process of improving the performance of image retrieval.

2. RELATED WORKS

This section provides a description of the past studies on Content-Based Image Retrieval (CBIR) that have been presented in the literature. These works are focused on retrieving images that are connected to the query image.

In (2018), Göksu and Aptoula [21] They suggested employing deep features obtained from a fine-tuned convolutional neural network model called CaffeNet. Dimensionality reduction methods (PCA and KPCA) were used to these deep features immediately following the feature extraction process. Using both the UCM Merced and RSSCN7 datasets, they assessed the efficacy of this approach. MAP on the UCM dataset yielded a value of 0.534 when PCA was applied, and 0.543 when KPCA was utilized. MAP on the RSSCN7 dataset produced a value of 0.535 when PCA was applied, and 0.522 when KPCA was utilized.

In (2019), Rian and Hendryli [22] They suggested augmenting the CBIR system through the implementation of a deep learning technique utilizing convolutional neural networks (VGG16) and an iNaturalist dataset for training purposes. The cosine similarity was used as a metric to measure the consistency between the photos. The retrieval method attained an average precision of 89.6% and an accuracy, measured by the F1-score, of 73%.

In (2020), Shraf et al. [23] They suggested texture features by combining the Gabor Wavelet Descriptor and Discrete Wavelet Transform (DWT), and they extracted color features using Color Moments for the HSV color space. Furthermore, characteristics were extracted by employing the color and edge directivity descriptor (CEDD) in conjunction with a Euclidean distance-based similarity measure. The mean precision score achieved for extracting 20 photos from the Corel-1k dataset was 0.875.

In (2021), Devulapalli et al. [24] In addition to devising a technique for retrieving images, they constructed the feature vector through the integration of handcrafted and deep learning features. To decrease the amount of feature dimensions, deep learning models used the pre-trained GoogleNet model as a feature extractor. The Gabor filter was employed to extract the manually created features. The two features were merged, and the resulting feature vectors were employed. The

Euclidean distance was computed between the query image and the images in the database. The precision score obtained was 0.91.

In (2022), Kalaivani et al. [25] A hyperparameter-tuned DL model for CBIR (HPTDLCBIR) was proposed, which employed a VGG-16 model as the feature extractor. Simultaneously, the water wave optimization (WWO) techniques were employed to precisely determine the hyperparameters of the VGG-16 model. This distance was measured using the Manhattan method. After obtaining 20 photos from the Corel-1k dataset, the average precision result was 0.86.

In (2022), Karthik et al. [26] An evolutionary optimization algorithm was suggested for implementation on the CBIR system, which incorporates squeeze networks and handcrafted features (EOCBIR-HFSN). Their approach involved utilizing` manually produced features derived from local binary patterns, in conjunction with deep features extracted from SqueezeNet. The SqueezeNet model's hyperparameters were fine-tuned with the help of the grasshopper optimization (GOA) technique. After that, we moved on to the Euclidean distance metric. The COREL-10K dataset was used to build this approach. A recall of 79.20 and a precision of 92.00 were the results.

3. CONTENT-BASED IMAGE RETRIEVAL

CBIR, or content-based image retrieval, is a significant issue pertinent to human activity in a number of scientific and commercial domains [27]. Another name for CBIR is query by image content [3]. CBIR, or Content-Based Image Retrieval, is a system that enables users to search for certain photos within a given collection by utilizing query images. CBIR, or Content-Based Image Retrieval, distinguishes itself from earlier image retrieval methods by utilizing a picture as input instead of relying on tags or text [28]. An important part of the CBIR system is the semantic gap, which shows how different low-level picture pixels are from high-level semantic ideas. By prioritizing the significant characteristics in the image, they provide a precise depiction of the image's contents, encompassing the color, texture, and shape [29]. Additionally, the CBIR system has the capability to record complex semantic ideas related to emotions, abstraction, and perception, which may be easily comprehended by people. This refers to a depiction of an image from a worldwide viewpoint, and it pertains to the definition of an image, where these characteristics can effectively imitate the human visual system [30]. Research articles pertaining to image retrieval primarily center on the semantic chasm. There are two main types of image retrieval methods: Personalized and expertly trained [31-33]. Typically, There will be two parts to the CBIR system: the offline and online parts[13]. In the offline stage, features are extracted from massive datasets that will later be used for system training and feature generation. The amount of training photos utilized to train the system determines how long this phase usually takes. During the online phase, the query image's features are extracted and the distance is calculated between the query image's features and the features of the dataset. This is done to measure the similarity between the images. Following this, the retrieval results for images with a small distance between them or a high degree of similarity are displayed to the user [13, 34]. Figure 1 illustrates a standard CBIR system.

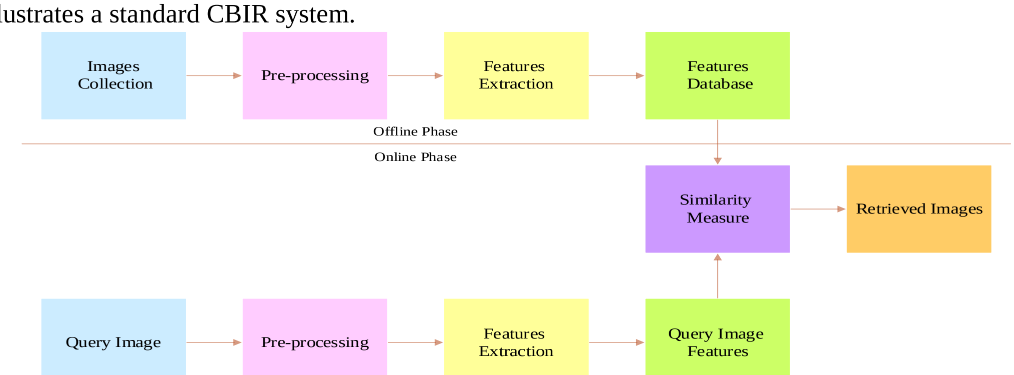


Fig 1. CBIR system

4. VISUAL GEOMETRY GROUP

Visual Geometry Group (VGG) [35, 36] refers to the team of mathematicians and artificial intelligence experts from Oxford University in England who are taking part in the ILSVRC competition. They established a collection of networks bearing the name of their organization in 2014. Their network was constructed as follows: two sets of three-by-three convolution layers, one set of convolution layers with a max pool, and another pair of three-by-three convolution layers with a max pool [37]. Following this, A max pool was added after three further sets of (3 × 3) convolution layers; three sets of (3 × 3) convolution layers were added after that, followed by a max pool; and three sets of (3 × 3) convolution layers were added after that [37, 38]. Furthermore, Two main, fully-connected layers made up the output layers; each layer had 4096 neurons. Ultimately, they obtained a 1000-neuron output layer that corresponds to each ImageNet class [39, 40]. The network was named VGG-16 because it had an input layer with a total of 2+2+3+3+3 units, followed by 2 completely connected layers, and finally an output layer [39, 41]. Figure 2 displays the VGG-16 architecture.

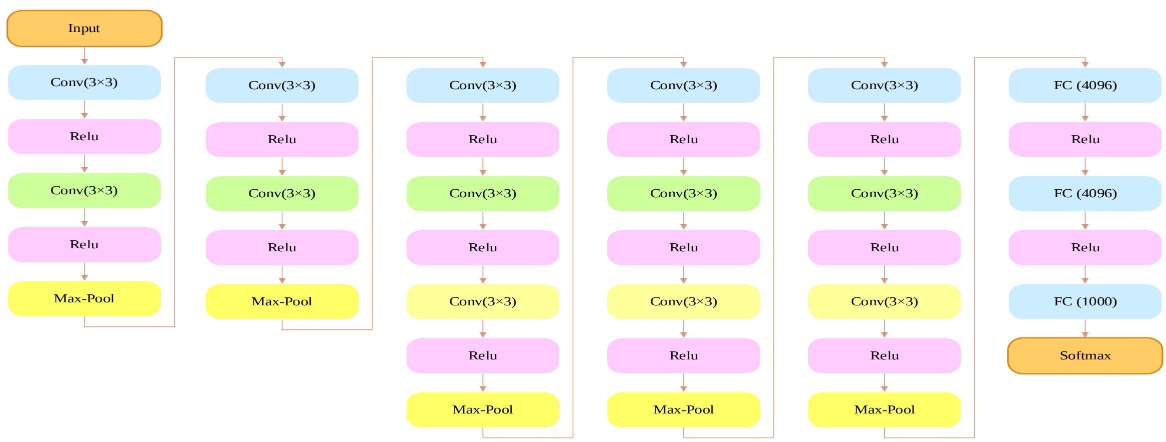


Fig 2. VGG-16

5. DIMENSIONALITY REDUCTION

Reducing the number of dimensions needed to describe data from a higher-dimensional feature to a lower-dimensional feature is called dimension reduction [42].

5.1 Singular Value Decomposition (SVD)

The quantity of storage space needed for digital photos might be substantial; therefore, it would be optimal if this requirement could be reduced without sacrificing image quality [43]. Massive datasets are modified with SVD in order to identify image features that contribute the least to the overall quality of the image [43]. By employing the Singular Value Decomposition (SVD) technique, it is possible to decrease the size of images while retaining crucial components. This enables images to be saved using a less amount of memory without compromising their quality [43]. The acronym SVD stands for Singular Value Decomposition, which is named for the singular values it involves [43]. The Singular Value Decomposition (SVD) can be employed to extract significant and distinct characteristics from the image. The resulting singular values can accurately represent the full image, thereby preserving its valuable details while occupying minimal storage space. The dimensionality reduction property is thus present in SVD [44]. In the field of subspace analysis, the SVD is a mathematical technique that is utilized to diagonalize the input matrix [44, 45]. For the most part, the area of image processing makes use of it to break down images and extract their key features. The matrix I, which has dimensions of m × n, is used to represent the image. Equation 1 provides the SVD of the variable I. The SVD matrices are depicted in Figure 3.

$$I = U \times S \times V^T \quad \text{------------} \ (1)$$

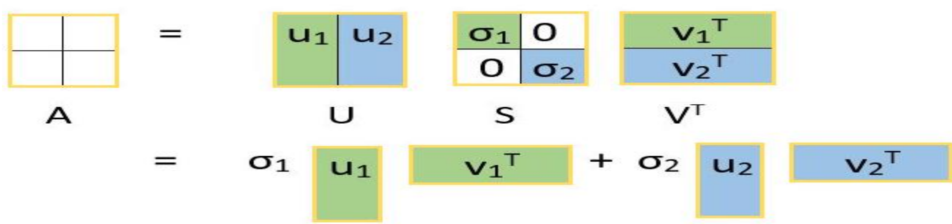


Fig 3. SVD matrices

U is a unitary matrix with dimensions m x m for left singular vectors and V^T is a unitary matrix with dimensions n x n for right singular vectors. The diagonal matrix S m x n shows the singular values, which are shown in decreasing rank [46]. Essentially, According to Equation 2, the SVD theorem converts the input matrix I into a k-low dimensional/rank form by using the truncated forms of U, S, and V^T, denoted as U (k), S_k, and V_k^T, respectively. Here, A, we only retain the highest k single values.Matrixes that have been shortened are displayed in Figure 4.

$$A = U_k \times S_k \times V_k^T \quad \text{------------ (2)}$$

Fig 4. Truncated matrices in SVD

Scikit-learn utilizes the 'TruncatedSVD' estimator to extract features in a linear manner using truncated SVD. With the exception of the first k biggest singular values, which stand for the number of components desired, all other singular values must be set to zero. with only the first k columns from U and V used. The effectiveness of SVD is evident while handling sparse matrices[47].

6. SIMILARITY MEASURE

The similarity metric is a critical component of the image retrieval procedure [48, 49]. Similarity measures are employed to identify photographs in the database that are comparable to the query image [50]. The efficacy of the similarity measures greatly influences the performance of the retrieval process outcomes [51]. The cosine measure, the Manhattan distance, and the Euclidean distance are only a few of the many similarity metrics used by image retrieval algorithms[48].

6.1 Euclidean Distance

The Euclidean distance is calculated by determining the magnitude of the difference between two vectors. Computing the square root of the total of the absolute differences, each squared [52, 53]. Equation 3 provides the Euclidean distance, which is calculated using the total number of data samples (n) and two picture attributes (x and y).

$$\text{Euclidean – distance } (x, y) = \sqrt{\sum_{i=1}^{n} (xi - yi)^2} \quad \text{---------- (3)}$$

6.2 City Block Distance

This distance metric is also known as the Manhattan distance. The calculation involves determining the total of the absolute differences between the feature vectors of two photos [54]. The Manhattan distance is defined by Equation 4, in which n is the overall count of samples, and x and y are two picture characteristics.

$$\text{Manhattan} - \text{distance}\ (x, y) = \sum_{i=1}^{n} |xi - yi| \quad \text{----------}\ (4)$$

6.3 Cosine Distance

Cosine similarity is a metric used to quantify the similarity between two feature vectors representing images based on the angle between them [55]. The Cosine similarity is defined by Equation 5, where x and y represent two picture attributes.

$$\text{Cosine} - \text{Similarity}\ (x, y) = \frac{x.y}{|x||y|} \quad \text{---------}\ (5)$$

7. EVALUATION METRICS

Standard evaluation metrics for image retrieval algorithms comprise precision, recall, and f-score. By employing these criteria, one can assess the performance of an image retrieval technique [56, 57].

7.1 Precision

Precision is determined by the ratio of true positive relevant images successfully retrieved to the total number of photos retrieved, including both true positives and false positives [58, 59]. Equation 2-9 provides the measure of precision for P.

$$P = \frac{TP}{TP+FP} \quad \text{----------}\ (6)$$

7.2 Average Precision

Each class's precision is abbreviated as P, the average precision as AP, and the total number of classes as n [60]. Equation 2-10 provides the expression for AP.

$$\mathrm{AP} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{P_i}}{\mathrm{n}} \quad \text{----------}\ (7)$$

7.3 Recall

The recall is calculated as the ratio of the number of true positive photos (successfully recovered relevant images) to the total number of relevant images in the database, which includes both true positives and false negatives [58, 59]. Equation 2-11 provides the formula for calculating the recall of R.

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad \text{----------}\ (8)$$

7.4 Average Recall

Where n is the total number of classes, AR is the average precision, and R is the recall for each class [60]. The equation that represents AR is Equation 2-12.

$$AR = \frac{\sum_{i=1}^{n} R_i}{n} \quad \text{----------} \ (9)$$

7.5 F-Score

The f-score is calculated by taking the harmonic mean of both the recall and precision [38, 61]. The f-score is defined by Equation 2-13:

$$\text{F-score} = 2 \times \frac{\text{P}\times\text{R}}{\text{P+R}} \quad \text{----------} \ (10)$$

8. THE PROPOSED METHOD

The suggested approach comprises three parts, with the first phase consisting of two steps. The initial phase of this investigation is providing a detailed description of the dataset that was utilized. The second phase involves preprocessing, which includes resizing photos, normalizing data, and separating data. The second phase has three steps. The initial phase involves the utilization of deep learning techniques, specifically VGG-16 fine tuning, to extract relevant features. The second and third steps involve dimensionality reduction using Singular Value Decomposition (SVD) and picture retrieval using three different distance measures: Cosine distance, City block distance, and Euclidean distance. The third phase involves assessing performance through the use of metrics like as precision, recall, and f-score. Further information regarding these stages is provided in the subsequent subsections. Figure 5 displays all the stages of the proposed approach.

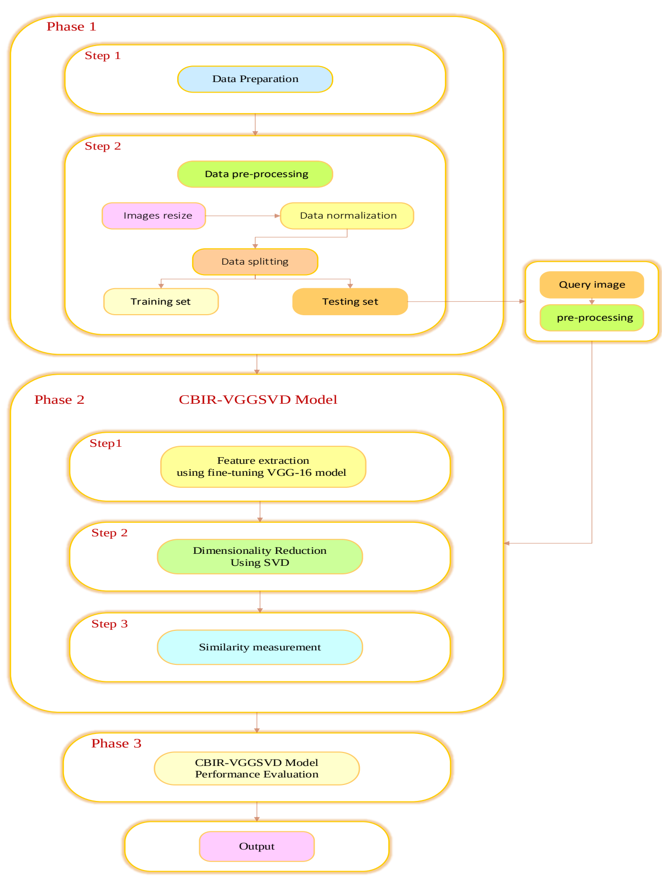


Fig 5. The proposed method

8.1 Corel - 1K Dataset

The Corel-1k dataset [62, 63] Among the ten semantic groups that make up the collection are African People, Beaches, Buildings, Buses, Food, Dinosaurs, Elephants, Flowers, Horses, and Mountains. The photographs themselves are 1000 JPEGs in color. Pictures that are sorted into 10 separate categories based on their semantic context. The dimensions of the photographs in each category range from 256 × 384 pixels to 384 × 256 pixels. This dataset is suitable for validating a Content-Based Image Retrieval (CBIR) system due to its diverse range of photos, encompassing natural landscapes, various activities, as well as a large array of animals and other subjects. Figure 6 displays a selection of photos from 10 different categories in the Corel-1k dataset.

Fig 6. Sample images from 10 classes of the Corel-1k dataset.

9. RESULTS

Based on observations, the Africa people class had the lowest values in the testing set during this division. The f-score, recall, and precision were 0.879032, 0.872000, and 0.875502, respectively, when 10 images were retrieved. The f-score, recall, and precision were all 0.841463, 0.828000, and 0.834677, respectively, after retrieving 20 images. The classes on dinosaurs possessed the greatest values. The retrieved ten photographs yielded f-scores of 1.0 for precision, recall, and precision. Similarly, when retrieving 20 images, the precision, recall, and f-score were all 1.0. The precision value is 0.948 on average. The precision of the image categories pertaining to dinosaurs, buses, flowers, foods, and animals is the highest among them. These images are distinguished by the fact that the backdrop consists primarily of grass, sky, or other single-toned surfaces; that is, the foreground and background are notably dissimilar. The CBIR-VGGSVD model, suggested in this study, demonstrates a robust capability to extract distinctive features specific to these images. However, the photographs of Africa depict people, beaches, elephants, and mountains. The foreground and backdrop of these images are not clearly distinguishable. The optimal divide is to allocate 75% of the data to the training set and 25% to the testing set. Table (1) displays the performance evaluation results for image retrieval, including precision, average precision, recall, average recall, F-score, and average F-score.

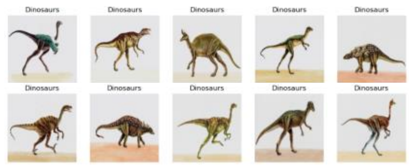
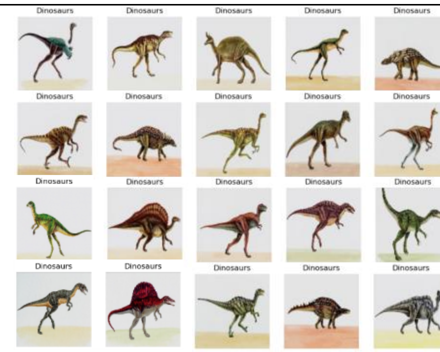
TABAL I. RETRIEVAL RESULTS OF 10 & 20 IMAGES USING MODEL.

Classes	**Precision**		**Recall**		**F-score**	
	10 Images	**20 Images**	**10 Images**	**20 Images**	**10 Images**	**20 Images**
Africa people	0.879032	0.841463	0.872000	0.828000	0.875502	0.834677
Beaches	0.912698	0.863095	0.920000	0.870000	0.916335	0.866534
Buildings	0.898734	0.886161	0.852000	0.794000	0.874743	0.837553
Bus	1.000000	0.988048	0.996000	0.992000	0.997996	0.990020
Dinosaurs	1.0	1.0	1.0	1.0	1.0	1.0
Elephants	0.922481	0.907336	0.952000	0.940000	0.937008	0.923379
Flowers	0.980315	0.957282	0.996000	0.986000	0.988095	0.971429
Foods	0.979508	0.964803	0.956000	0.932000	0.967611	0.948118
Horses	0.968872	0.961014	0.996000	0.986000	0.982249	0.973346
Mountains	0.936255	0.889524	0.940000	0.934000	0.938124	0.911220
Average	**0.948**	**0.926**	**0.948**	**0.926**	**0.948**	**0.926**

The queries employed successfully retrieved similar photos, providing proof of the effectiveness of the suggested CBIR-VGGSVD model in this study. A selection of query images from various classes, together with their corresponding results, is displayed in Table (2).

TABLE II. RETRIEVAL OF 10 & 20 IMAGES OF THE CBIR-VGGSVD MODEL BY A QUERY IMAGE.

cases	Query image	Top 10	Top 20
Spli 1			
			
Split 2			
			

cases	Query image	Top 10	Top 20
Split 3			
			

9.1 Distance Measurement

Essentially, in order to verify the CBIR-VGGSVD model for optimal similarity matching, three methods of distance measurement are employed. With the assistance of these methods, By comparing the query image's feature vector with the image dataset's feature vector, the least feasible distance may be extracted. The following are the practical results of various methods: a typical degree of accuracy of 0.948 when calculating distance using the cosine approach. The Manhattan distance yields an average precision value of 0.939. The mean precision of the distance measurement utilizing the Euclidean distance is 0.947. Table (3) displays a comparative outcome utilizing the cosine, Manhattan, and Euclidean distance measures for the CBIR-VGGSVD model.

TABLE III. COMPARISON OF p & AP OF THE CBIR-VGGSVD MODEL USING COSINE, MANHATTAN, AND EUCLIDEAN DISTANCE.

Classes	Distance measure		
	Cosine	**Manhattan**	**Euclidean**
Africa people	0.879032	0.864	0.883065
Beaches	0.912698	0.899194	0.906883
Buildings	0.898734	0.898734	0.895397
Bus	1	1	1
Dinosaurs	1	1	1
Elephants	0.922481	0.893939	0.919540
Flowers	0.980315	0.987805	0.980080
Foods	0.979508	0.967480	0.971660
Horses	0.968872	0.961390	0.972656
Mountains	0.936255	0.913386	0.940476
Average	**0.948**	0.939	0.947

9.2 Comparative Analysis

When compared to other studies, the CBIR-VGGSVD model's performance is also assessed in this area. It is seen that the dinosaur category yields the most favorable outcome, whereas the African population and beaches produce the least favorable outcome in most approaches.

TABLE IV. COMPARISON OF P OBTAINED BY THE PROPOSED METHOD AND OTHER RETRIEVAL METHODS

Classes	**our proposed**		**Ashraf,R.,[23]**	**Xiaobo,et al. [64]**	**Desai, p., et al. [65]**	**Desai, p., et al. [65]**	**Desai, et al. [65]66**	
	Top 10	**Top 20**	**Top 20**	**Top 10**	**Top 20**	**Top 20**	**Top 10**	**Top 20**
Africa people	**0.879032**	**0.841463**	0.90	0.608	0.9	0.66	0.84	0.84
Beaches	**0.912698**	**0.863095**	0.75	0.594	0.65	0.64	0.8406	0.8286
Buildings	**0.898734**	**0.886161**	0.95	0.628	0.7	0.68	0.8353	0.8526
Bus	**1**	**0.988048**	1	0.913	0.95	0.86	0.8273	0.8306
Dinosaurs	**1**	**1**	1	0.961	1	0.92	0.832	0.834
Elephants	**0.922481**	**0.907336**	0.95	0.467	0.65	0.82	0.8386	0.8353
Flowers	**0.980315**	**0.957282**	0.95	0.850	1	0.98	0.8308	0.836
Foods	**0.979508**	**0.964803**	0.60	0.584	0.65	0.8	0.8373	0.8226
Horses	**0.968872**	**0.961014**	0.95	0.761	0.95	0.98	0.8413	0.834
Mountains	**0.936255**	**0.889524**	0.70	0.397	0.5	0.72	0.838	0.8233

Table (4) presents a comparative analysis of the CBIR-VGGSVD model and several other retrieval methods, which were identified in previous studies and are based on handcrafted techniques, machine learning, deep learning, and hybrid approaches. The analysis was conducted on the Corel-1k dataset. When employing manual techniques, the Average Precision yielded a score of 0.6763 for retrieving the top 10 photographs, and scores of 0.875 and 0.795 for retrieving the top 20 images. The machine learning techniques yielded an Average Precision score of 0.882 when retrieving the top 20 photos. The utilization of deep learning techniques yielded Average Precision scores of 0.89 and 0.86 when fetching the top 20 photographs. for employing hybrid techniques, the Average Precision outcome was 0.8337 for retrieving the top 20 photos. The CBIR-VGGSVD model yields an Average Precision of 0.948 for obtaining the top 10 images and 0.926 for retrieving the top 20 photos. It is clear that the CBIR-VGGSVD model outperforms the most current methods. Table (5) presents a Comparison of the Average Precision (AP) of the CBIR-VGGSVD model with other known approaches on the Corel-1K dataset.

TABLE V. COMPARISON OF AP OF THE CBIR-VGGSVD MODEL WITH THE EXISTING METHODS ON CORE1-1K DATASET

Authors	**Methods**	**Average Precision**	
		Retrieve 10 images	**Retrieve 20 images**
shraf, R, et al.[23] (2020)	Handcraft	-	0.875
Xiaobo, et al.[64] (2021)	Handcraft	0.6763	-
Desai, P., et al.[65] (2021)	Handcraft	-	0.795
Desai, P., et al.[65] (2021)	Machine learning	-	0.882
Kokilambal, S.[67] (2021)	Deep learning	-	0. 89
Desai, et al.[66] (2021)	Hybrid	0.84	0.8337
Kalaivani, et al.[25] (2022)	Deep learning	-	0.86
our proposed	**Hybrid**	**0.948**	**0.926**

10. CONCLUSIONS

The effectiveness of any Content-Based Image Retrieval (CBIR) system mostly relies on the quality of the visually retrieved characteristics from the images. Consequently, the process of extracting image features is crucial in Content-Based Image Retrieval (CBIR). According to transfer learning, pre-trained networks have emerged; these architectures were trained using ImageNet, a large-scale benchmark dataset. The result is that these networks produce highly discriminative descriptors when used as feature extractors. The suggested method replaces manual feature extraction methodologies with an automated process that utilizes a transfer learning-based VGG-19 model to extract features. Due to the high dimensionality of deep features, the suggested method employs SVD to reduce the dimensionality of the input feature vector. This process yields a reduced set of ideal features. The cosine distance metric has been employed as the similarity measure for retrieving similar photos from the collection. The dinosaurs class yields superior results due to its homogenous background and absence of complicated imagery. The optimal dataset division is to allocate 75% for training and 25% for testing. Additionally, the most acceptable dimension for each dataset is 128×128, taking into consideration the appropriate dividing and dimensions for optimal performance. Retrieving 10 photos yields an average precision of 0.948 on the Corel-1k dataset, whereas retrieving 20 images yields an average precision of 0.926. When compared to the most recent techniques, it is clear that the CBIR-VGGSVD model presents a performance that is exceptionally superior.

Conflicts Of Interest

The absence of any competing relationships or biases that could affect the research is explicitly mentioned in the paper.

Funding

The author's paper asserts that the research was conducted on a voluntary basis and without any financial backing from institutions or sponsors.

Acknowledgment

The author acknowledges the institution for their commitment to fostering a research-oriented culture and providing a platform for knowledge dissemination.

References

[1] J. Pradhan, A. K. Pal, M. S. Obaidat, and S. H. Islam, "A post dynamic clustering approach for classification-based image retrieval," in *2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*, 2020, pp. 1-7: IEEE.

[2] M. L. Shuwandy *et al.*, "Sensor-Based Authentication in Smartphone; a Systematic Review," *Journal of Engineering Research*, 2024.

[3] M. Alrahhal and K. Supreethi, "Content-based image retrieval using local patterns and supervised machine learning techniques," in *2019 Amity International Conference on Artificial Intelligence (AICAI)*, 2019, pp. 118-124: IEEE.

[4] R. Abada, A. M. Abubakar, and M. T. Bilal, "An overview on deep leaning application of big data," *Mesopotamian Journal of Big Data*, vol. 2022, pp. 31-35, 2022.

[5] J. Yao, Y. Deng, Y. Yu, and C. Sun, "A fast image retrieval method with convolutional neural networks," in *2017 36th Chinese Control Conference (CCC)*, 2017, pp. 11110-11115: IEEE.

[6] J. Pradhan, A. Ajad, A. K. Pal, and H. Banka, "Multi-level colored directional motif histograms for content-based image retrieval," *The Visual Computer*, vol. 36, no. 9, pp. 1847-1868, 2020.

[7] E. Yakubchyk and I. Yurchak, "RESEARCH OF CONTENT-BASED IMAGE RETRIEVAL ALGORITHMS," 2021.

[8] O. Albahri *et al.*, "Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: Taxonomy analysis, challenges, future solutions and methodological aspects," *Journal of infection and public health,* vol. 13, no. 10, pp. 1381-1396, 2020.

[9] D. Agrawal, A. Agarwal, and D. K. Sharma, "Content-Based Image Retrieval (CBIR): A Review," *Recent Innovations in Computing*, pp. 439-452, 2022.

[10] M. H. Jasim *et al.*, "Emotion detection among Muslims and non-Muslims while listening to Quran recitation using EEG," *Int J Acad Res Bus Soc Sci*, vol. 9, p. 14, 2019.

[11] T. Musta, "PRACTICAL PERFORMANCE OF IMAGE RETRIEVAL METHODS," 2020.

[12] S. M. Samuri, T. V. Nova, B. Rahmatullah, S. L. Wang, and Z. T. Al-Qaysi, "Classification model for breast cancer mammograms," *IIUM Engineering Journal,* vol. 23, no. 1, pp. 187-199, 2022.

[13] X. Li, J. Yang, and J. Ma, "Recent developments of content-based image retrieval (CBIR)," *Neurocomputing*, vol. 452, pp. 675-689, 2021.

[14] R. Bibi, Z. Mehmood, R. M. Yousaf, T. Saba, M. Sardaraz, and A. Rehman, "Query-by-visual-search: multimodal framework for content-based image retrieval," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 11, pp. 5629-5648, 2020.

[15] A. Albahri *et al.*, "A Trustworthy and Explainable Framework for Benchmarking Hybrid Deep Learning Models Based on Chest X-Ray Analysis in CAD Systems," *International Journal of Information Technology & Decision Making*, 2024.
[16] M. A. Ahmed *et al.*, "Real-time sign language framework based on wearable device: analysis of MSL, DataGlove, and gesture recognition," *Soft Computing*, vol. 25, no. 16, pp. 11101-11122, 2021.
[17] N. Seth and S. Jindal, "A review on content based image retrieval," *INTERNATIONAL JOURNAL*, vol. 15, no. 14, 2018.
[18] Z. Al-qaysi, A. Albahri, M. Ahmed, and M. M. Salih, "Dynamic decision-making framework for benchmarking brain–computer interface applications: a fuzzy-weighted zero-inconsistency method for consistent weights and VIKOR for stable rank," *Neural Computing and Applications*, pp. 1-24, 2024.
[19] Y. Cai, Y. Li, C. Qiu, J. Ma, and X. Gao, "Medical image retrieval based on convolutional neural network and supervised hashing," *IEEE access*, vol. 7, pp. 51877-51885, 2019.
[20] S. Garfan *et al.*, "Telehealth utilization during the Covid-19 pandemic: A systematic review," *Computers in biology and medicine*, vol. 138, p. 104878, 2021.
[21] Ö. Göksu and E. Aptoula, "Content based image retrieval of remote sensing images based on deep features," in *2018 26th Signal Processing and Communications Applications Conference (SIU)*, 2018, pp. 1-4: IEEE.
[22] Z. Rian, V. Christanti, and J. Hendryli, "Content-based image retrieval using convolutional neural networks," in *2019 IEEE International Conference on Signals and Systems (ICSigSys)*, 2019, pp. 1-7: IEEE.
[23] R. Ashraf, M. Ahmed, U. Ahmad, M. A. Habib, S. Jabbar, and K. Naseer, "MDCBIR-MF: multimedia data for content-based image retrieval by using multiple features," *Multimedia tools and applications*, vol. 79, no. 13, pp. 8553-8579, 2020.
[24] S. Devulapalli, A. Potti, R. Krishnan, and M. S. Khan, "Experimental evaluation of unsupervised image retrieval application using hybrid feature extraction by integrating deep learning and handcrafted techniques," *Materials Today: Proceedings*, 2021.
[25] K. Kalaivani *et al.*, "A Novel Hyperparameter Tuned Deep Learning Model for Content based Image Retrieval," in *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, 2022, pp. 207-213: IEEE.
[26] T. Karthik, R. Krishna, T. R. Rao, V. Manoranjithem, S. Kalaiarasi, and B. Jegajothi, "Evolutionary Optimization Algorithm on Content based Image Retrieval System using Handcrafted features with Squeeze Networks," in *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, 2022, pp. 1425-1431: IEEE.
[27] P. Staszewski, M. Jaworski, J. Cao, and L. Rutkowski, "A new approach to descriptors generation for image retrieval by analyzing activations of deep neural network layers," *IEEE Transactions on neural networks and learning systems*, 2021.
[28] F. Mustafic, I. Prazina, and V. Ljubovic, "A new method for improving content-based image retrieval using deep learning," in *2019 XXVII International Conference on Information, Communication and Automation Technologies (ICAT)*, 2019, pp. 1-4: IEEE.
[29] S. Wang, K. Han, and J. Jin, "Review of image low-level feature extraction methods for content-based image retrieval," *Sensor Review*, 2019.
[30] M. M. Adnan, M. S. M. Rahim, A. Rehman, Z. Mehmood, T. Saba, and R. A. Naqvi, "Automatic image annotation based on deep learning models: a systematic review and future challenges," *IEEE Access*, vol. 9, pp. 50253-50264, 2021.
[31] A. J. Trappey, C. V. Trappey, and S. Shih, "An intelligent content-based image retrieval methodology using transfer learning for digital IP protection," *Advanced Engineering Informatics*, vol. 48, p. 101291, 2021.
[32] Z. Al-Qaysi *et al.*, "A systematic rank of smart training environment applications with motor imagery brain-computer interface," *Multimedia Tools and Applications*, vol. 82, no. 12, pp. 17905-17927, 2023.
[33] M. Azeem, B. M. Abualsoud, and D. Priyadarshana, "Mobile Big Data Analytics Using Deep Learning and Apache Spark," *Mesopotamian Journal of Big Data*, vol. 2023, pp. 16-28, 2023.
[34] A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical image retrieval using deep convolutional neural network," *Neurocomputing*, vol. 266, pp. 8-20, 2017.
[35] M. M. Salih, M. Ahmed, B. Al-Bander, K. F. Hasan, M. L. Shuwandy, and Z. Al-Qaysi, "Benchmarking framework for COVID-19 classification machine learning method based on fuzzy decision by opinion score method," *Iraqi Journal of Science*, pp. 922-943, 2023.
[36] A. Albahri *et al.*, "A systematic review of using deep learning technology in the steady-state visually evoked potential-based brain-computer interface applications: current trends and future trust methodology," *International Journal of Telemedicine and Applications*, vol. 2023, 2023.
[37] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272-279, 2019.

[38] Z. Al-Qaysi, M. M. Salih, M. L. Shuwandy, M. Ahmed, and Y. S. Altarazi, "Multi-Tiered CNN Model for Motor Imagery Analysis: Enhancing UAV Control in Smart City Infrastructure for Industry 5.0," *Applied Data Science and Analysis*, vol. 2023, pp. 88-101, 2023.
[39] B. Koonce, *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*. Springer, 2021.
[40] R. A. Aljanabi, Z. Al-Qaysi, M. Ahmed, and M. M. Salih, "Hybrid Model for Motor Imagery Biometric Identification," *Iraqi Journal For Computer Science and Mathematics*, vol. 5, no. 1, pp. 1-12, 2024.
[41] M. Hadid, Q. M. Hussein, Z. Al-Qaysi, M. Ahmed, and M. M. Salih, "An Overview of Content-Based Image Retrieval Methods And Techniques," *Iraqi Journal For Computer Science and Mathematics*, vol. 4, no. 3, pp. 66-78, 2023.
[42] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, "A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction," *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 56-70, 2020.
[43] E. A. Compton and S. L. Ernstberger, "Singular Value Decomposition: Applications to Image Processing," *Citations Journal of Undergraduate Research*, vol. 17, 2020.
[44] M. Majhi and A. K. Pal, "An image retrieval scheme based on block level hybrid dct-svd fused features," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 7271-7312, 2021.
[45] N. S. Baqer, H. A. Mohammed, A. Albahri, A. Zaidan, Z. Al-Qaysi, and O. Albahri, "Development of the Internet of Things sensory technology for ensuring proper indoor air quality in hospital facilities: Taxonomy analysis, challenges, motivations, open issues and recommended solution," *Measurement*, vol. 192, p. 110920, 2022.
[46] J. Yadav and K. Sehra, "Large scale dual tree complex wavelet transform based robust features in PCA and SVD subspace for digital image watermarking," *Procedia computer science*, vol. 132, pp. 863-872, 2018.
[47] F. Anowar, S. Sadaoui, and B. Selim, "Conceptual and empirical comparison of dimensionality reduction algorithms (pca, kpca, lda, mds, svd, lle, isomap, le, ica, t-sne)," *Computer Science Review*, vol. 40, p. 100378, 2021.
[48] S. Sudha and S. Aji, "A review on recent advances in remote sensing image retrieval techniques," *Journal of the Indian Society of Remote Sensing*, vol. 47, no. 12, pp. 2129-2139, 2019.
[49] M. Ahmed, B. Zaidan, A. Zaidan, M. M. Salih, Z. Al-Qaysi, and A. Alamoodi, "Based on wearable sensory device in 3D-printed humanoid: A new real-time sign language recognition system," *Measurement*, vol. 168, p. 108431, 2021.
[50] F. Baig *et al.*, "Boosting the performance of the BoVW model using SURF–CoHOG-based sparse features with relevance feedback for CBIR," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 44, no. 1, pp. 99-118, 2020.
[51] R. Kapoor, D. Sharma, and T. Gulati, "State of the art content based image retrieval techniques using deep learning: a survey," *Multimedia Tools and Applications*, vol. 80, no. 19, pp. 29561-29583, 2021.
[52] Y. D. Mistry, "Textural and color descriptor fusion for efficient content-based image retrieval algorithm," *Iran Journal of Computer Science*, vol. 3, no. 3, pp. 169-183, 2020.
[53] S. Agrawal, A. Chowdhary, S. Agarwala, V. Mayya, and S. Kamath S, "Content-based medical image retrieval system for lung diseases using deep CNNs," *International Journal of Information Technology*, vol. 14, no. 7, pp. 3619-3627, 2022.
[54] F. Malik and B. Baharudin, "Analysis of distance metrics in content-based image retrieval using statistical quantized histogram texture features in the DCT domain," *Journal of king saud university-computer and information sciences*, vol. 25, no. 2, pp. 207-218, 2013.
[55] Q. Zheng, X. Tian, M. Yang, and H. Wang, "Differential Learning: A Powerful Tool for Interactive Content-Based Image Retrieval," *Engineering Letters*, vol. 27, no. 1, 2019.
[56] S. R. Dubey, "A decade survey of content based image retrieval using deep learning," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 5, pp. 2687-2704, 2021.
[57] Z. Al-Qaysi, A. Al-Saegh, A. F. Hussein, and M. Ahmed, "Wavelet-based Hybrid learning framework for motor imagery classification," *Iraqi J Electr Electron Eng*, 2022.
[58] M. S. Ghaleb, H. M. Ebied, H. A. Shedeed, and M. F. Tolba, "Image Retrieval Based on Deep Learning," *Journal of System and Management Sciences*, vol. 12, no. 2, pp. 477-496, 2022.
[59] U. A. Khan, A. Javed, and R. Ashraf, "An effective hybrid framework for content based image retrieval (CBIR)," *Multimedia Tools and Applications*, vol. 80, no. 17, pp. 26911-26937, 2021.
[60] K. T. Ahmed, S. Ummesafi, and A. Iqbal, "Content based image retrieval using image features information fusion," *Information Fusion*, vol. 51, pp. 76-99, 2019.
[61] D. Sarvamangala and R. V. Kulkarni, "Convolutional neural networks in medical image understanding: a survey," *Evolutionary intelligence*, vol. 15, no. 1, pp. 1-22, 2022.
[62] "Corel-1K dataset, available on https://www.kaggle.com/datasets/elkamel/corel-images.."
[63] "Corel-1K dataset, available on http://wang.ist.psu.edu/docs/related.."

[64] Z. Xiaobo, P. Jinye, L. Tian, and A. Zhigang, "Image retrieval method based on improved local binary pattern," in *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*, 2021, pp. 256-260: IEEE.

[65] P. Desai, J. Pujari, and C. Sujatha, "Impact of multi-feature extraction on image retrieval and classification using machine learning technique," *SN Computer Science*, vol. 2, no. 3, pp. 1-9, 2021.

[66] P. Desai, J. Pujari, C. Sujatha, A. Kamble, and A. Kambli, "Hybrid Approach for Content-Based Image Retrieval using VGG16 Layered Architecture and SVM: An Application of Deep Learning," *SN Computer Science*, vol. 2, no. 3, pp. 1-9, 2021.

[67] S. Kokilambal, "Intelligent content based image retrieval model using adadelta optimized residual network," in *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 2021, pp. 1-5: IEEE.