Secure Content-Based Image Retrieval with Image Ambiguation

Shreesh Shankar Bhat
KLE Technological University

Padmashree Desai
KLE Technological University

Sujata C
KLE Technological University

이스마일 (M2021765)

DMLAB

Secure Content-Based Image Retrieval with Image Ambiguation

- Abstract
- Literature
- Method
- Results

Introduction

- Why Content Based Image Retrieval is necessary ?

Literature

1. Privacy-preserving CBIR (PP-CBIR)

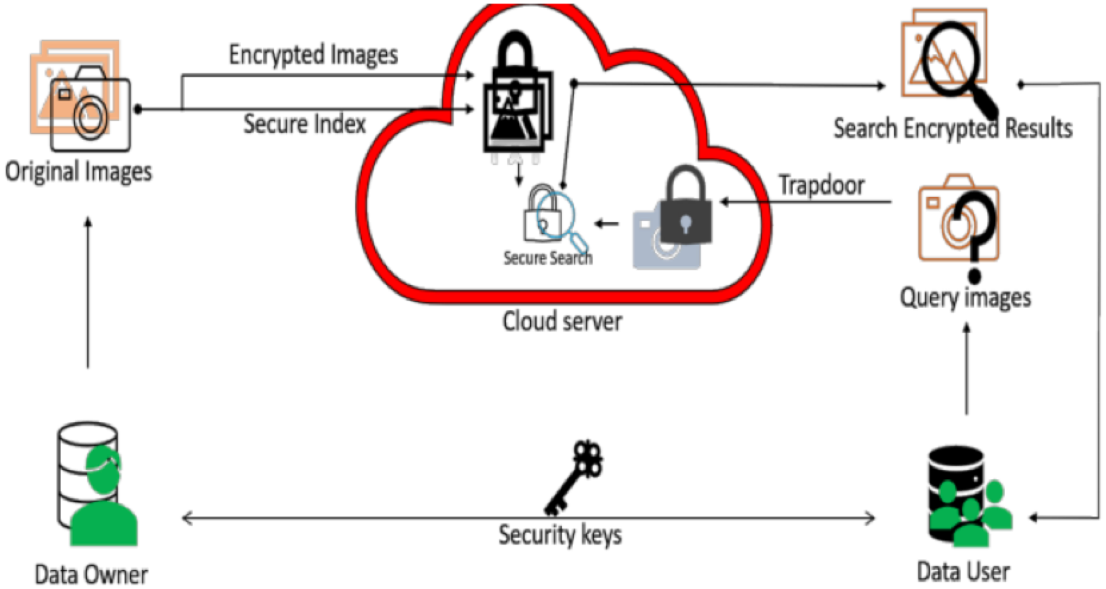


Fig. 2: Proposed scheme PP-CBIR

Privacy preserving Implemented a privacy-preserving CBIR (PP-CBIR) scheme that allows searching and retrieving image databases in a ciphertext format and using K-means clustering to group similar images together to enable faster search.

Literature

1. Privacy-preserving CBIR (PP-CBIR)

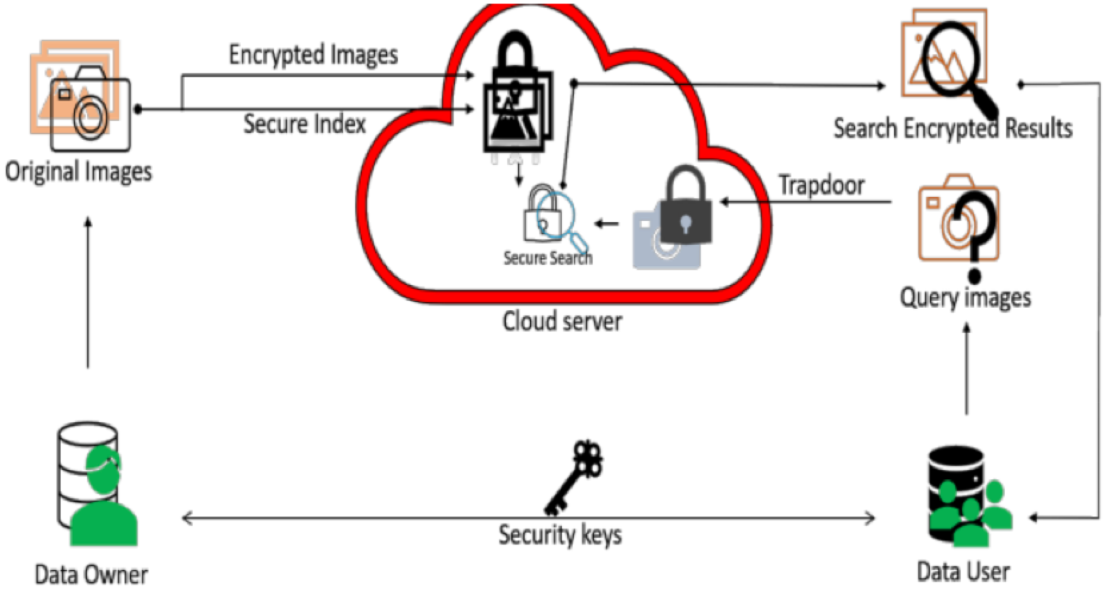


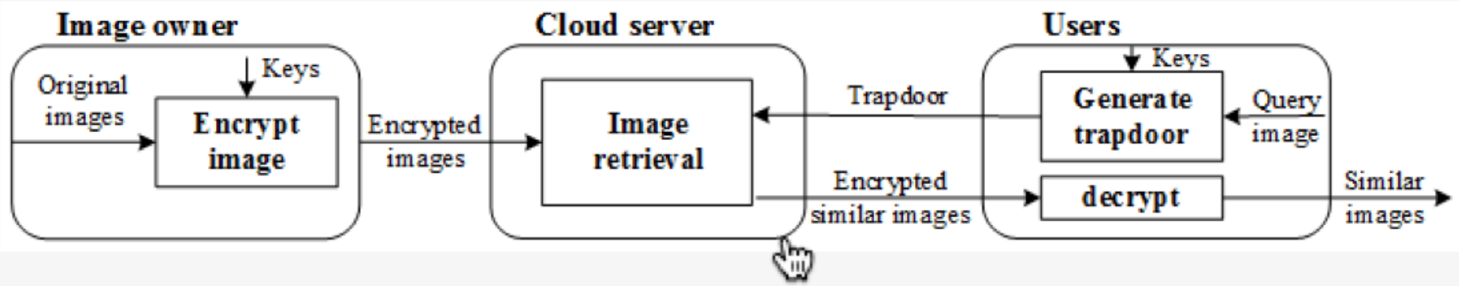
Fig. 2: Proposed scheme PP-CBIR

Each image is described as a single compact aggregated vector that is derived from local descriptors. This method significantly reduces the computation and communication costs. Authors proposed extracting the features on the cloud and encrypting them on the cloud in order to reduce the load on the client device, they also used the ASPE encryption mechanism to encrypt feature data. However, this model does not provide security while transmitting and before the extraction of features and encryption on the cloud

Literature

2. Encrypted Difference Histograms-CBIR)

Figure 1. System model.



EDHCBIR addressed privacy-oriented image retrieval models in the cloud and employs a custom- built encryption technique based on the Difference Histogram in which Encrypted images support feature extraction without decrypting. Retrieval is based on the trap door technique. **A trapdoor** is a method in which a client encrypts a query image and sends it to the server. The server extracts and compares the feature to the database using Euclidean distance.

Drawback

- Drawback of encryption is the high computational overhead while encrypting and decrypting. There needs to be another way of reliably storing the data without encryption and at the same time, preserving the image content.

1. Proposed Method

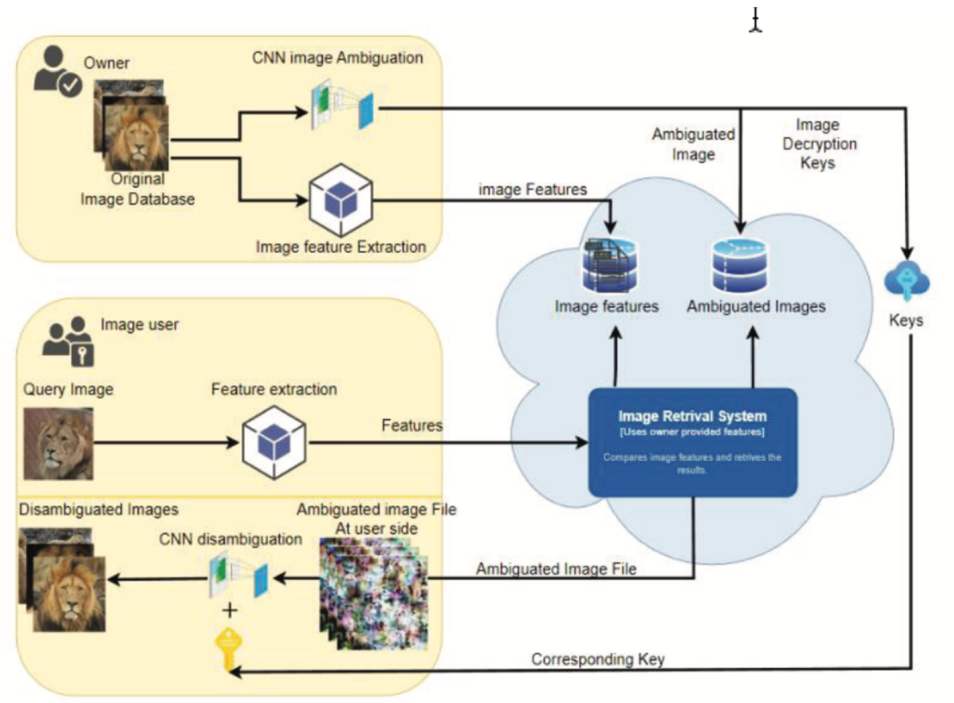


Fig. 1. Overall architecture of the system

A CNN autoencoder is used for image ambiguation. This method avoids the expensive Encryption and decryption to provide security, instead, uses image ambiguity and a key to disambiguate the image. The image can be reconstructed only by the keyholder. This approach has very little key-sharing overhead.

1. The Owner Side

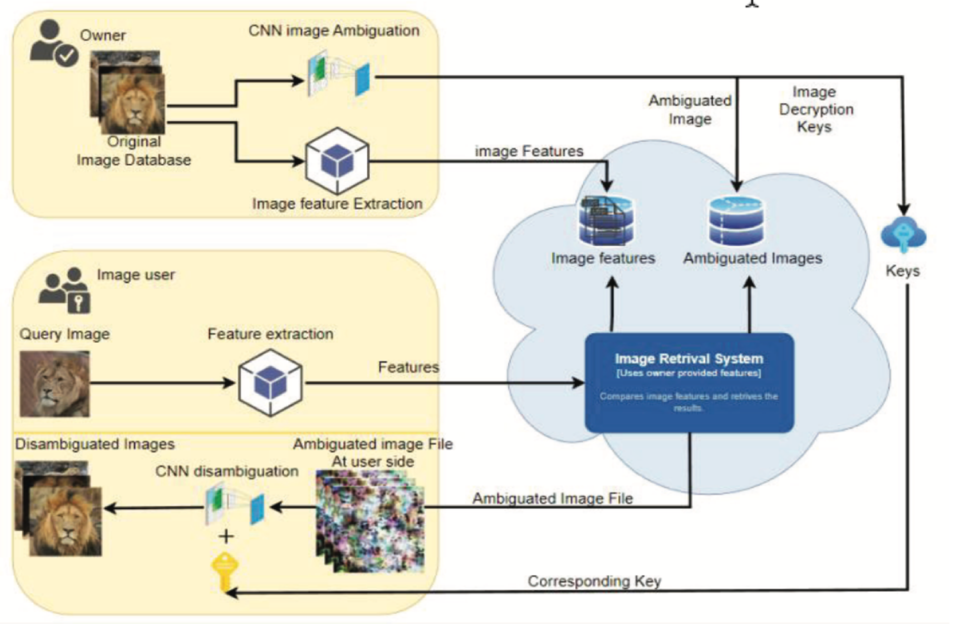


Fig. 1. Overall architecture of the system

On the owner side

- Autoencoder model & VGG16 for feature extraction
- Image database which assume to be owned by the owner

- In order to store the images into the cloud , we will go these steps

➢ We have to extract the image features using VGG16

➢ Then it’s time to ambiguate the images and we will ambiguate the images using the Auto-Encoder Model

➢ The Input of the model is the Original image and the output is ambiguated image with a give sparsity , In their experiment they have used sparsity of 128

1. The Client Side

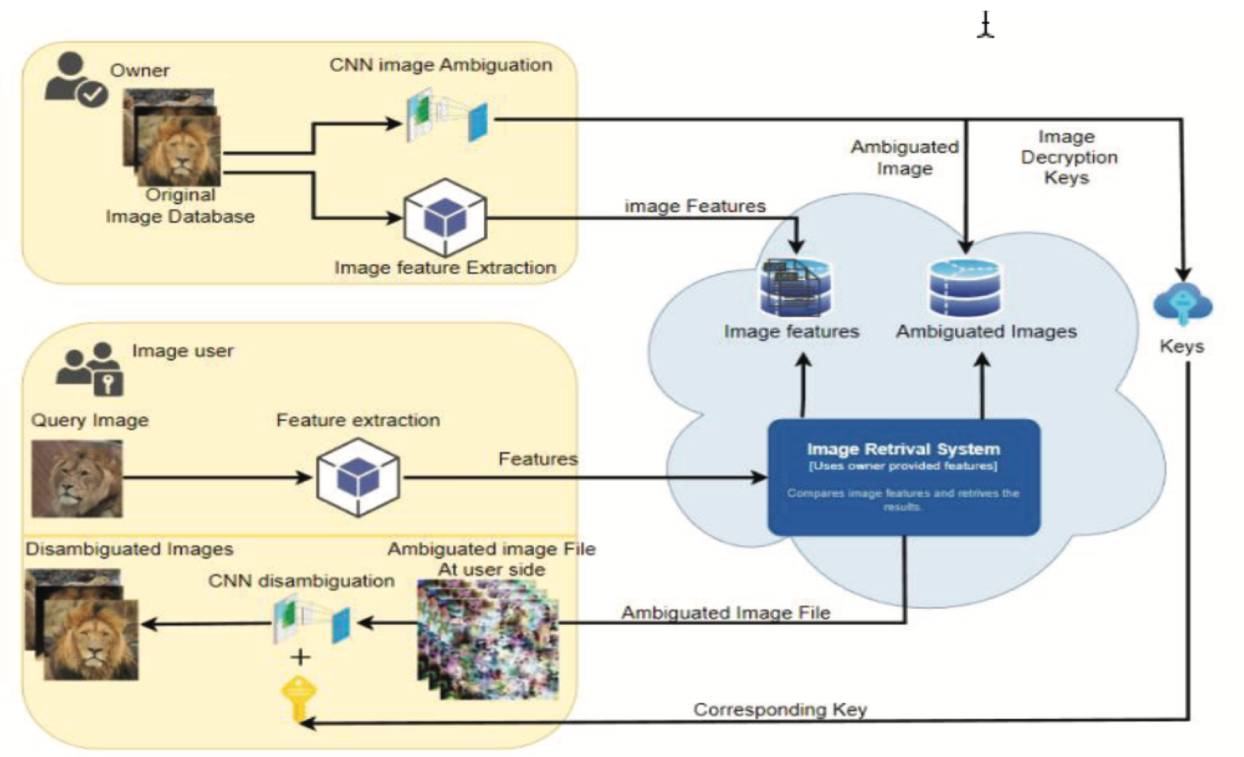


Fig. 1. Overall architecture of the system

On the Client side are similar to that of an owner

- Feature extractor to extract the features from a query image .
- The decoder model of the AE disambiguate images fetched from the cloud database & the image downloader .
- Here Feature extractor is the VGG16 model which is the same as on the owner side which extracts the feature of a given image. The image extracted is sent to the cloud for retrieval. Cloud responds with the list of images. the received list is then used by the downloader to download all the images from the database

IMAGE AMBIGUATION

- Image ambiguation is the process of transforming images into an unusable form. these images can be disambiguated later by the legitimate client with the given key.
- The chosen technique for image ambiguation is based on Autoencoder models. Autoencoders are neural networks that have the same input and output, used here to reduce the dimension of input images and generate a code that represents the image content.
- **Bottlenecked Autoencoder Model**: The chosen Autoencoder model is referred to as a "Bottlenecked Autoencoder," indicating that it specifically focuses on reducing the dimensionality of the input data
- **Sparsity Control**: The codes generated by the encoder are made k-sparse, with k being 128 in this experiment. Sparsity control affects the reconstruction performance of the autoencoder.

IMAGE RETRIEVAL

- For the feature extraction , They have used VGG16 features and Euclidean distance to calculate the distance between the feature vectors

$$distance(q,f) = ||q - f||_2 \sqrt{\sum_{i=1}^{n}(q_i - f_i)^2} \qquad (1)$$

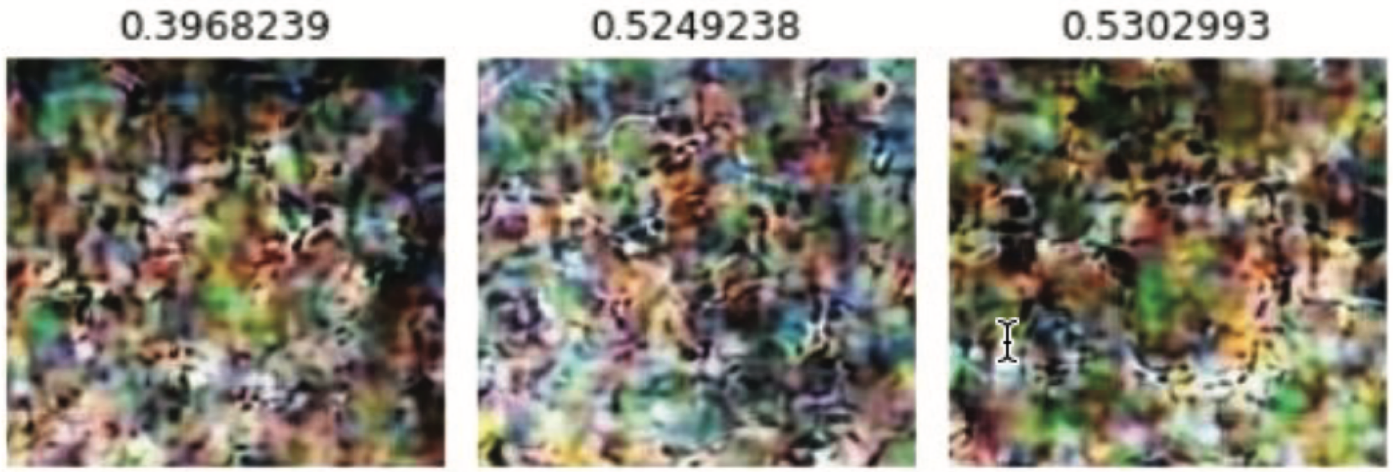
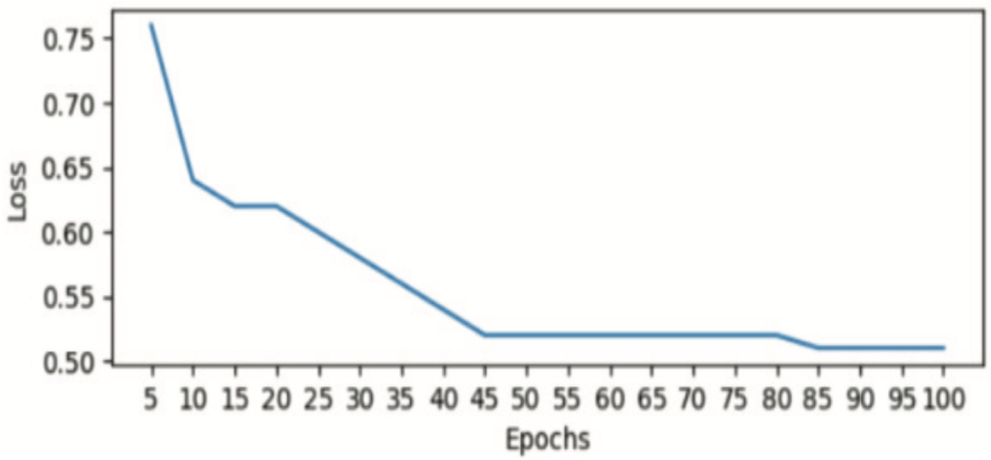


Fig. 6. Images Showing Retrieved images before disambiguation showing the Euclidian distance

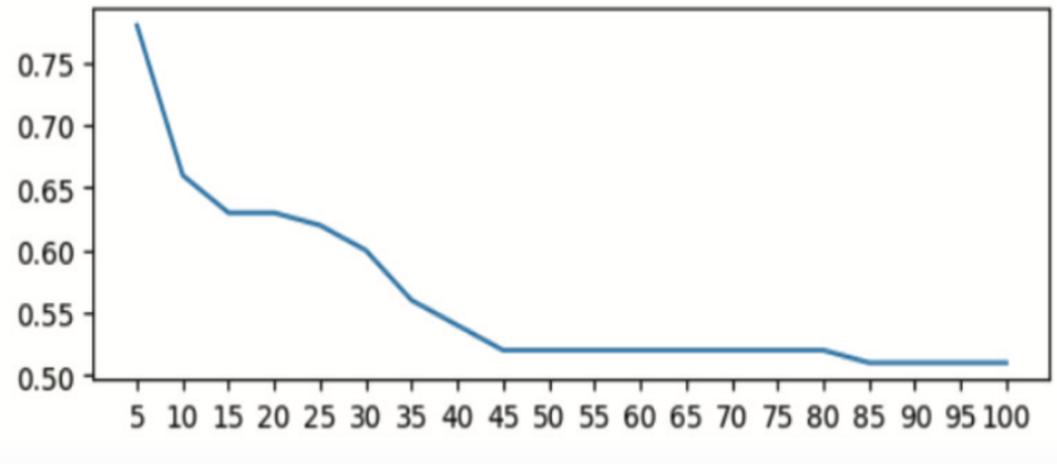
Performance of Image ambiguation

- **Training and Validation Loss**: The Autoencoder model was trained on two datasets, COREL and CBIR for 100 epochs. The reported validation loss for both datasets was 0.51.
- **Training Parameters**: The model was trained with specific hyperparameters, including a batch size of 80 and a learning rate of 0.002.
- **Evaluation Metrics**: To assess the reconstruction performance of the Autoencoder model two metrics were used: SSIM (Structural Similarity Index) and PSNR (Peak Signal-to-Noise Ratio). SSIM measures the structural similarity between two images with values ranging from -1 to +1 where higher values indicate better similarity. PSNR quantifies the quality of compressed or reconstructed images, with higher values indicating better quality.

Train Loss Graph



Validation Loss Graph



Performance of image retrieval

They have evaluated the performance of retrieval with the performance of the retrieval using the F1 score.

- The F1 score is the combination of precision and recall
- It is essentially combining the 2 metrics into one metric by taking the harmonic mean of the given precision and recall. The formula for the f1 score calculation is as in equation 2

$$\frac{2(P*R)}{P+R}$$

Results and Conclusion

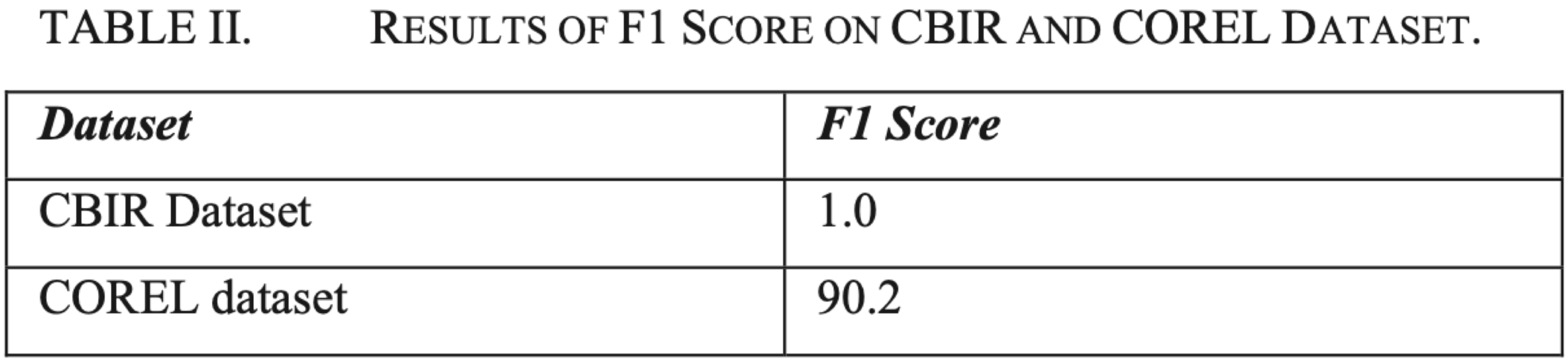


TABLE II. RESULTS OF F1 SCORE ON CBIR AND COREL DATASET.

Dataset	*F1 Score*
CBIR Dataset	1.0
COREL dataset	90.2

- They developed a simple yet effective solution to secure the image data on the cloud with Ambiguation of Images using Auto-Encoders.
- They also developed CBIR system using Euclidean distance for the ambiguated images and obtained the F1 score of 1.0 on the CBIR dataset and 0.902 on the COREL dataset.
- This type of system can be used with existing cloud computing service providers with simple deployment of searching and storage systems on the cloud. In future, we can extend this work on improving the reconstruction performance of the autoencoder model.