

# Content-Based Image Retrieval with Feature Extraction and Rotation Invariance

Nathanael Okoe Larsey, Raphael Mawufemor Kofi Ahiaklo-Kuz\*, Joseph Ncube

School of Information Engineering/Huzhou University, Zhejiang, China

Email: \*natlarsey1@gmail.com

**How to cite this paper:** Larsey, N.O., Ahiaklo-Kuz, R.M.K. and Ncube, J. (2022) Content-Based Image Retrieval with Feature Extraction and Rotation Invariance. *Journal of Computer and Communications*, 10, 24-31. <https://doi.org/10.4236/jcc.2022.104003>

**Received:** March 5, 2022

**Accepted:** April 8, 2022

**Published:** April 11, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Over recent years, Convolutional Neural Networks (CNN) has improved performance on practically every image-based task, including Content-Based Image Retrieval (CBIR). Nevertheless, since features of CNN have altered orientation, training a CBIR system to detect and correct the angle is complex. While it is possible to construct rotation-invariant features by hand, retrieval accuracy will be low because hand engineering only creates low-level features, while deep learning methods build high-level and low-level features simultaneously. This paper presents a novel approach that combines a deep learning orientation angle detection model with the CBIR feature extraction model to correct the rotation angle of any image. This offers a unique construction of a rotation-invariant CBIR system that handles the CNN features that are not rotation invariant. This research also proposes a further study on how a rotation-invariant deep CBIR can recover images from the dataset in real-time. The final results of this system show significant improvement as compared to a default CNN feature extraction model without the OAD.

## Keywords

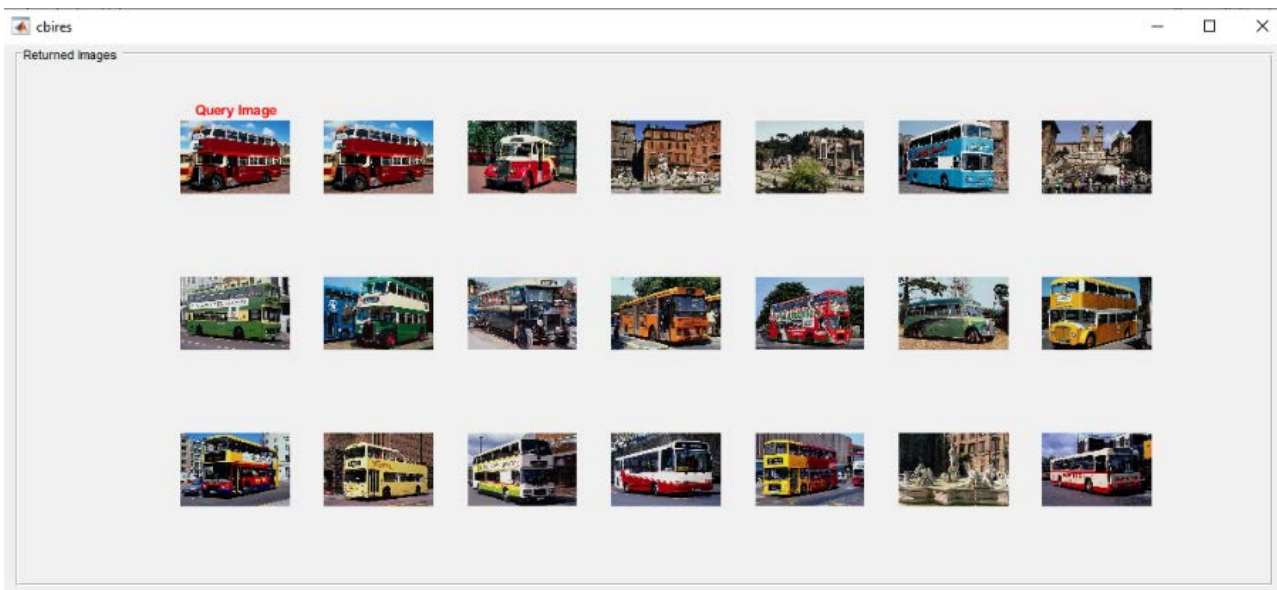
Rotation Invariant CBIR, Image Orientation Angle Detection, Convolutional Neural Network, Deep Learning, Real-Time CBIR, Information Retrieval

## 1. Introduction

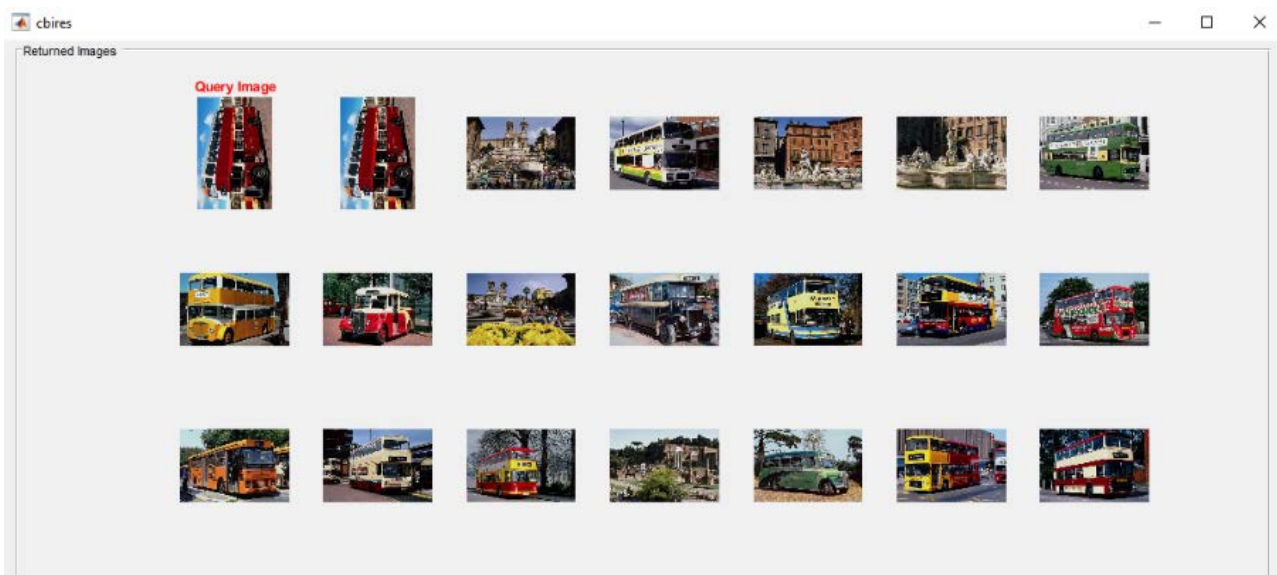
To date, the technique or method used to correct rotation angles in the case of retrieving images from a dataset has not been applied to a pre-trained Convolutional Neural Network to make its rotation invariant. That is to say, if the orientation of the query image changes, the returned images will also change in that regard. In addition, retrieval is hampered by rotated images. **Figure 1** illustrates the point. To make CNN acknowledge rotation angles, it will call for quite a

number of techniques of which several developments are currently underway. One technique is to replace CNN's merging layer with a Spatial Transformer Module [1], which makes it read it in a three-dimensional way and, ultimately, arrives at making it rotation invariant. Unfortunately, the spatial module solution does not give a pre-trained or well-defined architecture. So, to make the retrieval system rotation invariant, a different approach is used.

The Rotation Invariant CBIR system has received very little attention. Bharti *et al.* [2] suggested a method for extracting rotation-invariant curvelet features by analyzing curvelet transform and performing some relevant derivations. Vandhana *et al.* [3] presented a method that uses a Scale-Up Robust Features



(a)



(b)

**Figure 1.** (a) ImageData1; (b) ImageData2.

(SURF) detector to find prominent points. By conducting a mapping from cartesian to logarithmic-polar coordinates, projecting this mapping onto two 1D signature vectors, and computing their power spectra coefficients, Milanese *et al.* [4] proposed a technique for obtaining an image signature that is retrieved from the Fourier power spectrum. Fountain *et al.* [5] used Fourier expansion algorithm to propose a CBIR rotation invariant system which has to do with manipulating the histogram of intensity gradient. To make CBIR rotation invariant, Tzagkarakis *et al.* [6] proposed a method based on a texture data conversion using a steerable pyramid. Chifa *et al.* [7] proposed an approach that involves applying circular masks of various sizes to a picture, extracting the color descriptor from the viewable region on the mask, and merging the results. Krishnamoorthi *et al.* [8] proposed a technique using an orthogonal polynomials model on CBIR. This extracts surface features that have to do with the gray level discrepancies and frequency band of the image being analyzed and the resulting surface feature vector which then becomes rotation invariant.

From related literature, it is evident that models used could not predict the angle of the rotated image and so cannot satisfy accuracy and check the difference in improvement after the orientation angle is corrected. In this research, a separate model is designed and trained to correct the incorrect rotation invariance.

### The Problem

Using only pre-trained, CNN will definitely be a rotation variant and will retrieve images that are still similar to the dataset given and the probability of confusing the query is high. **Figure 1(a)**—ImageData1 query picture is rotationally accurate and has an accuracy of 1. And for the second approach, **Figure 1(b)**—The same query image is rotated 90 degrees anti-clockwise, resulting in essentially identical images with a precision of 0.45. Therefore, the technique is to re-train the pre-trained CNN by correcting the angle back to its rotation accuracy for it to see the image in the accuracy of 1 even though the image is rotated. This method automatically makes the rotation variant pre-trained CNN invariant.

## 2. Methodology

By employing a single pre-trained OAD model, any angle spanning between zero and 359 degrees of every image in the collection is marked. Paper 3 version 1 of the OAD model will be used for the process, where a given dataset for the content-based image retrieval can have any of its images rotated at any arbitrary angle. The images stored as a dataset will then be rotated according to the model's expected orientation angle. The images will then be processed through the pre-trained CBIR model, in this case, InceptionResNetV2 [9] in order to extract their features. The query image will be processed through both models at the moment of retrieval: the first Orientation Angle Detection Model and the second CBIR Model. Then, using the query image's extracted characteristics, a similarity measure is computed and the results are returned. In that regard, every

skewed image will be automatically orientated by the OAD model, yielding identical results as if the image were not tilted at all. If the OAD model fails to compute the correct orientation angle due to its prediction error, the program detects the orientation of the image w.r.t the perpendicular axis and straightens the image. If the prediction error is greater than 1 the image is adjusted to the left to align to the perpendicular axis. If less than 1, the image is adjusted to the right in order to align to the perpendicular axis. **Figure 2** depicts a rotation-invariant CBIR system.

## 2.1. Datasets Used

The CBIR model was used on the image datasets below:

- *ImageData1*: A dataset containing 200 images divided into four groups—Nature, People, Buses, and buildings, making 50 images in each group.
- *ImageData2*: This dataset has 200 pictures divided into two groups. This makes 100 images in each group—Buses and buildings.

## 2.2. OAD Model on CBIR System

To solve the difficulty indicated in **Figure 1**, the OAD model is engineered to rotate the image. As seen in **Figure 3**, the OAD model corrects the image's orientation. The image after OAD correction in **Figure 3(b)** and the previous image in **Figure 1** is now quite comparable.

As a result, it extracted similar images while boosting the accuracy from 0.45 to 1. *ImageData1* and *ImageData2* datasets are compared to further justify the improvement. The objective is to apply an arbitrary angle to the  $n\%$  dataset of images. Next, apply an extraction technique on CBIR features. This can be done by processing the images via the OAD model and then forwarding them into the CBIR model.

**Figure 4** depicts the benefit of combining the OAD and CBIR models on the *ImageData1* and *ImageData2*.

**Table 1** above shows the result. Here,

Improvement = Correctness with OAD Model – Correctness without OAD Model (1)

The OAD model improves the performance of images rotated by different percentages on **Table 1(a)**—*ImageData1* and **Table 1(b)**—*ImageData2* where each number is expressed in percentages.

Thus, when combined with the CBIR model, the precision value increases; yet, when used alone, the precision value decreases. As seen in **Figure 4**, utilizing the OAD model considerably improves the accuracy of the CBIR system when the dataset comprises a high number of rotated images, whereas using the OAD model has a negligible effect when the dataset images are rotationally precise. Additionally, about 5% of rotated photos in both datasets fall between the range of positive and negative improvement. Thus, if a database contains between 5% and 10% rotated images, the correctness level will be high when combining the OAD and CBIR models. By picking a random sample, we can determine the number of images rotated in large datasets.

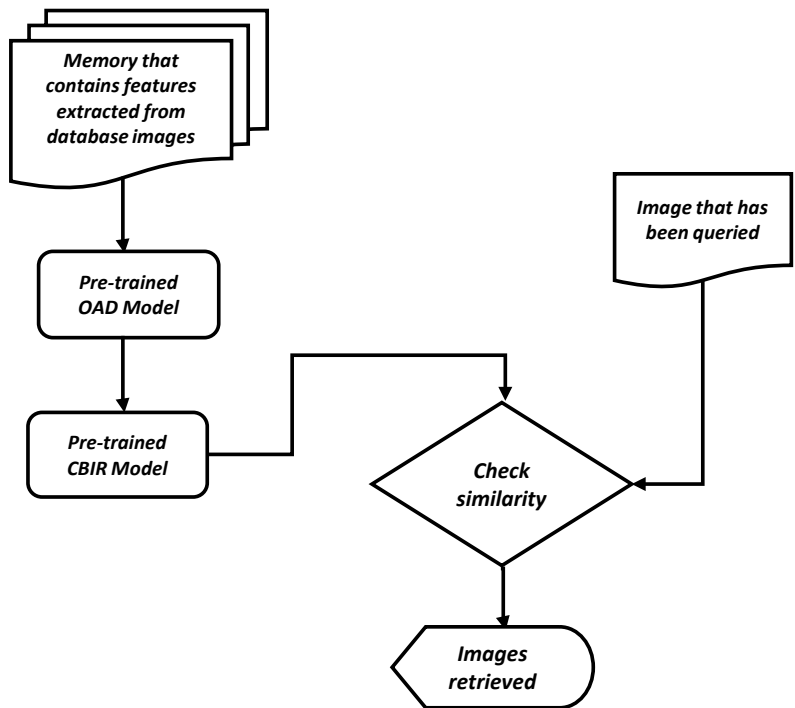


Figure 2. Flowchart of a Rotation Invariant CBIR system.

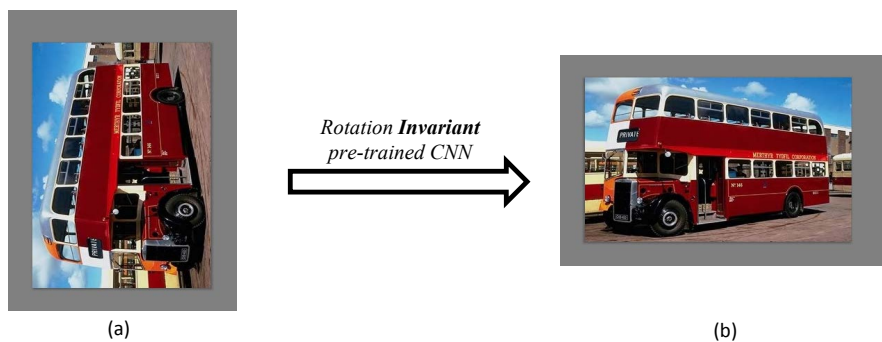
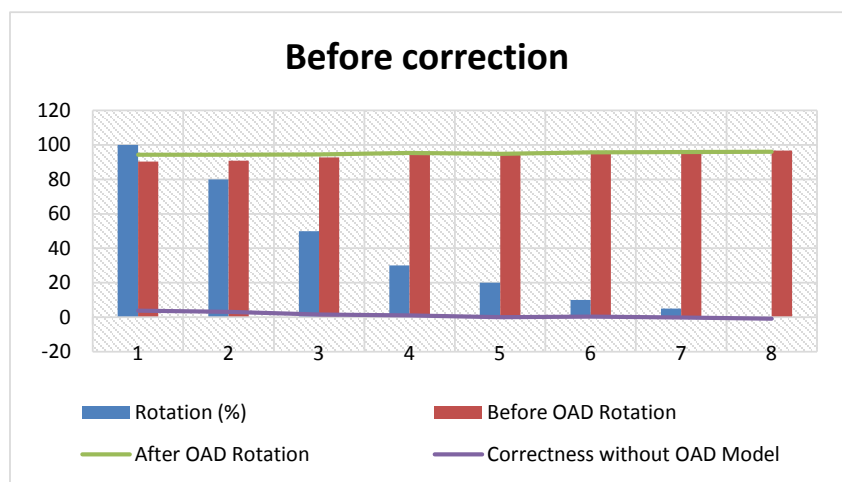
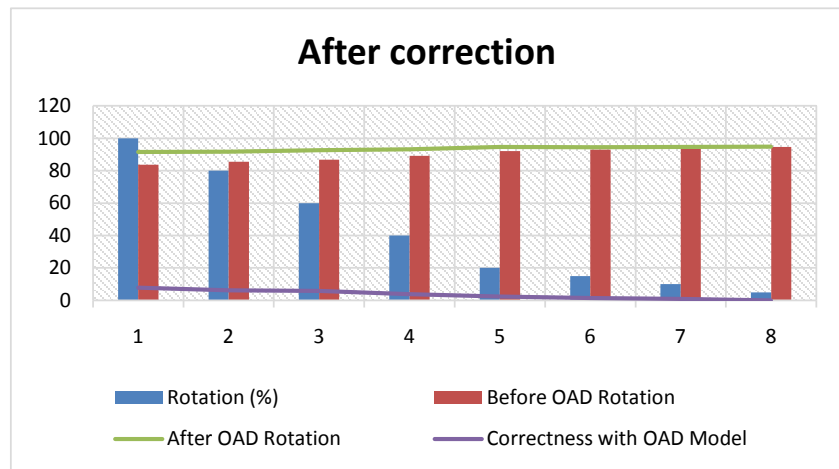


Figure 3. (a) OAD query image; (b) OAD correction of angle rotation.



(a)



(b)

**Figures 4.** (a) and (b) chart shows the performance level of correctness when the OAD model is used on ImageData1 and ImageData2 over different values of percentage of rotated images.

**Table 1.** (a) Correctness without OAD Model. (b) Correctness with OAD Model.

(a)

Rotation (%)	Before OAD Rotation	After OAD Rotation	Correctness without OAD Model
100	90.295	94.175	3.88
80	90.8675	94.135	3.2675
50	92.7195	94.3135	1.594
30	94.125	95.215	1.09
20	94.66	94.71	0.05
10	95.095	95.575	0.48
5	95.7865	95.765	-0.0215
0	96.635	95.875	-0.76

(b)

Rotation (%)	Before OAD Rotation	After OAD Rotation	Correctness with OAD Model
100	83.6	91.5	7.9
80	85.45	91.755	6.305
60	86.685	92.54	5.855
40	89.145	93.14	3.995
20	92.145	94.71	2.565
15	92.975	94.525	1.55
10	93.56	94.625	1.065
5	94.685	94.85	0.165
0	96.125	94.77	-1.355

### 3. Conclusions

By combining a transitional deep learning model to correct the rotation angle of any image, this study offered a unique construction of a rotation-invariant CBIR system that handles the CNN features that are not rotation invariant. Lastly, it demonstrates that combining this extra correction model with the previous CBIR model had no noticeable impact on real-time image retrieval. The inclusion of additional models considerably enhanced the results although image retrieval time remains an issue.

For further research, the pre-trained CNN scale-invariant system has a chance of retrieving images in real-time for ImageData1 and ImageData2. The average query image retrieval time for a scope may be calculated using the two image datasets. The datasets may be processed using OAD and CBIR models, and the recovered features can then be kept in memory as a feature bank. After retrieving and processing the query images, the same models may be used to get and process the original image. The query image's extracted features may then be matched against its feature lists. This explains that the CBIR model which is now rotation invariant can recover images in real-time.

### Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

### References

- [1] Jaderberg, M., Simonyan, K., Vedaldi, A. and Zisserman, A. (2015) Reading Text in the Wild with Convolutional Neural Networks. *International Journal of Computer Vision*, **116**, 1-20. <https://doi.org/10.1007/s11263-015-0823-z>
- [2] Vijaya Bharati, P. and Rama Krishna, A. (2014) Rotation Invariant Content-Based Image Retrieval System. *International Journal of Engineering Trends and Technology*, **17**, 429-438.
- [3] Raja Vadhana, P., Venugopal, N. and Kavitha, S. (2015) Optimized Rotation Invariant Content Based Image Retrieval with Local Binary Pattern. 2015 *International Conference on Computing and Communications Technologies (ICCCCT)*, Chennai, 26-27 February 2015, 306-311. <https://doi.org/10.1109/ICCCCT2.2015.7292766>
- [4] Milanese, R. and Cherbuliez, M. (1999) A Rotation, Translation, and Scale-Invariant Approach to Content-Based Image Retrieval. *Journal of Visual Communication and Image Representation*, **10**, 186-196. <https://doi.org/10.1006/jvci.1999.0411>
- [5] Fountain, S.R. and Tan, T.N. (1998) Efficient Rotation Invariant Texture Features for Content-Based Image Retrieval. *Pattern Recognition*, **31**, 1725-1732. [https://doi.org/10.1016/S0031-3203\(98\)00015-6](https://doi.org/10.1016/S0031-3203(98)00015-6)
- [6] Tzagkarakis, G., Beferull-Lozano, B. and Tsakalides, P. (2006) Rotation-Invariant Texture Retrieval with Gaussianized Steerable Pyramids. *IEEE Transactions on Image Processing*, **15**, 2702-2718. <https://doi.org/10.1109/TIP.2006.877356>
- [7] Chifa, N., Badri, A. and Ruichek, Y. (2019) Rotation-Invariant Approach Using Mask to Content-Based Image Retrieval. *Proceedings of the 2019 5th International Conference on Computer and Technology Applications*, New York, 16 April 2019,

11-14. <https://doi.org/10.1145/3323933.3324066>

- [8] Krishnamoorthi, R. and Devi, S.S. (2013) Rotation Invariant Texture Image Retrieval with Orthogonal Polynomials Model. *Intelligent Computer Vision and Image Processing*, 239-261. <https://doi.org/10.4018/978-1-4666-3906-5.ch017>
- [9] Maji, S. and Bose, S. (2021) CBIR Using Features Derived by Deep Learning. *ACM/IMS Transactions on Data Science*, **2**, 1-24. <https://doi.org/10.1145/3470568>