How to Cite:
Hire, D. N., & Patil, A. V. (2022). Geometrically transformed image retrieval with transfer learning. *International Journal of Health Sciences*, *6*(S2), 11244–11254.
https://doi.org/10.53730/ijhs.v6nS2.8024

Geometrically transformed image retrieval with transfer learning

Mrs. D. N. Hire
E&TC Engineering, DYPCOE, Akurdi Pune, India
Corresponding author email: dnyanadahire@gmail.com

Dr. Mrs. A. V. Patil
E&TC Engineering DYPIEMR, Akurdi Pune, India
Email: anupamav4@gmail.com

***Abstract*---**Image retrieval as per their features is an important topic since last decades. For image processing, classification and retrieval deep learning (DL) techniques utilized as state-of-the art techniques. For the images geometric transformation especially, rotation is common thing. However, Convolutional Neural Networks (CNNs) not supporting rotation invariance features due to use of anisotropic filters for convolution. So, it an important emerging field in image processing that, rotations in images should not affect on its characteristics or features. In this paper, we propose to use features pre-trained CNN model combinations, which trained for large image database containing rotated as well as scaled images for classification & similar image retrieval. This approach produces superior result of classification and retrieval for Corel dataset for original query images and comparably good results by considering rotated & scaled query images.

***Keywords*---**content-based image retrieval, CNN, ResNet18, VGG19, denseNet201.

Introduction

Due to greater computer performance, higher storage capacity, and wider methods of data collecting, DL has been one of the most prominent disciplines in machine learning (ML) during the last decade. In most cases, artificial intelligence applications, particularly computer vision, employ in-depth reading algorithms. Many computer vision applications rely on deep communication networks, which are frequently utilized in visual categorization, acquisition, separation, and other tasks. Convolutional neural networks (CNNs) have emerged as image processing champions due to their ability to rapidly comprehend location-related data in a

Manuscript submitted: 27 March 2022, Manuscript revised: 09 April 2022, Accepted for publication: 18 May 2022

picture. An essential aspect of CNN is that it can record translation consistency in its position encoded characteristics to a certain degree using integration and integration functions. It's adaptable. While restricting CNN's performance, the property lends familiarity to its results compared to fully connected networks, the number of parameters.

The frequency variations are a significant static characteristic that CNNs fail to capture. Rotation is the most common occurrence in the case of photographs. Therefore, the rotation of the rotation will be useful for computer vision systems where the rotation is unfounded in the image the aircraft doesn't properly affect the natural information enclosed in an object or scene affected. Such cases are often observed on aerial/ satellite and small images. The objective of the paper is to penetrate the power of CNN networks to achieve rotation flexibility of 45 and 90 degree and scaling of 0.40 and 0.60 with some extent.

When we look at various research articles on this topic, several methods have been used to deal with problems, the problem of fluctuations. The paper [1] explains how to complete the rotation Flexible CBIR using the Curvelet texture and performance feature into different parts of the same development teams efficiency of recovery. Combine all these features, make a feature vector, and retrieve identical images based on similarities check. This method gave better results even in photos is available in a variety of forms. But here only texture features taken into consideration for measuring similarity index. If we compare more features then there may be chances of increase in retrieval result. Also the experimental results are only calculated for images present in dataset and that of only specific cases of query images. In [2] they use the transformation of the directional pyramid to obtain a more balanced representation and a more sophisticated shape of the text images. Then they employed a hidden Markov tree (HMT) model to represent dependence on all scales and orientations, as well as the statistical aspects of the conversion coefficients, which is an excellent tool for defining the texture element. They used a Brodatz texture site and remote sensor photos to test the algorithm retrieval's success. Experiments have proven that this method restores the image satisfactorily and is less sensitive to the shape of the structure. DL techniques provided state-of-the-art methods [3,4] for image classification and retrieval but CNN fail to detect rotation & scaling changes in query images due to use of anisotropic filters in the convolution process. In [13] author implemented OAD technique to find the orientation angle and correct the angle and then retrieved similar images but here the whole dataset searched for retrieving similar images so the retrieval time will be more. As well as the image first go to OAD model and then deep learning model for similar image retrieval which will further increase processing time. The main objectives of this paper are to implement generic image classification & retrieval technique which will be robust to various geometric transformations and should be easily incorporated to existing DL models. Also, the clustering of images as per classes with classification technique and then retrieving images from that cluster instead of searching whole dataset, to reduce the retrieval time.

Deep Learning Models

Pre-trained models are techniques that AI engineers can use to try to build a deep neural network (DNN) using a framework that already exists. Naturally, building a model from scratch is not possible due to calculation or time constraints. Nowadays, AI engineers prefer to utilize advanced training because there are a few well-developed CNNs models, such as AlexNet, LeNet, VGGNet, and others, and the built-in models are not the same for certain job models. Krizhevskyet al. [7] proposed AlexNet in 2012, competing with another CNN model and winning the ImageNet Large-Scale Visual Recognition Competition (ILSVRC 2012). AlexNet has a similar architecture to LeNet, but with several unique methods such as multi-integration and ReLU nonlinearity. Karen Simonyan [8] won in 2014 using a novel CNN model termed VGGNet, which consisted of 19 DNN. ILSVRC2014, 2nd place. In the same year, Szegedy et al. [9] nominated the ILSVRC2014 GoogleNet winner, which featured new ideas such as 1 * 1 convolutions, which can help to reduce feature map size and make the network deeper and wider. Okay, we ILSVRC 2015, ResNet [10] winner, created a 152-layer network with a new notion that overrides at least two layers or a direct link. Huang et al.[11] established a new concept called DenseBlock in 2016, which allowed the feature to be reused across networks.

We used DenseNet Multi-sclerosis acquisition in this proposed technique since it had the best performance based on the ImageNet segmentation function, as demonstrated in Figure 1. In Figure 1, we compare the outcomes of the single crop top-1 verification error of those trained models through [12]. DenseNet clearly outperforms other pre-trained models. As a result, we chose DenseNet-201 as our primary model for testing.

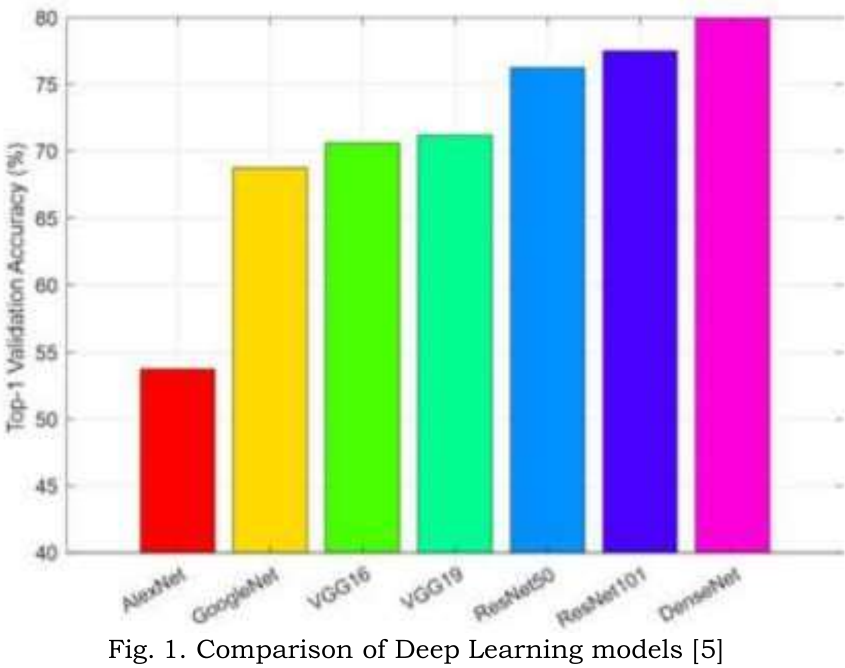

Fig. 1. Comparison of Deep Learning models [5]

DenseNet201Model

The DenseNet-201 is a convolutional neural network with 201 layers. We may use the ImageNet database [6], which has been trained on over a million photographs, to import a pre-trained version of the network. The network can categories photographs into 1000 different object categories, such as keyboard, mice, pencils, and other animals. As a result, the network has accumulated comprehensive feature representations for a wide range of objects. The DenseNet model was shown to be the most effective DL model for image categorization [6].

Methodology

- Stage-I: Extracting features of images from ResNet18 deep learning model. High level features gives more accurate results so the features collected from pooling layer 5 of this pre-trained ResNet18 model. Out of collected features 80% features selected from each category of images for further process by k- mean clustering technique. This will also reduce the processing time for classification as well as retrieval.
- Stage-II: These bag-of-features given to DenseNet201 pre-trained deep learning model for classification purpose. Here the training & testing dataset considered as 70% and 30% respectively. DenseNet201 model trained for 1 epoch with 150 iterations and learning rate of 0.0003.
- Stage-III: After classification, for retrieving similar images, the query image given to classification model DenseNet201 to classify it among anyone category of database images. And display the class and maximum similarity score of query image with selected class database images.
- Stage-IV: Once the class decided then by using cosine similarity, retrieve top 12 similar images at output from selected class. So, if the class of query image is predicted correctly then the accuracy of retrieval will be 100% but if the classification is incorrect the accuracy is 0%. The significant property of vector cosine angle is that it produces a metric of similarity between two vectors, as opposed to Euclidean distance, which gives metrics of variance. Cosine likeness is given by cosine angle between two vectors defined in eq. 1 as,

$$Similarity(x, y) = \frac{x.y}{|x||y|} ---------- (1)$$

If this value is 0, the two vectors are orthogonal (at 90 degrees to each other) and have no match. The lower the angle and the better the match between vectors, the closer the cosine value is to 1.

Database Images Classification

Fig. 2. Image Classification with DenseNet201 CNN model

As mentioned in figure 2 features of dataset images extracted using ResNet18 pooling layer 5 and given to DenseNet201 model. DenseNet201 trained for Corel-1000 dataset with 10 classes and in each class 100 different images were trained with mode. To incorporate rotation and scale invariance, DesnseNet201 trained with 45-degree, 90 degree, 40% scaling and 60% scaling as well. So, Total images used for training under each category of Corel dataset is 500. So total images used to train the network are 5000. For testing the classification of proposed system in real sense, we tested it with query images out of dataset.

Image Retrieval

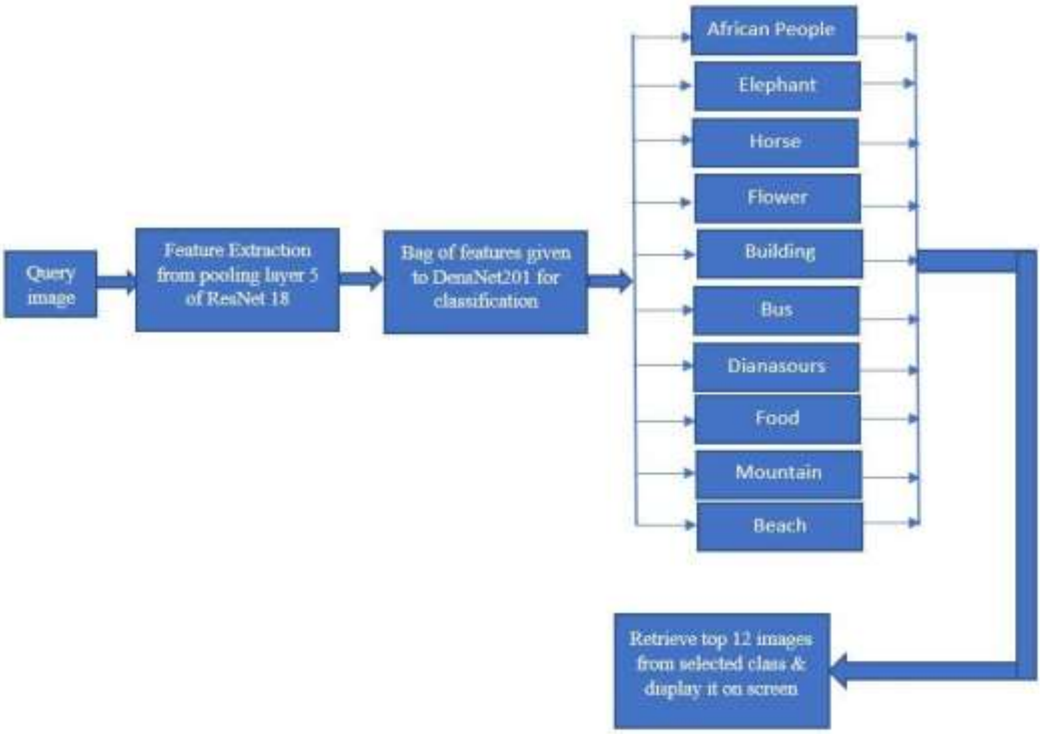

Fig. 3. Image Retrieval with DenseNet201 CNN model

For experimentation, to achieve the rotation and scale invariance simply a method applied where we are training existing CNN networks with images of different rotations and scales. Corel dataset used for analysis of network with 10 different classes viz., African people, beach, building, bus, dinosaurs, food, flower, mountain, elephant, horse. In each class 500 images with 45-degree, 90-degree rotation and 40%, 60% scaling are trained with ResNet18 pre-trained model. Features of all images extracted from pooling layer 5 of ResNet18. Extracted features given to DenseNet201 CNN model for classification and retrieval of images. If the predicted class is correct then the retrieval precision and recall is 100% and 12% (as we retrieved only 12 images) respectively that is maximum values for both performance parameters. But for incorrect predicted class the precision as well as recall values will be zero. For experimentation, the query images will be selected randomly from dataset as well as images form Internet for better analysis of proposed method.

System Configuration

- Single GPU
- Item: GeForce MX350
- CUDA cores: 640
- Memory Data rate: 7.01 Gbps
- Memory Bandwidth: 56.06Gbps

Result and Discussion

Proposed system trained for Corel dataset with 10 categories. For evaluation of system in real manner, we randomly selected query images from dataset as well as any arbitrary image form Internet as reflected in fig 6 and 7. As it is observer from fig 3 that proposed technique provides much better result of retrieval as compared to method, which used original dataset for training model. Fig 4 shows result for all four categories of query images, viz., 45-degree rotation, 90-degree rotation, 60 percent scaling & 40 percent scaling.

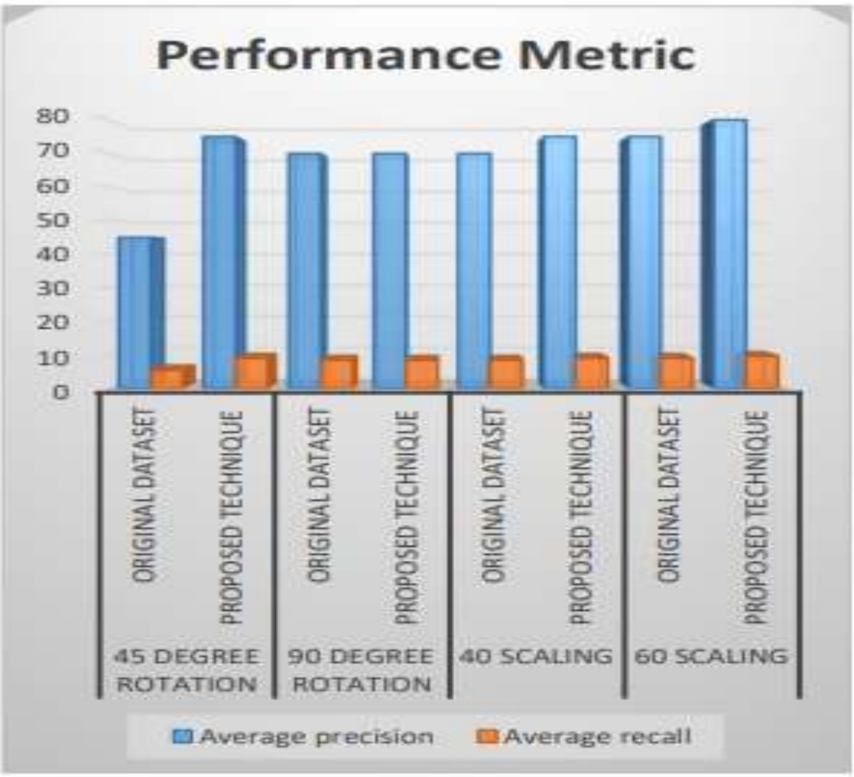

Fig. 4. Precision-recall graph for proposed method

Performance metric shows that the proposed technique provides better result in case of rotation and scaling as compared to previous techniques. But the computational complexity increases on other hand as the above fig 5 shows the precision and recall performance parameters values for 10 epoch whereas proposed technique parameters calculated for only 1 epoch. If we increase epoch value then we may get even better result but at the same time the training time & computational complexity will be more.

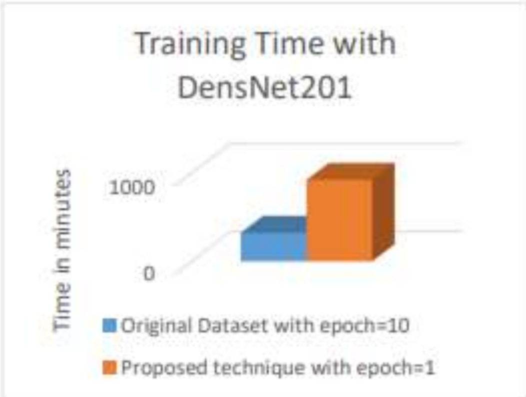

Fig. 5. Training time required for model to train dataset of proposed method

Above fig. 5 proved that time complexity for proposed method is increased extensively in cost of increase in accuracy value as compared to the original database technique used.

Fig. 6. Query images from Internet for each category

Fig. 6 indicates query images manually selected from Internet with maximum difficulty level for classification & retrieval of proposed system.

Fig. 7. Query images from Dataset for each category

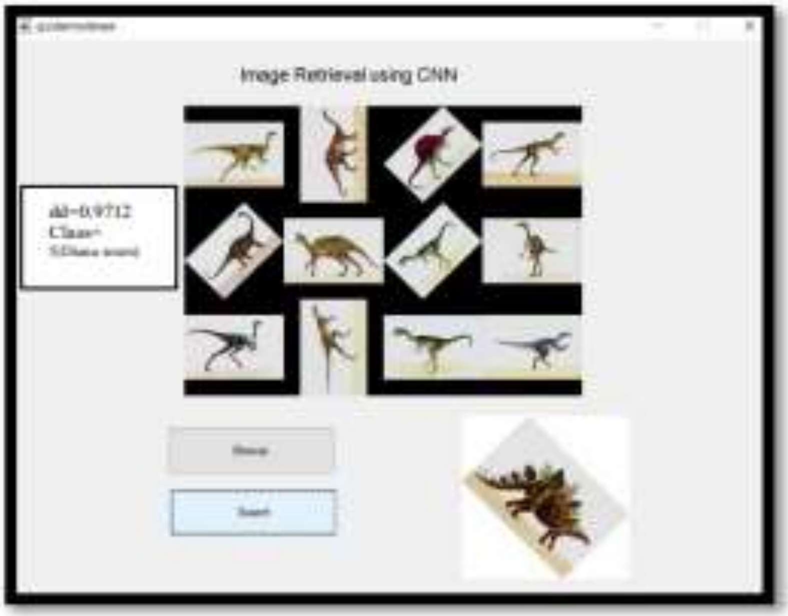

Fig. 8. Classification & Retrieval of 45-degree rotated query image of Dinosaurs from Corel dataset

Figure 8 indicates the result of 45- degree rotated query image of dinosaur from dataset, which is classified correctly, and similarity index is 0.9712. Fig.9 shows example of wrongly classified and retrieved images for 45-degree rotated query image of building from Internet. Table 1 and 2 shows significant improvement in average precision as well as average recall values with proposed technique.

Fig. 9. Classification & Retrieval of 45-degree rotated query image of building from Internet

Query Image	With Original Dataset	With Proposed System	Improvement (%)
45-degree rotation	45	75	30
90-degree rotation	70	70	0
40-scaling	70	75	5
60-scaling	75	80	5

Table 1. Improvement in Average Precision in Percentage

Query Image	With Original Dataset	With Proposed System	Improvement (%)
45-degree rotation	5.4	9	3.6
90-degree rotation	8.4	8.4	0
40-scaling	8.4	9	0.6
60-scaling	9	9.6	0.6

Table 2. Improvement in Average Recall in Percentage

Conclusion

Based on the results of the experimental study of training set, classification test, and image retrieval performed on this CBIR programme using the CNN method, we can conclude that our trained DenseNet201 with 0.0003 learning rate, 1 epoch, was successful in classifying the image in the validation dataset with an accuracy (F1-score) of 96.66 percent, and an average of precision in retrieval considering 45 degree rotated original query images is 75 percent, whereas for 90 degree rotated original query image value is 70 percent. In the validation dataset, there are three classes that are frequently misclassified at the stage of predicting the class or image classification and retrieval such as blossom, construction, foods, and Africans. So, we can conclude from experiment result as DenseNet201 model can be used for rotation & scale invariance image retrieval by training it with rotated & scaled images.

Acknowledgement

This research work is supported by Dr. D. Y. Patil Institute of Engineering, Management and Research, Akurdi, Pune by providing access to the reputed journal papers.

Conflict of Interest

There is no any conflict of interest

References

1. Fatma, Dr Gulnaz & Pirzada, Nahla & Begum, Sameena. (2022). Problems, Illusions and Challenges Faced by a non -Arabic Speaker in Understanding Quran: A Sub-Continental Study. 5422-5426.
2. ChandraSekhar, P. N. R. L., SuryaPrasad, P.; Kumar, M. Vinodh; Santosh, D. Hari Hara, "Image Retrieval with Rotation Invarience", International Conference on Electronics Computer Technology - Image retrieval with rotation invariance. , April2011, 194–198, DOI: 10.1109/ICECTECH.2011.5941683.
3. Congcong Miao, Yindi Zhao, "Rotation-invariant image retrieval using hidden Markov tree for remote sensing data", Proceedings Volume 9263, Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications; 92631W (2014) https://doi.org/10.1117/12.2068889 Event: SPIE Asia-Pacific Remote Sensing, 2014, Beijing, China.
4. Tripathi, M. A., Tripathi, R., Sharma, N., Singhal, S., Jindal, M., & Aarif, M. (2022). A brief study on entrepreneurship and its classification. International Journal of Health Sciences, 6(S2). https://doi.org/10.53730/ijhs.v6nS2.6907
5. Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep Learning", Nature 521, pp.436-444,2015.
6. Fuyong Xing, Yuanpu Xie, Hai Su, Fujun Liu and Lin Yang, "Deep Learning in Microscopy Image Analysis: A Survey", IEEE Transactions on Neural Networks and Learning Systems, vol.29, no.10, pp.4550-4568, 2018.
7. Suihua Wang, Yu-Dong Zang, "DenseNet-201-Based Deep Neural Network with Composite Learning Factor and Precomputation for Multiple Sclerosis Classification", ACM Transaction on Multimedia Computing, Communications and Applications, 16(2s):1-19, June-2020, DOI:10.1145/3341095
8. ImageNet. http://www.image-net.org.
9. Alalmai, Ali & Fatma, Dr Gulnaz & A., Arun & Aarif, Mohd. (2022). Significance and Challenges of Online Education during and After Covid-19. Türk Fizyoterapi ve Rehabilitasyon Dergisi/Turkish Journal of Physiotherapy and Rehabilitation. 32. 6509-6520.

10. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks” In Proceedings of the 25th International Conference on Neural Information Processing Systems” - Volume 1 (Lake Tahoe, Nevada, 2012). Curran Associates Inc., 2999257, 1097–1105
11. A. Z. Karen Simonyan, “Very deep convolutional networks for large-scale image recognition”, ICLR-2015
12. Alalmai, Ali A., and Mohd Aarif. "Importance of Effective Business Communication for Promoting and Developing Hospitality Industry in Saudi Arabia A Case Study of Gizan (Jazan)." (2019).
13. Alalmai, Ali & A., Arun & Aarif, Mohd. (2022). Social Media Advertising Impact on the Consumer Purchasing.
14. C. Szegedy, L. Wei, J. Yangqing, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions” In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1–9, 2015.
15. K. He, X. Zhang, S. Ren, and J. Sun. 2016. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2016), 9.
16. G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger. 2017. Densely connected convolutional networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2261–2269.
17. X. B. Jin, T. L. Su, J. L. Kong, Y. T. Bai, B. B. Miao, and C. Dou. 2018. State-of-the-art mobile intelligence: Enabling robots to move like humans by estimating mobility with artificial intelligence. Applied Sciences-Basel 8, 3 (2018), 39.
18. Subhadip Maji, Smarajit Bose. Rotation Invariant Deep CBIR. https://arxiv.org/pdf/2006.13046.pdf, Preprint-June 2021