**International Journal of Software Innovation**
Volume 10 • Issue 2

# Analysis of Image Similarity Using CNN and ANNOY

Jun-Ki Hong, Pai Chai University, South Korea*

## ABSTRACT

This article proposes an algorithm to more efficiently search for clothing product images that are similar to a new input clothing product image. Convolutional Neural Network (CNN) and Artificial Neural Network Oh Yeah (ANNOY) technologies were applied to a database of 60,000 clothing images, and the similarity and the processing rates of the two technologies were compared. The conventional CNN technology searches similar images by exploring all the pixels of an image, while the ANNOY technology uses a binary tree node, which is the similarity distance measured between images. The ANNOY technology can drastically reduce image search time, although the image similarity accuracy is slightly decreased. The reduction in image search time saves costs, and the rapid search processing rate enables the technology to be applied to various kinds of online services, including product search, product comparison, and product recommendation.

## KEYWORDS

ANNOY, Clothing, CNN, Image Similarity

## 1. INTRODUCTION

With recent developments in deep learning technique, an increasing number of studies are being conducted on images and image recognition (Al-Halah, Z., Stiefelhagen R., Grauman K. 2017; Mihm, B. 2010; Lee, H., Seol J., Sang-goo Lee. 2017). Among these, efforts are focused on methods of searching clothing product images, that can identify fashion trends and clothing features. Such a system should be capable of analyzing the fashion trends of many clothing products quickly. In the present study, a clothing product image search system is proposed and its performance evaluated. The system collects clothing product images from about 300 clothing-related shopping malls on a real-time basis, rapidly classifies the clothing images relative to a new input image, and analyzes the trend in clothing products.

Figure 1 shows an example of a similar image search for the 'S' Shopping Mall. The images in Figure 1 were searched for specific clothing product patterns, poses and tags. A feature-based analytical technique is employed. The deep learning model provides a simple comparison of characteristics, stable service while handling numerous products, rapid updating of new products, and can process a large number of search inquiries.

DOI: 10.4018/IJSI.289593 *Corresponding Author

**Figure 1. Similar image results of clothing**

Establishing a deep learning system for precise searches based on a huge number of images can be high cost, which small companies can hardly afford. Many recent studies on image searches have been conducted using Convolutional Neural Network (CNN). However, similar image analysis with CNN is performed by exploring all the pixels of an image, and thus requires countlessly repeated learning procedures, and a long time to measure similarity (Lee, S., Choi M.-H, Lee H.-J., et al., 2019; Seo, M., Lee K.-Y., 2019). In addition, if the frequency of the repeated learning procedures is decreased to increase computation speed, the level of similarity is reduced.

To overcome these problems related to CNN, this article proposes a model for similar image searches based on the Artificial Neural Network Oh Yeah (ANNOY) using a binary tree, with a little trade-off in image similarity (Bernhardsson E., 2020). The image similarity and the image search measurement time using the CNN and the ANNOY were compared. For the performance comparison, a clothing product dataset was constructed to collect images from about 300 clothing-related shopping malls. The similarity among the fashion products were compare and analyzed using the two techniques. Python was used to collect and update the images of fashion products from about 300 shopping malls on a real-time basis. The image dataset was constructed using 60,000 collected images: 40,000 images (67%) were used for transfer learning, and the other 20,000 images (33%) were used as the input images for the similar image search. The images' similarity and measurement times for the two techniques were analyzed and compared.

## 2. CNN AND ANNOY

### 2.1 CNN

The CNN extracts features at gradually increasing levels, from a lower level to a higher level. The CNN is highly extendable, and has been employed in many applications, especially for large-scale image comparisons, because of its excellent performance in the field. For example, recent large-scale image searches using the CNN have been conducted in the medical field, where it has been used to analyze the disease diagnosis rate based on a disease case dataset, or to detect errors in mechanical equipment (Krizhevsky, A., Sutskever I., Hinton G. E. 2012; Seo, K., Choi, H-Y. 2017).

Figure 2 shows the image analysis algorithm based on the CNN. An image to be analyzed is input through the convolutional layer to extract the minor features in the image, and the effective values are verified through a pooling layer. This procedure is then repeatedly performed to model all of the features of the input image.

### 2.2 ANNOY

The ANNOY analyzes the pixels of an image to search for similarity using the binary tree technique. Depending on the similarity of each image, it allocates a domain with a vector value that is most