Study on Infringement Identification of Art Works Based on CNN Image Recognition Technology

Yan Chen 1 and Rui Qu 2, *

1 Shenzhen University College of Management, Shenzhen, China
2 Shenzhen University Law School, Shenzhen, China

*Corresponding author e-mail: 2018031188@email.szu.edu.cn

Abstract. In view of the current situation where the criteria for judging the infringement of art works are still vague and the court discretionary process is highly subjective, this article introduces the working principle of convolutional neural network (CNN) which is based on image recognition technology, and uses CNN algorithm to extract and store the characteristics of related works, and use the method of cosine similarity criterion to compare and calculate the degree of similarity, to assist in judging the standard of substantial similarity, and to improve the objectivity of the criterion for identifying copyright infringement of artworks.

1. Raising the problem
Artwork refers to a work of plastic art in a plane or a three-dimensional form with aesthetic significance, such as painting, calligraphy, sculpture, etc., which is composed of lines, colors or other means.[1] Substantial similarity is a legal principle that is open and reasonable application in copyright law, belongs to the scope of court discretion, and it is also a core legal issue to be resolved in handling infringement cases. [2] Especially in the context of non-verbal similarity, the subjectivity and fuzziness of the substantial similarity criteria make the application process difficult. At present, Chinese courts mainly use the "holistic perception method" and the "abstract-filter-comparison" three-step method as the main methods of substantial similarity standards. The "overall contrast method" refers to determining whether subsequent works use the core and essence of previous works based on ordinary people's overall perception of the work. It is highly subjective and ignores the copyright law attributes of the creative elements in the work. The three-step judgment of "abstract-Filter-comparison" is to first remove the "idea" part, filter out the same part which belongs to the public domain, and finally compare the rest parts. Even so, the interference of subjective judgment still cannot be eliminated. [3] For artwork, judicial practice is mainly based on the main features for substantial similarity judgments and comparisons, for example "head and facial features: big eyes, glasses in the upper two-thirds of the head... The burgundy nose is large and convex oval; lips are raised; the mouth and nose are forward in a white peach-shaped pattern."[4]; "The head and body are about two to three; the skin on the body and the deep tail are yellow, with black stripes on it." [5]; "There are eyeballs and upper eyelids that can move up and down; the air intake grille is a large flat mouth with white teeth exposed." [6] The artificial subjective inference of the similarity degree through the elements of the work is also of great uncertainty, which is not conducive to safeguarding the interests of the parties and the authority of the judiciary.
With the development of image recognition technology, the introduction of algorithm technology conforms to the background of the in-depth integration of artificial intelligence and the legal system, and also brings a relatively objective standard for the determination of infringement.

2. Analysis of the principle and method of content-based image retrieval
The image representation method based on convolutional neural network (CNN) makes up for the inefficiency and inaccuracy of traditional content-based image retrieval tasks (CBIR). In this paper, the idea of using CNN involves a set of image retrieval systems-apply CNN to extract image features to build a feature database, and then use the method of cosine similarity to calculate the feature similarity between images.

2.1. Working principle of CNN
CNN is mainly used for the recognition of optical image targets, which consists of convolutional layer, pooling layer, nonlinear element and full connection layer. Each layer converts the input value into the output value through its own function to achieve the corresponding task requirements. After inputting a picture into the first layer, each time the picture passes through a new layer, the most meaningful features in the picture will be extracted and transferred to the next layer. With each layer, the semantic representation of the new image becomes more multidimensional. Typically, the first layer of a neural network extracts low-level features (such as edge features), and the higher layers combine those features into more abstract features. Each filter is a unique set of weights that are multiplied by the pixel values in the input image to produce a new image, a process known as convolution, in which the depth of the newly generated image channel encapsulates the learned features.

2.2. Residual structure model
A short circuit mechanism is introduced into the structure of the residual structure so that the output of the model can be expressed as a linear superposition of the input and the nonlinear transformation of the input. Generally, the stronger the deep network learning ability is, the richer the acquired features are, and the more explicit the semantic information is. However, with the increase of network layers, the phenomenon of network degradation will occur. To solve this problem, the residual structure introduces the identity and other quick links, and directly takes the identity mapping as a part of the network, so as to change the learning objective into learning a residual function:

\[ F(x) = H(x) - x \]  

(1)

When \( F(x) = 0 \), that is, the stacked network layer only does the identity mapping. The residual structure will ensure that the model performance will not decrease after the network is deepened. In fact, \( F(x) \neq 0 \), \( F(x) \) will also be input New features are learned on the basis of features, and the residual structure is superimposed with \( F(x) \) to strengthen feature reuse, so as to have better performance.

Figure 1. Residual module for ResNet
2.3 Calculation of feature similarity

Content-based image retrieval is carried out by calculating the similarity of the visual features between the query and the database image. Therefore, it is necessary to define an appropriate visual feature similarity measurement method. The vector space model (VSM) was proposed by Salton et al. in the 1970s. VSM simplifies the processing of visual feature similarity to vector calculation in vector space, that is, measures the similarity between image features by calculating the closeness between two points.

Cosine similarity evaluates their similarity by calculating the cosine of the angle between two vectors. Cosine similarity draws the vector into the vector space according to the coordinate value. After the image passes through the fully connected layer of CNN, a vector is obtained, and the cosine distance between the vectors is calculated to characterize the similarity of two pictures. Assuming that the feature vector of the image to be retrieved is \( A \) and the feature vector of the database image is \( B \), the cosine similarity is:

\[
\cos \theta = \frac{a \cdot b}{|a||b|}
\]

Cosine similarity uses the cosine value of the angle between two vectors in the vector space as a measure of the difference between two individuals. By calculating the cosine value of the angle \( \theta \) between two vectors in the vector space, and using this to measure the difference between two individuals, the closer the cosine value is to 1, the closer the angle number is to 0, indicating that the two vectors are more similar.

![Figure 2](image)

**Figure 2.** The value range of cosine similarity is between -1 and 1. The value is 1 when exactly the same, -1 when the opposite is true, and 0 when orthogonal or irrelevant.

3. Conclusion

This article uses CNN-based image recognition technology to provide a new idea for the judicial practice of copyright law, as the Figure 3 shows:

![Figure 3](image)

**Figure 3.** The process of CNN-based image recognition
In the process of judging the substantial similarity of the infringing works involved, first, extract the characteristics of the plaintiff’s work, store the characteristics to form a database, and then, input the entire and part of the characteristics of the defendant’s work involved in the infringement. Retrieving the plaintiff’s work can assist in proving that there is a substantial similarity; otherwise, it does not exist. Through this method, the problems caused by the subjectivity and ambiguity of the existing "substantial similarity" standards for artwork can be resolved to a great extent.

Acknowledgments
This paper is funded by Law School, Shenzhen University. We would like to thank Professor Xianwei Zhang for his suggestions in writing this paper.

References
[1] Article 4, Paragraph 1, Item 8 of the Regulations for the Implementation of the Copyright Law
[2] Liang Zhiwen. "Typification of Fair Use of Copyright." Journal of East China University of Political Science and Law .03(2012):34-45.
[3] Xu Bo. "Determination of Copyright Protection Scope and Judgment of Substantial Similarity: From the Perspective of Historical Script Works." Intellectual Property .02(2012):28-34.
[4] Shenzhen Mengshiqi Trading Co., Ltd. v. Tianjin Ninghe County Ze'an Trading Co., Ltd. and Tianjin Ninghe County Ze'an Trading Co., Ltd., the Civil Judgment of Tianjin Supreme People’s Court (2015) Jin Gaomin San Zhong Zi No. 0018, May 22, 2015
[5] Bei Lesheng Co., Ltd. v. Guangdong Taimao Affairs Co., Ltd., etc. Infringement of reproduction rights, distribution rights, information network dissemination rights disputes, Shanghai No. 1 Intermediate People's Court Civil Judgment (2012) Huyi Zhongmin 5 (zhi) Chu Zi No. 132 Number, May 3rd, 2013
[6] Disney Enterprises Inc. v. Xiamen Blue Flame Film and Television Animation Co., Ltd. and others for copyright infringement and unfair competition disputes, Shanghai Pudong New Area People’s Court Civil Judgment (2015) Pumin San ,Chu Zi No. 1896, December 29, 2016
[7] Yan Chunman, Wang Cheng. (202). Convolutional Neural Network Model Development and Application. Computer Science and Exploration
[8] Hochreiter S, Schmidhuber J.LongShortTerm Memory[J]. Neural computation, 1997,