

# Design of a MobilNetV2-Based Retrieval System for Traditional Cultural Artworks

Zhenjiang Cao, Shandong University of Arts, China\*

Zhenhai Cao, Jingdezhen University, China

## ABSTRACT

Aiming at the problem that it is difficult for art teachers to take into account each student in the art appreciation education in colleges and universities, this paper proposes a retrieval system for traditional cultural works of art. Dense connections are used to replace residual connections between bottlenecks in MobileNetV2 network and gradient transmission in the network. The dilution factor is used to control the size of the network to solve the problem of the rapid increase in the number of network channels. In addition, the non-local attention mechanism is effectively combined with the improved MobileNetV2 network structure, which effectively improves the classification accuracy of the network. Compared with VGG16, ResNet18, and ResNet34, the classification accuracy is increased by 21.3%, 9.2%, and 3%, respectively. The method in this paper has achieved good results in the classification of art works. According to the images of art works to be appreciated, it helps students understand the relevant cultural knowledge independently and reduce the burden of teachers.

## KEYWORDS

Art Education in Higher Education, Attention Mechanisms, Deep Learning, MobilNetV2, Traditional Culture

## 1. INTRODUCTION

New media technologies, represented by computer vision, are increasingly becoming new variables that affect people's lives and work. As a new technology leading to changes in people's lives and work, artificial intelligence has the same disruptive potential in the work of education and teaching. In current Chinese university education, art education is responsible for enhancing students' aesthetic skills and passing on traditional culture, an area in which science and technology students are undoubtedly lacking compared to those majoring in art (Feng et al, 2021). Art majors have systematic professional art teaching, usually in small classes, and teachers can generally take care of every student. However, as students of science and technology, due to the different training requirements, art education in schools is usually given in the form of elective courses, mainly in the form of appreciation of art works, and the excessive number of elective courses makes it difficult to ensure that each student has a detailed understanding of the works to be appreciated in class. Utilizing computer vision technology

DOI: 10.4018/IJGCMS.334700

\*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

to categorize artworks featuring traditional culture holds immense significance in imparting associated cultural knowledge within college art education. Nonetheless, artworks embedding traditional cultural elements often feature diminutive representations of the cultural essence within the image, thereby heightening the challenge for network recognition.

Neural networks exhibit commendable performance in tackling nonlinear conundrums, exemplified by their proficiency in tasks like image recognition. However, they are not without their design challenges, including the intricacies of determining network architecture, the selection of training parameters, and the delicate calibration of network weights. These quandaries necessitate substantial human involvement in the neural network design process. Within the realm of neural network structure exploration, the parameters constituting the search space play a pivotal role in defining the network's structure (Chen et al, 2023). This encompasses critical aspects such as the number of network layers, the dimensions of convolutional kernels, and the choice of operators for each layer. The quest for an optimal network structure has emerged as a focal point in deep learning research in recent years. By employing efficient and cost-effective search algorithms, it becomes possible to automatically ascertain a neural network structure endowed with robust generalization capabilities and hardware-friendly attributes. This approach not only obviates the need for manual network design but also outperforms early, labor-intensive design paradigms across a spectrum of application scenarios (Zhong et al, 2023).

Feature extraction from images constitutes a pivotal phase in the realm of traditional image classification techniques. Researchers invariably find themselves in the position of crafting custom feature extraction methods tailored to the unique characteristics of the images within a specific classification task. A case in point is the Scale-Invariant Feature Transform (SIFT) (Kasiselvanathan, 2020), which excels in discerning object features for the purpose of object recognition. SIFT leverages the presence of salient points in an image, regardless of whether they are distant or nearby, blurred or sharp. Similarly, the Local Binary Pattern (LBP) operator (Cheng et al, 2022) is a stalwart in extracting edge information from facial images, primarily for face recognition. LBP capitalizes on the relationships between pixel gradients and their immediate neighbors to uncover valuable facial features.

For pedestrian detection, the Histogram of Oriented Gradients (HOG) feature (Yang & Wei, 2022) becomes the method of choice. HOG operates by capturing alterations in directional gradients at the juncture of the target object and its background. This approach effectively teases out the edge gradient features of the human body in images, making it indispensable for pedestrian detection tasks.

Besides the conventional image classification methods mentioned above, numerous alternative methods have delivered outstanding results in diverse image classification tasks. However, one common challenge with these methods is the necessity for manual parameter configuration during feature extraction. This reliance on manual intervention presents a formidable drawback as it does not guarantee consistent classification performance when transitioning from one image to another. In the backdrop of progressively intricate image classification tasks, it becomes increasingly evident that methods hinging on manual parameter tuning are no longer equipped to meet the demands of practical applications (Xu et al, 2023). As such, there is a pressing need for the development of more automated and adaptive feature extraction techniques to ensure robust and reliable classification results across various image datasets.

The advent of convolutional neural networks (CNNs) has ushered in a revolutionary era of remarkable advancements in image classification. This transformative progress can be attributed to the rapid evolution of deep learning, which has injected fresh vitality into the increasingly intricate landscape of image classification tasks. For instance, the VGG architecture, as exemplified by VGG(Thepade et al, 2022), leverages small convolutional kernels to significantly deepen the network. This strategic design has led to substantial improvements in performance. By virtue of this architectural innovation, VGG has not only expanded the network's capacity to extract intricate features but has also elevated its overall classification accuracy. In a similar vein, ResNet(Zhao et al, 2022) and DenseNet(Albahli et al, 2021) have introduced groundbreaking concepts in neural network

15 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/article/design-of-a-mobilnetv2-based-retrieval-system-for-traditional-cultural-artworks/334700](http://www.igi-global.com/article/design-of-a-mobilnetv2-based-retrieval-system-for-traditional-cultural-artworks/334700)

## Related Content

---

### Towards a Mobile Learning Pedagogy

Scott E. Hamm, Jason Drysdale and Diana Moore (2015). *Gamification: Concepts, Methodologies, Tools, and Applications* (pp. 2167-2185).

[www.irma-international.org/chapter/towards-a-mobile-learning-pedagogy/126165](http://www.irma-international.org/chapter/towards-a-mobile-learning-pedagogy/126165)

### Thresholds of Transmedia Storytelling: Applying Gérard Genette's Paratextual Theory to The 39 Clues Series for Young Readers

Amy Nottingham-Martin (2015). *Gamification: Concepts, Methodologies, Tools, and Applications* (pp. 826-851).

[www.irma-international.org/chapter/thresholds-of-transmedia-storytelling/126091](http://www.irma-international.org/chapter/thresholds-of-transmedia-storytelling/126091)

### Videogames and the Ethics of Care

John Murphy and José Zagal (2011). *International Journal of Gaming and Computer-Mediated Simulations* (pp. 69-81).

[www.irma-international.org/article/videogames-ethics-care/56339](http://www.irma-international.org/article/videogames-ethics-care/56339)

### The Fall of the Fourth Wall: Designing and Evaluating Interactive Spectator Experiences

Samantha Stahlke, James Robb and Pejman Mirza-Babaei (2018). *International Journal of Gaming and Computer-Mediated Simulations* (pp. 42-62).

[www.irma-international.org/article/the-fall-of-the-fourth-wall/210207](http://www.irma-international.org/article/the-fall-of-the-fourth-wall/210207)

### Book Review: The Esports Education Playbook Empowering Every Learner Through Inclusive Gaming

Alvaro Brito (2022). *International Journal of eSports Multidisciplinary Research* (pp. 1-4).

[www.irma-international.org/article/book-review/302077](http://www.irma-international.org/article/book-review/302077)