2023 【海報論文】 □一般組 ☑研究生組 □大學組
11th 台灣 TAIWAN 工業與應用數學會年會 TWSIAM Annual Meeting

Image Style Transfer Using Convolutional Neural Networks With Compressed Data

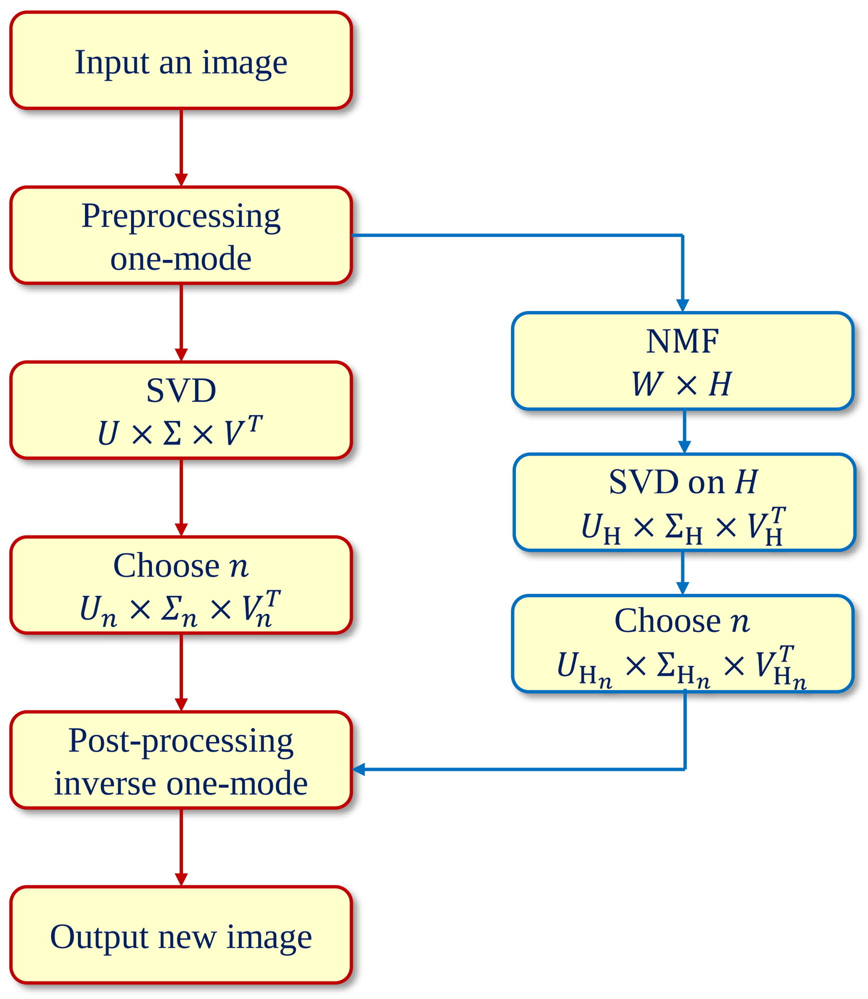
Pan, Yi-Ting (潘薏婷), *Department of Applied Mathematics, National Sun Yat-sen University, Taiwan*
(honeylydia8827@g-mail.nsysu.edu.tw)
Advisor: Prof. Tsung-Lin Lee (李宗錂)

Abstract

In recent years, the problem of style transfer has become increasingly complex. Due to the mixture of too many styles, the original content image often becomes blurry and requires precise processing. Additionally, the data of color images is relatively large, resulting in various issues in calculation and storage, such as slow convergence and large space occupation. Therefore, this paper aims to explore how to use compression technology for style transfer, in order to achieve more efficient and precise image processing. At the same time, we also explore how to make the input content image of a person more striking, thereby improving the effectiveness of image processing.

Style Transfer

This study utilizes SVD and NMF decomposition for image data compression to reduce storage space demand. The com-pressed data is then used as input to the model to identify Feature maps for style transformation. The total loss is employed as the training objective of the model, which calculates individual loss to minimize the total loss and generates a new style transformed image. This approach effectively preserves human contours while reducing image size, improving processing efficiency.



Algorithm(Style transfer):

Input: style image S, content image C, and model M

Output: new image Y

Step 1 (Initialization):

Define the style layers S_{layer} in M,
And the content layers C_{layer} in M

Set $Y = C$

Step 2 (Feature extraction):

Extract the style features $S_{feature}$ from M by S_{layer},
and the content features $C_{feature}$ from M by C_{layer}

Step 3 (Optimization):

Find loss L_{total} for training the model

for $i = 1, 2, \ldots$
    $C_{feature} = M(C)$
    $L_{style} = L_{content} = 0$

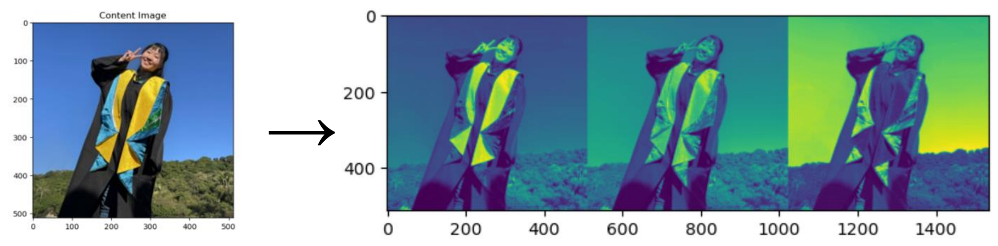
    for j in S_{layer}
        $C_{style_{gram}} = C_{feature}[j] \times C_{feature}[j]^{\mathrm{T}}$
        $S_{style_{gram}} = S_{feature}[j] \times S_{feature}[j]^T$
        $L_{style} = \left(C_{style_{gram}}[j] - S_{style_{gram}}[j]\right)^2$

    for k in C_{layer}
        $L_{content} = \left(C_{feature}[k] - S_{feature}[k]\right)^2$

    $L_{total} = \alpha L_{style} + \beta L_{content}$
    $Y = Y - \lambda \times \nabla Y L_{total}$

Compression

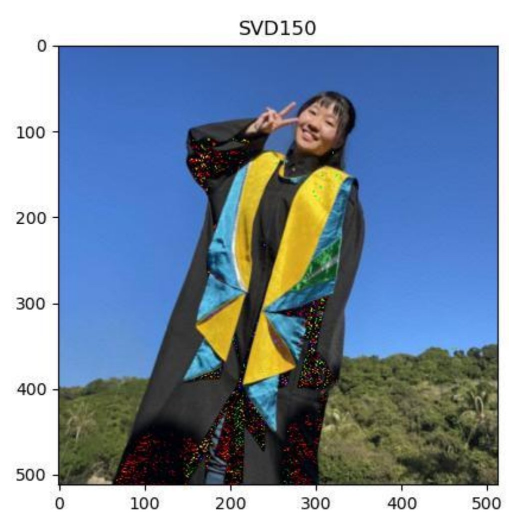
In compressing image data, we utilize SVD or NMF decomposition to reduce the dimensiona-lity of the content image. Since these methods can only handle matrices, the original three-dimensional tensor needs to be converted into matrix form before the decom-position process can commence. To achieve this, we use the one-mode method to transform the three-dimensional tensor into a flatter matrix.



$$X \longrightarrow X_{(1)} = [X_{::1}\ X_{::2}\ X_{::3}]$$

HOSVD

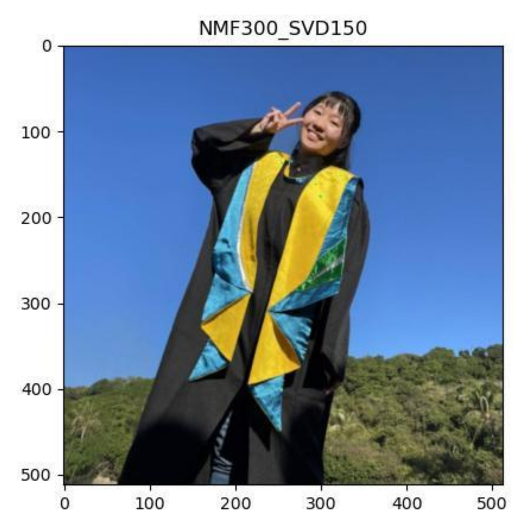
Performing data compression using SVD.
Let $X_{(1)} = U \times \Sigma \times V^T$ be given,
then we choose n singulars from $X_{(1)}$.
Therefore, we can get $U_n \times \Sigma_n \times V_n^T = [Y_{::1}\ Y_{::2}\ Y_{::3}]$.
We obtain a new matrix $Y_{(1)}$ and perform inverse one-mode to transform it back into a new tensor image Y.



Performing HOSVD with n = 150, the singular values account for 91.91% of the total variance, and MSE between the processed con-tent and the original content is 0.0001.

HONMF+SVD

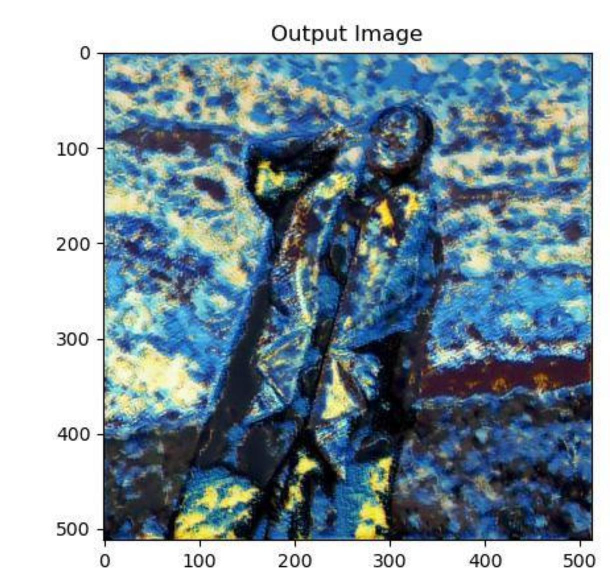
Performing data compression using HONMF.
Let $X_{(1)} = W_r \times H_r$ be given,
then we choose n singulars from H_r.
Therefore, we can get $U_n \times \Sigma_n \times V_n^T = H_r'$.
We obtain a new matrix $Y_{(1)} = W_r \times H_r'$ and perform inverse one-mode to transform it back into a new tensor image Y.



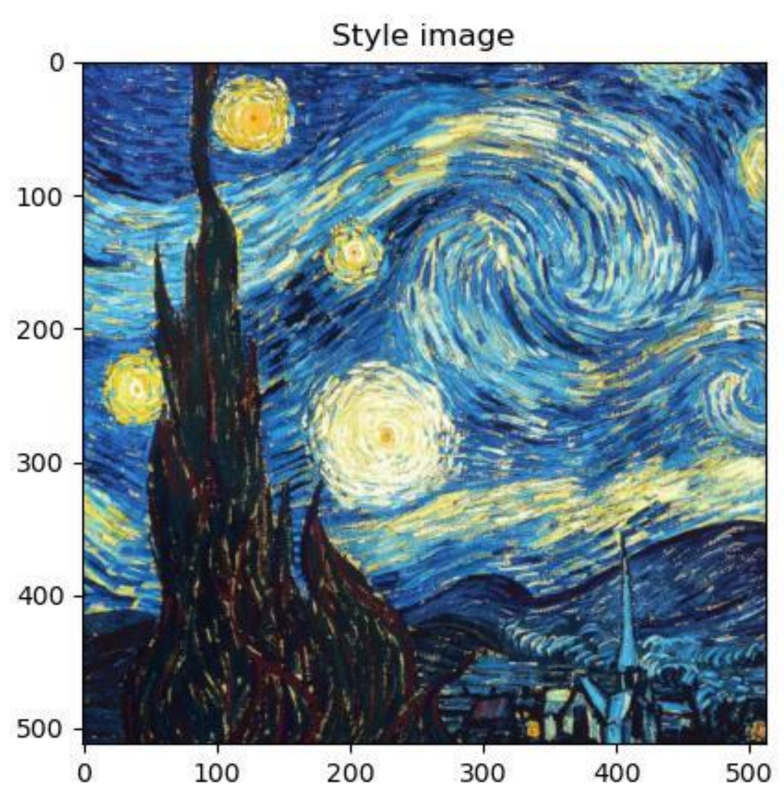
Performing HONMF +SVD with r=300 and n = 150, the sin-gular values account for 92.1% of the total variance, and MSE between the processed con-tent and the original content is 0.0001.

Experiment Results

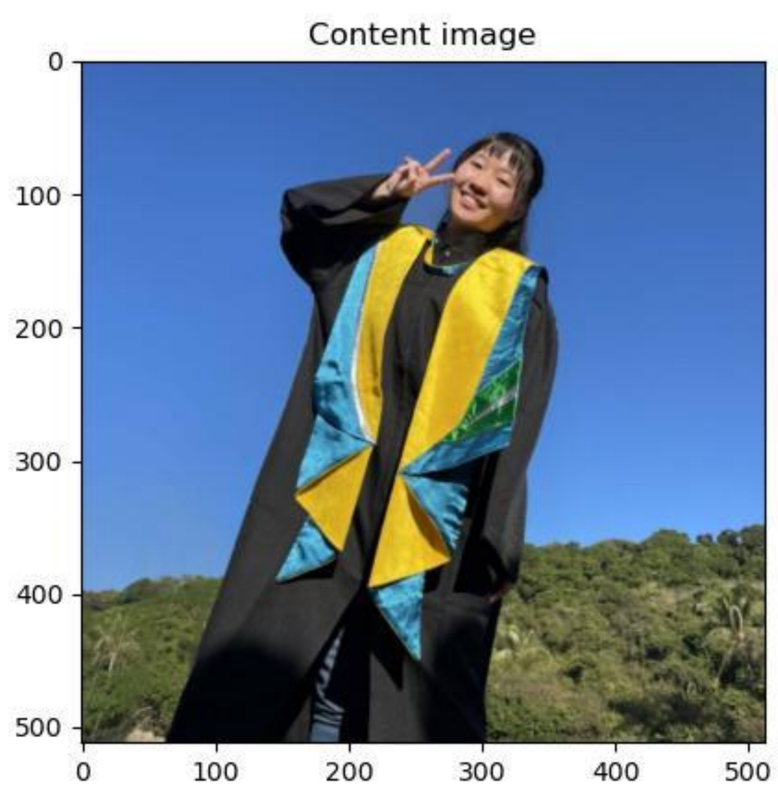
Upper right image: The image shown in the upper right corner represents the outcome achieved by employing HOSVD with n = 150. In comparison to the uncompressed image, it is apparent that the compressed image exhibits significantly enhanced clarity and distinctiveness in terms of facial features and contours.
Lower right image: The image depicted in the lower right corner corresponds to the application of HOSVD with n = 150. Compared to the image solely subjected to HOSVD, it is apparent that the compressed image enables clearer visibility of facial features and contours.

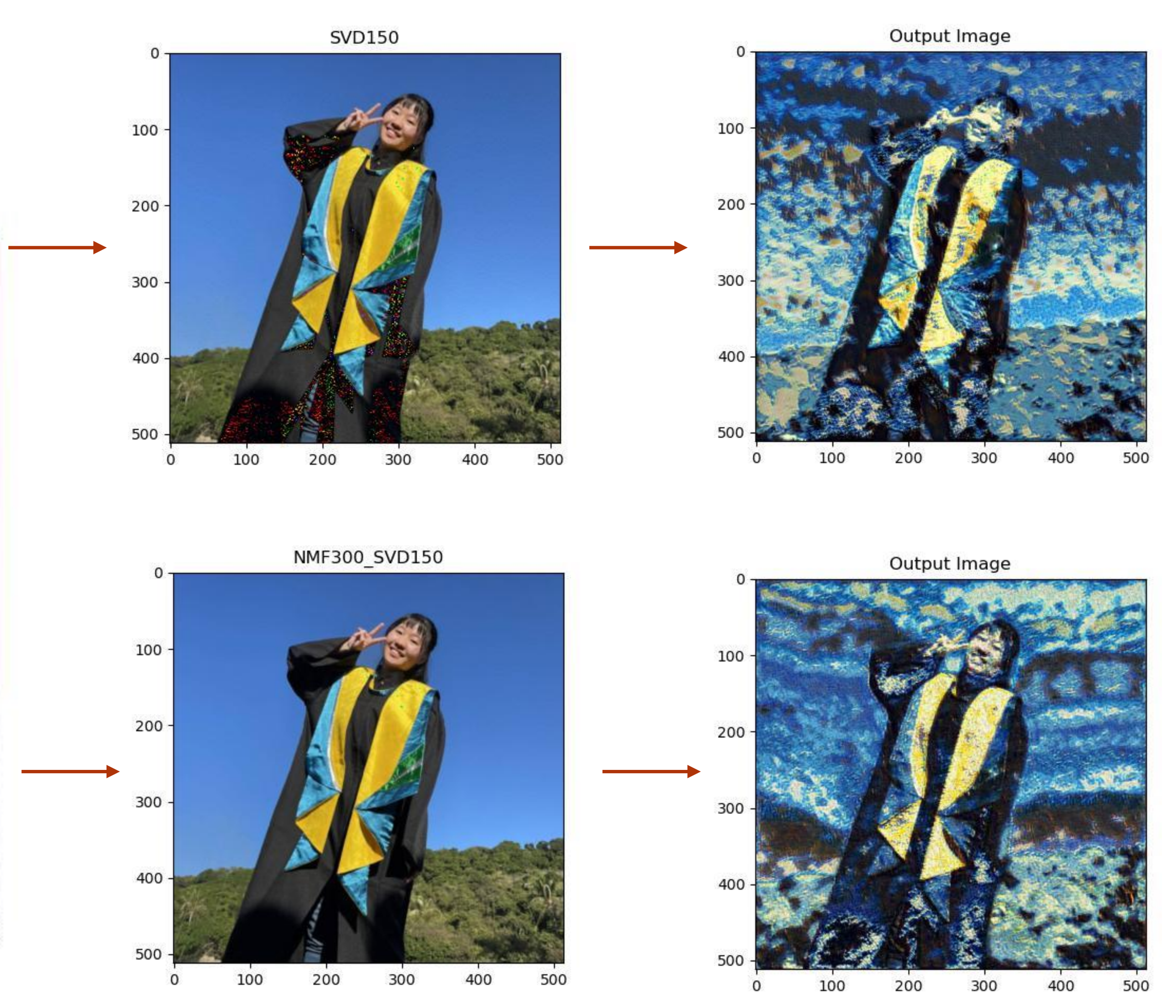


Upper image: The image presented in the upper right corner is obtained by directly inputting the data into the model's output without utilizing any data compression. As a result, the contours of the person's face are not clearly discernible in this uncompressed image.



+





Conclusion

Based on the observed experimental results, it is evident that directly processing uncompressed images with the model does not produce satisfactory outcomes, as the output fails to clearly distinguish the contours of the individuals. However, both of the employed methods show significant improvements compared to the uncompressed images, allowing for a clear representation of the facial contours rather than presenting unrecognizable stylistic effects. In particular, when applying the HONMF+SVD method, the improvement effect is even more pronounced, enabling better capture of fine details in the images and resulting in sharper and clearer delineation of the facial contours, while also performing well in removing stylistic effects.

References

[1] Leon A. Gatys, Alexander S. Ecker and Matthias Bethge, Image Style Transfer Using Convolutional Neural Networks, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.

[2] Dumoulin, Vincent and Shlens, Jonathon and Kudlur, Manjunath, A learned representation for artistic style, arXiv preprint arXiv:1610.07629, 2016.

[3] Hoecker, Andreas and Kartvelishvili, Vakhtang, SVD approach to data unfolding, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,1996.

[4] Feng Gao, Xiaopeng Liu, Junyu Dong, Guoqiang Zhong and Muwei Jian, Change Detection in SAR Images Based on Deep Semi-NMF and SVD Networks, Remote Sensing, 2017.