Semnan University

Journal of Modeling in Engineering

Journal homepage: https://modelling.semnan.ac.ir/

ISSN: 2783-2538

Research Article

Combining Style Transfer in a Frame with Random Rotation Using Convolutional Neural Networks

Mahdi Musa Semnani [a], Kourosh Kiani [b,*]

[a] *MSc Student, Department Electrical and Computer Engineering, Semnan University, Semnan, Iran*
[b] *Associate Professor, Department Electrical and Computer Engineering, Semnan University, Semnan, Iran*

PAPER INFO

Paper history:

Received: 14 April 2024
Revised: 22 May 2024
Accepted: 05 June 2024

Keywords:

Style transfer,
Random rotation,
Deep learning,
Convolutional Neural Networks (CNN).

ABSTRACT

Style transfer is a research area that has garnered significant attention. This technology allows the style of one image to be transferred into the content of another image. Extensive research has been conducted in the field of style transfer, aiming to accelerate processing and produce beautiful and high-quality images. One application of this technology could be in generating beautiful designs for printing in industries such as tile and carpet manufacturing, where creating such intricate patterns by human hands would be challenging. In this article, a simple method is proposed for transferring multiple styles that have been randomly rotated, resulting in more unique designs compared to using a single style for transfer. This method can enhance the diversity and quality of the generated images. Evaluation of this method was conducted by surveying 25 individuals who compared our method to other existing methods. Ultimately, our method received the highest approval rating. The outcome of this research facilitates the creation of artworks more efficiently, offering a viable alternative to human-designed patterns.

DOI: https://doi.org/10.22075/jme.2024.33791.2653

© 2024 Published by Semnan University Press.
This is an open access article under the CC-BY 4.0 license.(https://creativecommons.org/licenses/by/4.0/)

*** Corresponding Author.**
E-mail address: kourosh.kiani@semnan.ac.ir

How to cite this article:
Musa Semnani, M., & kiani, K. (2024). Combining Style Transfer in a Frame with Random Rotation Using Convolutional Neural Networks. Journal of Modeling in Engineering, 22(78), 275-281. doi: 10.22075/jme.2024.33791.2653

مقاله پژوهشی

انتقال سبک ترکیبی در یک قاب با دوران تصادفی با استفاده از شبکه‌های کانولوشنی

مهدی موسی سمنانی[۱] ، کوروش کیانی[۲،*]

اطلاعات مقاله

دریافت مقاله: ۱۴۰۳/۰۱/۲۶
بازنگری مقاله: ۱۴۰۳/۰۳/۰۲
پذیرش مقاله: ۱۴۰۳/۰۳/۱۶

واژگان کلیدی:
انتقال سبک،
دوران تصادفی،
یادگیری عمیق،
شبکه‌های کانولوشنی عمیق.

چکیده

انتقال سبک یک حوزه تحقیقاتی است که توجه فراوانی را به خود جلب کرده است. این فناوری اجازه می‌دهد تا سبک یک تصویر را بر روی تصویر محتوا منتقل کنیم. تحقیقات گسترده‌ای در زمینه انتقال سبک انجام شده است، هدف از آن تسریع در پردازش و یا تولید تصاویر زیبا و با کیفیت بالا است. یکی از کاربردهای این حوزه می‌تواند در زمینه تولید طرح‌های زیبا برای چاپ در طراحی کاشی، فرش و هر صنعتی که به دنبال چاپ طرح‌هایی است، که بشر توانایی خلق آن را ندارد. در این مقاله یک روش ساده برای انتقال سبک‌های چندگانه که به طور تصادفی دوران پیدا کرده‌اند، پیشنهاد می‌شود که منجر به تولید طرح‌های خاص‌تر نسبت به استفاده از یک سبک برای انتقال می‌شود. این روش می‌تواند باعث ایجاد تنوع و کیفیت بالاتری در تصاویر تولید شده گردد، که ارزیابی آن توسط ۲۵ نفر و نظرسنجی از ایشان برای مقایسه بین روش ما و روش‌های دیگر صورت گرفته است، که در نهایت روش ما بیشترین رای موافق را دریافت نمود. حاصل این پژوهش تسریع در امر خلق آثار را پدید می‌آورد که می‌تواند جایگزین خوبی برای طراحی‌های انسان باشد.

DOI: https://doi.org/10.22075/jme.2024.33791.2653

© 2024 Published by Semnan University Press.
This is an open access article under the CC-BY 4.0 license.(https://creativecommons.org/licenses/by/4.0/)

۱- مقدمه[۱]

انتقال سبک، تکنیکی در زمینه تولید هنر کامپیوتری است که هدف آن بازآفرینی یک تصویر محتوایی (Content Image) با استفاده از یک تصویر مرجع به نام سبک (style image) می‌باشد. این تکنیک کاربرد گسترده‌ای در خلق آثار هنری با کمک کامپیوتر دارد. شبکه‌های عصبی عمیق نه تنها محتوا بلکه اطلاعات سبک تصویر را نیز رمزگذاری می‌کنند [۲]. علاوه بر این، سبک و محتوای تصویر تا حدودی قابل جدا شدن هستند. روش های انتقال سبک انعطاف پذیر بوده و قابلیت تعمیم به تصاویر دلخواه را نیز دارند. کار اصلی گاتیس و همکاران [۱] نشان داد که همبستگی بین ویژگی‌های کدگذاری شده توسط یک شبکه عصبی کانولوشنی عمیق [۳] می‌تواند الگوهای سبک را به خوبی ثبت کند. روش های انتقال سبک مطرح شده نیز فرض می‌کنند که می‌توان ویژگی‌های استخراج شده از شبکه کانولوشنی عمیق را در فرم ماتریس گرام [۱] برای محاسبه تابع هزینه مورد استفاده قرار داد. گرچه در برخی از شبکه‌ها به جای استفاده از ماتریس گرام از روابط دیگری مانند همبستگی پیرسون، کوواریانس، فاصله اقلیدسی، شباهت کسینوسی نیز استفاده شده است [۵]. این مدل‌ها را مدل‌های یک مدل، یک سبک می‌نامند زیرا مدل آموخته شده تنها برای یک سبک خاص ترکیب می‌شود. روش های

* پست الکترونیک نویسنده مسئول: kourosh.kiani@semnan.ac.ir
۱. دانشجوی کارشناسی ارشد، دانشکده برق و کامپیوتر، دانشگاه سمنان، سمنان، ایران
۲. دانشیار، دانشکده برق و کامپیوتر، دانشگاه سمنان، سمنان، ایران

استناد به این مقاله:
موسی سمنانی، مهدی، و کیانی، کورش. (۱۴۰۳). انتقال سبک ترکیبی در یک قاب با دوران تصادفی با استفاده از شبکه های کانولوشنی. مدل سازی در مهندسی، ۲۲(۷۸)، ۲۷۵-۲۸۱. doi: 10.22075/jme.2024.33791.2653

انتقال سبک مبتنی بر شبکه‌های کانولوشنی اکنون می توانند تصاویر تقلیدی با کیفیت بالا تولید کنند. با این حال، این روش ها بر انتقال تصویر اصلی به سبک ارائه شده توسط یک تصویر سبک دیگر (معمولاً یک نقاشی)، تمرکز می کنند. در عمل، زمانی که کاربر از یک منظره زیبا عکس می‌گیرد، ممکن است امیدوار باشد که آن را دوباره بر روی بوم ارائه کند، به طوری که به نظر برسد که توسط یک هنرمند نقاشی شده است [۴]. این تحقیق منجر به یک ایده نوآورانه جهت انتقال چند سبک برروی یک تصویر شد که پیش‌تر در تحقیقات برخی پیگیری شده بود.

۱. ایجاد تنوع و خلق آثاری با استفاده از چند سبک با دوران تصادفی هرسبک

۲. تغییر مقیاس تصاویرو چیدن آن‌ها در کنار هم به عنوان یک تصویر سبک

۳. بررسی کارایی انتقال سبک در طراحی کاشی

۴. توانایی انتقال سبک برروی تصاویر بدون وابستگی به رزولوشن خاص

۲- پژوهش‌های پیشین

گاتیس و همکاران [۱] یک الگوریتم عصبی سبک هنری را معرفی می‌کنند که می‌تواند محتوای تصویر و سبک تصاویر طبیعی را دوباره ترکیب کند. این الگوریتم اجازه می‌دهد تا تصاویر جدیدی با کیفیت ادراکی بالا تولید کنیم که محتوای یک عکس دلخواه را با ظاهر آثار هنری شناخته شده متعدد ترکیب می‌کند. نتایج آن‌ها بینش‌های جدیدی را در مورد بازنمایی‌های تصویر عمیق آموخته شده توسط شبکه‌های عصبی کانولوشنی ارائه می‌کند و پتانسیل آن‌ها را برای سنتز و دستکاری تصویر سطح بالا نشان می‌دهد.

دی یونگ و همکاران [۶] نیز در تحقیقی یک شبکه جدید سبک توجه (SANet) را معرفی می‌کنند که به طور کارآمد و انعطاف‌پذیر الگوهای سبک محلی را با توجه به توزیع فضایی معنایی تصویر محتوا یکپارچه می‌کند. یک تابع جدید از دست دادن هویت و تعبیه ویژگی‌های چند سطحی، (SANet) و رمزگشا را قادر می‌سازد تا ساختار محتوای مشارکت یکسان را تا حد امکان حفظ کند و در عین حال الگوهای سبک را غنی‌تر کند.

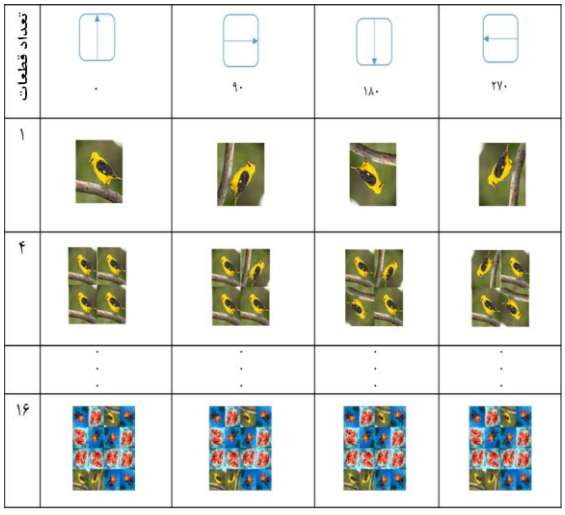
چن و همکارانش [۴] نیز با استفاده از شبکه های مولد تخاصمی کار انتقال سبک در تصاویر را انجام داده‌اند، به این گونه که آن‌ها در مقاله خود از دو شبکه مولد و متمایز کننده استفاده می‌کنند که نوع جدیدی از انتقال سبک است. به عبارت دیگر، شبکه آموزش می‌بیند به جای آن که تصویر محتوا را تغییر دهد، تصاویر در استایل خاص را تولید کند.

ژان هانگ و همکارانش [۲] مدل عادی سازی نمونه تطبیقی (AdaIN) را ارائه نموده‌اند. در این مدل با پارامترها، میانگین و واریانس، محتوا را با توجه به ورودی سبک تنظیم می‌کنند. از طریق آزمایش‌ها، متوجه می‌شویم که AdaIN به طور مؤثری محتوا و سبک را با انتقال آمار ویژگی‌ها ترکیب می‌کند. سپس یک شبکه رمزگشا، می‌آموزد که با برگرداندن خروجی AdaIN به فضای تصویر، تصویر استایل شده نهایی را تولید نماید. این روش به طور حدودی، سه مرتبه سریع‌تر از [۱] است.

سون و همکارانش [۷] روش چرخش نقشه ویژگی‌ها در زوایای مختلف به نام چرخش ویژگی عمیق (DFR) را معرفی می‌کنند. انگیزه اساسی ایجاد ویژگی‌های جدید از ویژگی‌های استخراج شده از VGG19 [۳] است که به خوبی آموزش دیده است، در نتیجه تغییر در سبک سازی‌ها را مشاهده می‌کنیم. در این میان، راه‌های زیادی برای تبدیل ویژگی‌های میانی در انتقال سبک وجود دارد. با این حال، آثار آن‌ها تنها یک یا چند خروجی از یک جفت تصویر محتوا و سبک تولید می کند.

۳- روش پیشنهادی

در مرحله اول، برای آماده‌سازی سبک برای ورودی شبکه عمیق کانولوشنی، سبک‌های مختلف با چرخش‌های تصادفی در یک تصویر قاب ترکیب می‌گردد.

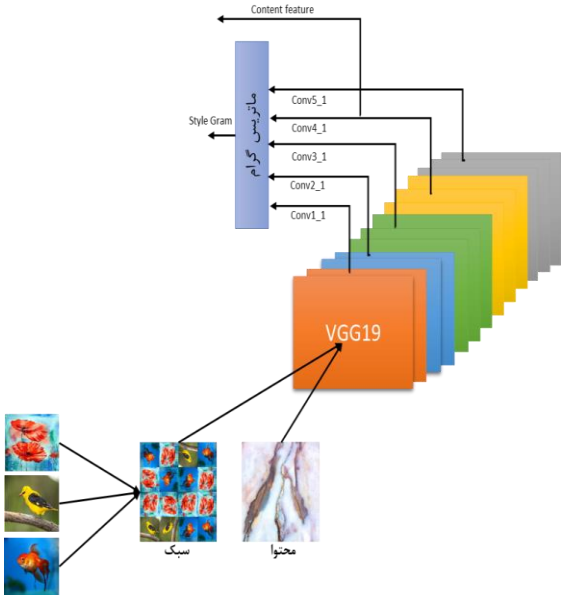


شکل ۱- روش‌های مختلف آماده‌سازی سبک برای ورودی مدل یادگیری

شکل (۱) تصویر قاب‌های تکی، چهارتایی و شانزده تایی را نشان می‌دهد. هر قطعه از این تصویر قاب، یک سبک

کوچک شده با چرخش می‌باشد. تصویر قاب مورد نظر همیشه باید مجذور اعداد طبیعی باشد، در غیر این صورت تصویر قاب کاملی نخواهیم داشت و قطعه‌ای از این تصویر قاب، خالی خواهد ماند.

تعداد سبک‌ها می‌تواند برابر یا کوچک‌تر با تعداد قطعات باشد. پس از آماده سازی تصویر قاب سبک‌ها، آن را به شبکه کانولوشنی می‌دهیم. نقشه ویژگی ۵ لایه را گرفته و ۵ ماتریس گرام برای قاب سبک‌ها محاسبه نموده و آن را ذخیره می‌نماییم. شکل (۲) این شبکه کانولوشنی را نشان می‌دهد. در این مدل شبکه از قبل آموزش دیده VGG19 [۳] برای استخراج نقشه ویژگی‌ها مورد استفاده قرار می‌گیرد.



شکل۲- تولید نقشه ویژگی‌های سبک و محتوا

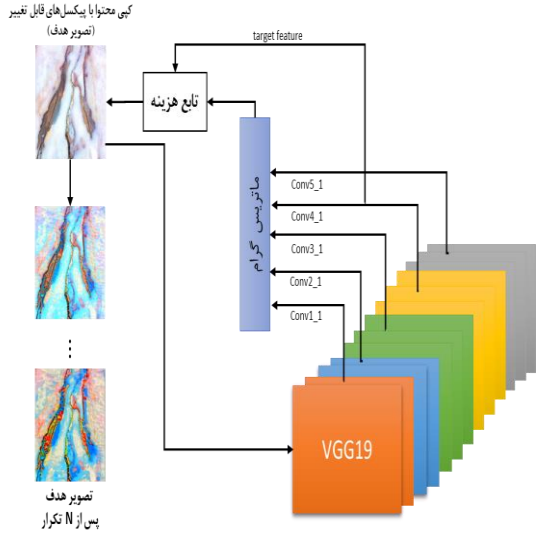
سپس ورودی محتوا را به طور جداگانه به این شبکه کانولوشنی از قبل آموزش دیده می‌دهیم و مانند تصویر قاب سبک‌ها، نقشه ویژگی‌های ۵ لایه را گرفته و ذخیره می‌نماییم.

معادله زیر محاسبه گرام هر لایه را نشان می‌دهد:

$$[G^l(F^l)]_{ij} = \sum_k F^l_{ik} F^l_{jk} \qquad (۱)$$

$F^l(*)$ نقشه ویژگی هر لایه می‌باشد. * می‌تواند ورودی تصویر سبک I_s یا محتوا I_c باشد. اطلاعات نقشه ویژگی‌های محتوا در تصویر هدف کپی می‌گردد تا در تکرارهای بعدی وارد شبکه شکل (۳) گردد.

(۴)



شکل۳- ادامه روند آموزش شبکه

معادله زیر محاسبه گرام هر لایه را نشان می‌دهد:

$$L_{content}(contect\ , target) = \frac{1}{2}||F^{l=conv4_1}\ (content) - F^{l=conv4_1}\ (target)||_2^2 \qquad (۲)$$

در معادله شماره ۲ عبارت $(F^{l=conv4_1}\ (content))$ که نمایش نقشه ویژگی لایه‌های محتوا می‌باشد، فقط یک بار محاسبه می‌گردد (شکل ۲). ولی عبارت $F^{l=conv4_1}\ (target)$ که نماینده نقشه ویژگی‌های هدف می‌باشد در هر تکرار به روز می‌گردد.

$$L_{style}(target\ , style\ ) = \sum_{l=1}^{L} \frac{\omega_l}{4D_l^2 M_l^2}\ ||G^l(F^l\ (\text{target})) - G^l\ (F^l\ (\text{style}))||_2^2 \qquad (۳)$$

در معادله شماره ۳ عبارت $(G^l\ (F^l\ (style)))$ که نمایش ماتریس گرام لایه‌های سبک می‌باشد فقط یک بار محاسبه می‌گردد (شکل ۲). ولی عبارت $(G^l(F^l\ (target)))$ که نماینده ماتریس گرام لایه‌های هدف می‌باشد، در هر تکرار به روز می‌گردد.

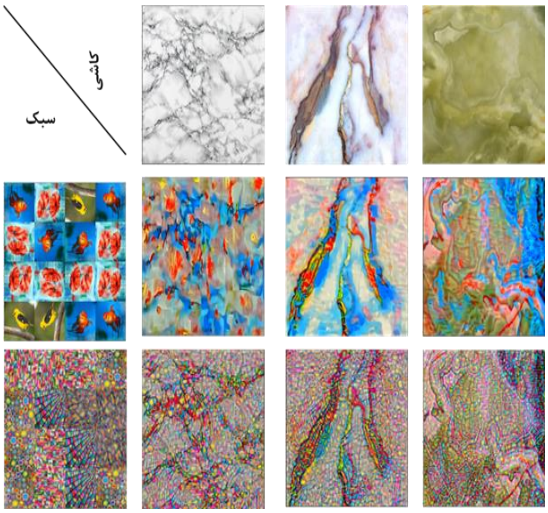
همچنین D_l تعداد نقشه ویژگی‌های هر لایه و M_l حاصل ضرب طول و عرض نقشه ویژگی است. ($M_l = H \times W$) و ω_l وزن تاثیر گذاری هرلایه می‌باشد.

در محاسبه هزینه کل، برای تاثیر سبک و محتوا وزن‌های α و β در نظر گرفته شده است.

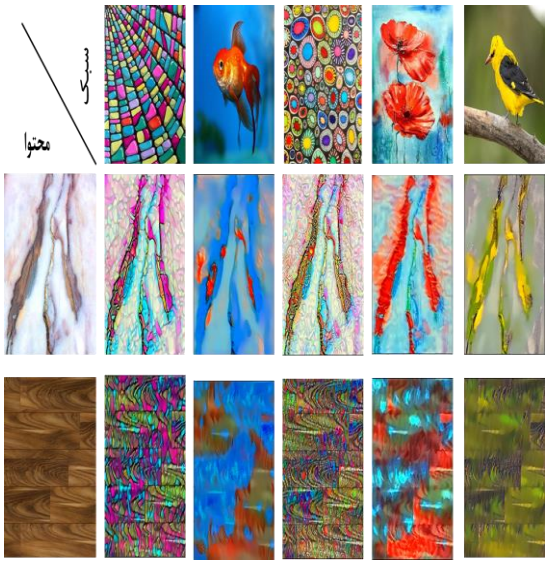
$$L_{total} = \alpha\, L_{content} + \beta L_{style}$$

۴- ارزیابی

در پژوهش‌هایی با محوریت انتقال سبک که خروجی مدل یک تصویر است، معیارهای ارزیابی عددی قابل استنادی وجود ندارد. بدین منظور برای سنجش خروجی‌های انتقال سبک از تکنیک نظرسنجی استفاده شده است. در شکل (۴) نمونه‌ای از تصویر طرح‌های کاشی، سبک‌های ترکیبی و تصویر هدف نشان داده شده است.

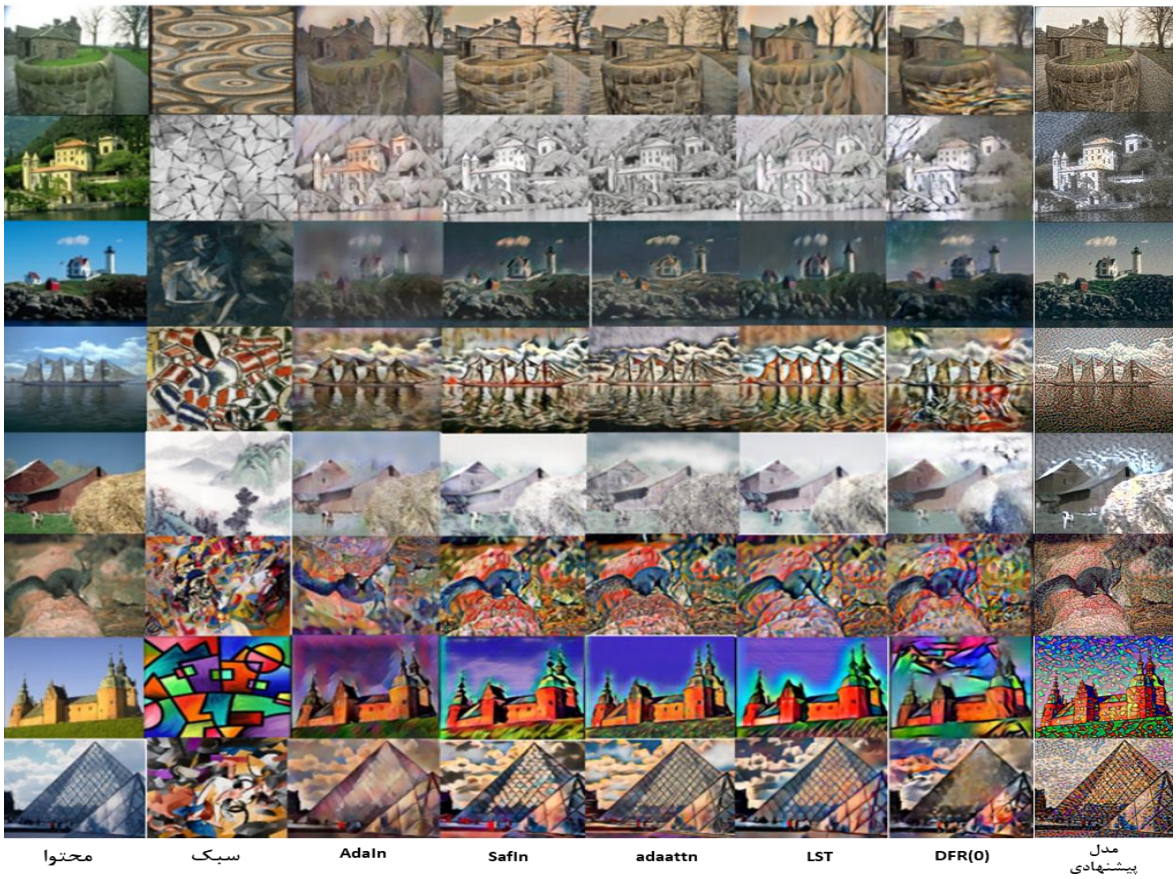


شکل۴- نمونه‌ای از خروجی‌های مدل در طراحی کاشی

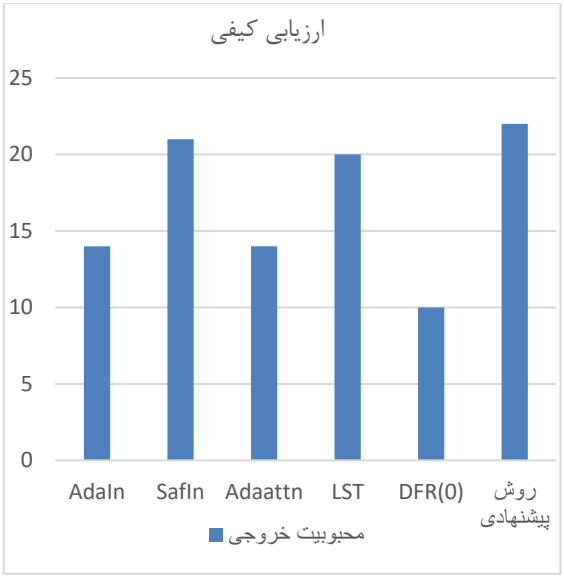


شکل ۵ - تولید طرح با یک سبک

همچنین خروجی مدل خود را که از کنار هم قرار دادن ۱۶ قطعه یکسان از تصویر سبک‌های موجود در شکل شماره (۵) تولید شده است. در شکل (۶) خروجی مدل ما با خروجی مدل های Adain [۲]، Safin [۹]، Adaattn [۱۰]، LST [۱۱]، DFR [۷] با تصاویر سبک و محتوای یکسان مقایسه شده است.



شکل ۶- مقایسه خروجی مدل ما با دیگر مدل‌های مربوط در این حوزه



شکل ۷- نتایج نظر سنجی برای ارزیابی کیفی خروجی‌های مدل

در این پژوهش ما از ۲۵ نفر از افراد متشکل از دانشجویان و اساتید رشته‌های هنری و رشته‌های غیر مرتبط جهت انتخاب بهترین خروجی نظرسنجی به عمل آوردیم. ما خروجی های مدل های مختلف را بدون ذکر مرجع و مدل آن به افراد شرکت کننده در این نظر سنجی نشان دادیم و از آن‌ها خواسته شد که تصویر مورد علاقه خود را انتخاب کنند. پس از جمع‌آوری کلیه نظرها نتایج نظرسنجی مربوطه در نمودار شکل (۷) قابل مشاهده است.

۵- نتیجه‌گیری

در این پژوهش، ما یک انتقال سبک ترکیبی با چرخش تصادفی در هر قطعه را پیشنهاد کرده‌ایم، این سبک ترکیبی با رزولوشن‌های بالای تصویر محتوا و سبک‌های متنوع ویژگی‌های محلی بهتری به دست می‌آید و ساختار محتوا کم‌تر مخدوش می‌گردد. روش پیشنهادی ما باعث ایجاد تنوع بسیار زیادی در تولید تصاویر خروجی می‌شود، این طرح‌ها می‌تواند درصنایع کاشی‌سازی، قالی‌بافی، نساجی و تمامی صنایع مشابه مورد استفاده قرار بگیرد. لازم به ذکر است انسان در زمینه تولید تصاویر هنری برای کاربردهای نام برده شده، نمی‌تواند تنوعی مانند انتقال سبک هوشمند را خلق نماید و در نتیجه این روش، روند طراحی نقشه‌های چاپ را تسریع می‌کند. در نهایت می‌توان گفت روش پیشنهادی ما هم برای انتقال تک سبک و هم برای انتقال چند سبکی کاربردی خواهد بود.

برای پژوهش‌های آینده پیشنهاد می‌گردد، از شبکه‌های زایشی نظیر شبکه‌های Diffusion و یا شبکه‌های تخاصمی مولد استفاده شود.

مراجع

[1] L.A. Gatys, A.S. Ecker, and M. Bethge. "Image style transfer using convolutional neural networks." *In Proceedings of the IEEE Conference On Computer Vision and Pattern Recognition*, pp. 2414-2423. 2016.

[2] X. Huang, and S. Belongie. "Arbitrary style transfer in real-time with adaptive instance normalization." *In Proceedings of the IEEE International Conference on Computer Vision*, pp. 1501-1510. 2017.

[3] K. Simonyan, and A. Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[4] X. Chen, C. Xu, X. Yang, L. Song, and D. Tao. "Gated-gan: Adversarial gated networks for multi-collection style transfer." *IEEE Transactions on Image Processing* 28, no. 2 (2018): 546-560.

[5] N.Q. Tuyen, S.T. Nguyen, T.J. Choi, and V.Q. Dinh. "Deep correlation multimodal neural style transfer." *Ieee Access* 9 (2021): 141329-141338.

[6] D.Y. Park, and K.H. Lee. "Arbitrary style transfer with style-attentional networks." *In proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, pp. 5880-5888. 2019.

[7] S. Truong Nguyen, N. Quang Tuyen, and N. Hong Phuc. "Deep Feature Rotation for Multimodal Image Style Transfer." *arXiv e-prints* (2022): arXiv-2202.

[8] C. Li, and M. Wand. "Combining markov random fields and convolutional neural networks for image synthesis." *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2479-2486. 2016.

[9] A. Singh, S. Hingane, X. Gong, and Z. Wang. "SAFIN: arbitrary style transfer with self-attentive factorized instance normalization." *In 2021 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1-6. IEEE, 2021.

[10] S. Liu, T. Lin, D. He, F. Li, M. Wang, X. Li, Z. Sun, Q. Li, and E. Ding. "Adaattn: Revisit attention mechanism in arbitrary neural style transfer." *In Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6649-6658. 2021.

[11] X. Li, S. Liu, J. Kautz, and M.H. Yang. "Learning linear transformations for fast arbitrary style transfer." *arXiv preprint arXiv:1808.04537* (2018).