Bridging Art and Fashion: AI-Driven Techniques for Seamless Design Integration

Rishita Mantri, Prof. Sofia Francis
Computer Engineering Department
Mukesh Patel School of Technology Management and Engineering,
SVKM's Narsee Monjee Institute of Management Studies (NMIMS) Deemed-to-be University, Mumbai-400056, India
rishita.mantri75@nmims.in, sofia.francis@nmims.edu

***Abstract*— This research explores the successful integration of artwork onto T-shirts using deep learning models, in particular ResNet18 and Pix2Pix, which allow for seamless integration of artwork onto T-shirts. That is, using the extracted feature capabilities of ResNet18, this model can retain and save the texture and contours of garments for integrity in the blending process. The Pix2Pix model, more specifically an image-to-image translation model, allows for it to map artwork properly onto the T-shirt and thus result in visually coherent and high-quality outputs. It contributes to the automation of fashion design with minimized manual effort, zero wastage, and rapid prototyping. The comparative advantage of the two models is depicted through quantitative measures, such as MSE and inference time. MSE for the former is 0.0486, but its speed is faster than that of Pix2Pix. Results vindicate AI-based approaches in fashion design and form a trajectory toward sustainable practices. It goes further to talk about the implications for the fashion industry and to discuss future extensions and broader applications of the methodology.**

Keywords*— Fashion Design, Deep Learning, Resnet18, Pix2Pix, Blending.

I. INTRODUCTION

The fashion industry, as gigantic as it is and with its constant innovating, is still far from being able to function without concerns over too much waste, outdated designing methods, and the quick and ever-growing demand to tailor apparel according to exact customer needs. Artificial intelligence has transformed many industries but is still not fully utilized in fashion. The traditional way of doing things quite often fails to satisfy these demands, so challenges open opportunities for new solutions. From this study, an attempt was made to explore how AI can facilitate customization, reduce reliance on manual labor, and accelerate the design process, while merging technology and fashion. To address these critical issues, we start with leveling on the wave of AI tools that can introduce solutions not only efficiency-enhancing but also leading to sustainability practice in the business.

To integrate art onto the T-shirt, and to create beautiful designs we evaluated two different models in this study Resnet18 and Pix2Pix. The first model is to have used a deep residual learning structure, which is a Convolutional Neural Network (CNN) model [1]. ResNet18 successfully overcomes the vanishing gradient problem existing in deep networks [2]. This is achieved using ResNet18 to extract complex features from the T-shirt so that the texture and shape of the clothing are preserved as blending occurs to keep the garment's structure accurate.

The second used model is Pix2Pix, which happens to be a Generative Adversarial Network (GAN) especially suited for image-to-image translation tasks [3]. Using the Pix2Pix model, a discriminator and a generator are merged together in such a way that it provides photorealistic, high-quality images. For this project, it is essential to make use of Pix2Pix to transfer the artwork to the T-shirt in such an integrated manner that all original features of the garment, such as the color or texture, become an integral part of it. Together, these models produce visually appealing T-shirts with seamlessly integrated artwork.

Literature Survey

A number of models have tried to bypass the difficulty of transplanting art onto clothing. Fast Neural Style Transfer (FastNST) uses CNNs for real-time style transfer but lacks the enforcer clothing texture and detail structural, leading to ungrounded fashion practice [4]. FastNST is efficient in transferring artistic styles on images but not so great with matching designs to the intricate textures and shapes of clothes. While great in the ability to create images with renewed visual designs, it overlooks the natural curves of material and gives up an aesthetic appeal in art that would almost seem visually disparate from the original contours of the garment.
Conversely, FashionGAN and StyleGAN represent models explicitly designed for the generation of new fashion items. FashionGAN was created for designers to conceive clothing items based on sketches or written descriptions [5]. However, its limited viability in this respect relates to the restriction: it designs rather than incorporates existing works of art into clothing. StyleGAN features the power that generates grand and detailed images, enabling design manipulation, which is useful in generating innovative fashion ideas. The second point is that similar to FashionGAN, StyleGAN generates innovative designs better than smooth mixtures of external art, integrated with other existing clothing, along the lines of once again, no emphasis is placed on keeping garment textures while introducing the art.

ResNet18, a popular deep residual network, is one of the most used backbone networks in fashion tasks such as clothing recognition and feature extraction [7]. ResNet18 is exceptionally good at capturing fine-grained details and we use it to maintain the structural appearance of clothing when adding outside patterns. Pix2Pix is a generative adversarial network (GAN) which is used for image-to-image translation. We use a Pix2Pix-based model to add textures and patterns on clothing. It produces very realistic results, in which the added patterns look natural and consistent in shape and texture with the original garments.

These designs have greatly impacted fashion creation, dealing with different issues, and highlighting the continued difficulties in blending artwork with clothing seamlessly.

Contributions

a. *Resnet18*: ResNet18 was utilized for feature extraction from the T-shirt images, ensuring the preservation of key structural and textural attributes of the garment during the artwork integration process.

b. *Pix2Pix*: The Pix2Pix model performed image-to-image translation by mapping the artwork to the T-shirt with precise adherence to texture and visual coherence of the artwork and apparel [6].

Figure 1: *Output as Blended Tshirts*

The ResNet18 in collaboration with Pix2Pix showed reliable results by combining artworks to print upon T-shirts. The output of ResNet18 captured the detailed features of the t-shirt. Thus, the important structural details like texture of the fabric and outlines were clearly preserved with an MSE of 0.8319. The performance of the Pix2Pix model was excellent in transferring artworks to T-shirts by keeping the MSE at its minimum value of 0.0486 besides maintaining color uniformity as well as shape. From the MSE analysis, ResNet18 maintains the original T-shirt characteristics well and is good as a model for pixel-to-pixel mapping, whereas Pix2Pix is good for seam blending with greater accuracy. The combination of these models produced stylish T-shirts where artwork blended in very well.

In this paper, we present and examine a few of the most important sections. Under Methodology, we describe the dataset and architectures of the models: ResNet18 and Pix2Pix, discussing design choices and processing and training which led to results achieved. Under Results and Discussion, we detail model output, effectivity on embeddings of artwork onto T-shirts, and discuss on successes, limitations, and issues confronted. Limitations and Future Work This part of the work discusses the limitations of this method: that is, restrictions of the dataset and model effectiveness, for improvement recommendations in the application of artwear on different clothing types. It discusses how AI can improve brand equity, diminish manual labor, reduce wastage, and help streamline aspects of design to be both creative and have business applicability. Finally, the Conclusion summarizes the key findings of this research, reflecting on AI's broader implications for the fashion industry.

II. METHODOLOGY

In this section, we delve into the datasets, preprocessing techniques, and model architectures used in this research in great detail. The conversation starts with organizing the artwork and T-shirt data sets, then moving on to discuss the preprocessing methods that guarantee the data is correctly structured for training. Then we present the design of ResNet18 and Pix2Pix and examine the use of these models in addressing the challenge posed by smooth integration of artwork onto T-shirt. In this section, we look at the mechanics in detail, so that the reader may understand the essence that drives this endeavor.

A. *Dataset and Preprocessing*

This research utilizes a pair of datasets. The initial dataset consists of 8,683 images of famous artworks, representing a range of styles and textures. The pictures are in JPG form, offering a variety of but restricted examples for training the model. The second set of data includes a pair of JPEG images showing a white T-shirt; one image displays the T-shirt itself, while the other shows a masked section marking where artwork can be applied. Random images are chosen from the art collection to be blended onto the T-shirt. Data augmentation techniques like flips and rotations are used to improve generalization and increase variability in the dataset for the Pix2Pix model.

Preparation for ResNet18 involves resizing T-shirt images to 224x224 pixels, converting them to RGB format, and transforming them into tensors. The pixel range is then normalized using standard mean and deviation values for consistency, and an additional batch dimension is added to ready the images for model processing.

The Pix2Pix model is meant to enhance images with flips and rotations to produce further training data [8]. The sizes of the artwork images and T-shirt images have been resized and normalized to ensure the alignment and consistency of the translation from image to image.

B. *Proposed Architecture*

The Proposed Architecture section will accommodate the data flow diagram along with an overview of how ResNet18 and Pix2Pix function to naturally weld artworks onto T-shirts. Different model configurations, including training parameters and system specifications in use, will be documented for optimization of functionality.

1. *Diagram*

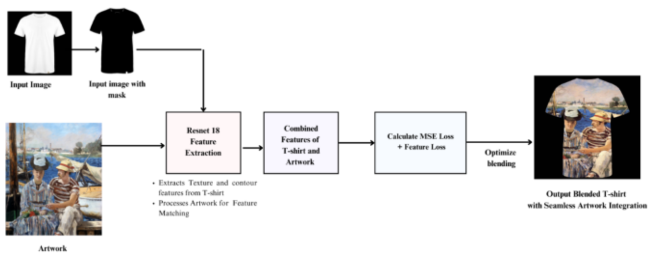


Figure 2: *Resnet18 Architecture*

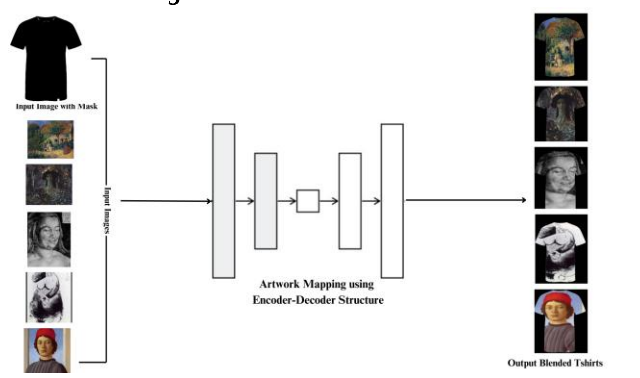


Figure 3: *Pix2Pix Architecture*

2. *Working of the Models*

ResNet18 addresses the vanishing gradient issue in deep networks by implementing deep residual learning within a convolutional neural network (CNN) [7]. ResNet18 plays a crucial part in capturing high-level characteristics from images of T-shirts in this study. These features, such as texture and outlines of fabric, are necessary to maintain the original shape integrity of the dress when superimposing artwork. ResNet18 maintains the shape and texture of the T-shirt so that the applied artwork does not alter the design of the fabric. The capability of the model to capture minute details makes the application seem as though it is in its original form closely on the surface of the T-shirt.

Pix2Pix is a type of conditional generative adversarial network (GAN) that is utilized for tasks involving translating images into other images [9]. Pix2Pix is utilized in this research to convert artwork onto T-shirt photos by creating a lifelike representation of the artwork on the clothing. In Pix2Pix, the generator creates a merged design, and the discriminator determines its accuracy through a comparison with the actual data. While the specific characteristics of the textile and T-shirt's shape were manipulated, the painting transfer has become beneficial to such operations. The pix2pix method worked superbly as it addressed many discrepancies in clothing design while adding artistic elaboration in a seamless and organic way. The system integrates a pix2pix model that efficiently streamlines the process, thereby reducing any owner-dependent and manual adjustments.

3. *Machine Configurations*

The experiment was performed and executed on models using MacBook Air M1 (2020), power provided by Apple's M1 chip with 8-core CPU, 7-core GPU, and 16-core Neural Engine. M1 chip is specifically designed for machine learning applications and thus yields superior performance for deep learning functionalities. Feature extraction and image translation are also computationally intensive tasks that this system works well on with 8GB of unified memory in ResNet18 and Pix2Pix models [7,10].

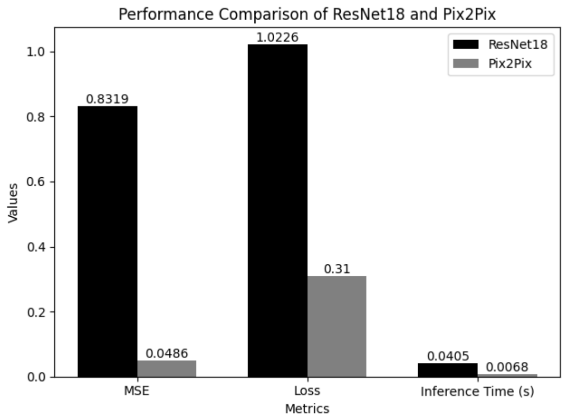
This utilization of local hardware was complemented with running computationally demanding models on Google Colab. With such infrastructure provided by Colab, specifically GPU acceleration, training and testing of the models proceeded much quicker. Major coding happened through Jupyter Notebook because it laid an interactive platform for seamlessly integrating preprocessing of data, model training, and evaluation in an integrated development environment. Pairing of the M1 technology of Apple with cloud resources from Google Colab was imperative to ensure that the testing process was efficient and expandable as necessary, thereby making it easier to run all required calculations.

III. Result and Discussion

The primary objective of the paper will be the successful transfer of images onto T-shirts using models ResNet18 and Pix2Pix. To evaluate their performance, we will make a comparison as to how well they can transfer images on to a T-shirt without losing quality and structural integrity. The significant measurements include Mean Squared Error (MSE), Structural Similarity Index (SSIM), and inference time apart from loss values for ResNet18 and Pix2Pix models.

ResNet18 Performance: The best mean squared error (MSE) reached by the ResNet18 model is 0.8319, while the value of the loss is about 1.0226. Thus, the overall accuracy of the extracted features and the preservation of the texture is reasonable. The model achieved a SSIM score of 0.9185, indicating a strong resemblance between the original T-shirt and the artwork-incorporated version, retaining essential structural elements and texture [11]. The model can process images quickly in real-time applications where speed is important, with an inference time of 0.0405 seconds.

Pix2Pix Performance: The losses for the generator and discriminator are found to be D_fake=0.286 and D_real=0.688, while G_loss was measured at 15. The balance between the actual and artificial bias shows that there was successful achievement of the reality of the artwork fusion by the generator while efficiently pushing the discriminator. The MSE and perceptual losses of 0.0486 and 0.31 respectively show how well the model preserves intricate details while translating images. Pix2Pix proved to be more efficient than ResNet18 in terms of speed, with an inference time of 0.0068 seconds, making it ideal for generating visually cohesive designs.



Graph 1: *Performance Comparison*

Graph 1 compares the ResNet18 model with the Pix2Pix model based on three important criteria: MSE, loss, and inference time. MSE for the Pix2Pix model is much lower compared to the value for the ResNet18 model at 0.0486 and 0.8319 respectively, meaning that Pix2Pix has a better reducing error from the output and target images. This makes Pix2Pix produce a more accurate visual rendition of the artwork T-shirt, and thus, a better candidate for tasks like the requirement of precise alignment and high visual fidelity.

The perceptual loss in Pix2Pix is 0.31. Much lower than the 1.0226 loss seen with ResNet18. Such loss might be said to characterize how Pix2Pix captures and retains the visual specifics of the image, including texture and color continuity with output closer to reality. Although ResNet18 proves to be relatively good at extracting features, its increased loss value shows that it might outperform Pix2Pix in many respects but could not lead towards quality visual outputs. The plot further suggests a vast time gap between processing these two models: the inference time of Pix2Pix comes out to be amazingly fast, that is, 0.0068 seconds, whereas the same for ResNet18 runs around 0.0405 seconds. This would make Pix2Pix efficient enough for real-time applications like automatic design processes and user interfaces for customization applications. Though GANs are computationally much more expensive, in terms of serving up productions where fast output generation is paramount, Pix2Pix with immediate processing of inputs has an edge over GANs [3].

In general, the analysis shows that Pix2Pix performs well in both speed and visual precision, making it a perfect option for creating art-infused designs. ResNet18's capability in extracting features could still be beneficial in situations where preserving the structural integrity of the clothing is important [12]. This brings attention to the trade-offs between two models, providing valuable information for brands and designers when choosing AI tools for automating fashion design.

Implications for the Fashion Industry: These results indicate that utilizing both ResNet18 and Pix2Pix together can simplify the integration of artwork onto clothing. The combination of ResNet18's ability to maintain intricate fabric details and Pix2Pix's skill in creating realistic designs provides a well-rounded solution for brands wishing to automate their design processes [10]. These advancements have the potential to greatly decrease the time and money required by traditional design methods, allowing for quicker production and more personalized product choices.

IV. LIMITATIONS AND FUTURE WORK

Although this study introduces a unique method for incorporating art onto clothing, it is important to recognize specific constraints. A limitation is present in the diversity of the dataset. The art dataset utilized in this research has a restricted quantity of images, which might not completely showcase the diverse array of styles and textures found in actual fashion. That is, this constraint might compromise the generalizability of the models to varied designs of complexity. While the ResNet18 and Pix2Pix perform well on designing T-shirts, there is likely a need for fine tuning on more complex patterns and alternate textures or unusual shapes of clothing items. The system in its current design is tailored towards a simple structure of garments and will not be entirely suitable for complicated clothing such as dresses or layered wear [13].

Beyond printing artwork on T-shirts, the possible uses of this study go further. Techniques developed in this study can be tailored to many items of fashion products, from handbags, footwear, and ornaments, among others, which gives brands a new avenue to put designs quickly and automatically. This has the potential to change conventional design methods, enabling businesses to experiment with imaginative opportunities more quickly and flexibly. Moreover, being able to test and modify designs digitally before entering production can greatly reduce the need for physical prototypes, leading to a reduction in material wastage and simplification of the prototyping process.

From a strategic standpoint, implementing AI-powered design processes can assist companies in minimizing manual labor in activities such as print design, leading to quicker completion times and uniform quality. This is especially important for small and medium-sized businesses seeking to expand, and for large manufacturers wanting to reduce expenses and enhance productivity. In addition to that, process automation will lead to lower labor costs and better use of the resources available since it allows for process execution. This, therefore, will boost profitability in an extremely competitive business. With the fashion industry ever more committed to sustainability, such a digital-first design approach would benefit it by having less waste in material and energy usage. Eventually, AI technologies advancements are going to play a crucial role in the fashion industry, maybe in order to create greater possibilities for creativity, greater accountability toward the environment, and greater responsiveness to what consumers have to say. In this vein, the future potential does not only view AI as a design tool but also as an industry-orientated driver for change.

V. CONCLUSION

This research investigated how ResNet18 and Pix2Pix models could be used to incorporate artwork onto T-shirts, offering a novel method for fashion design driven by AI. ResNet18 successfully captured clothing details, preserving fabric textures and shapes, whereas Pix2Pix was exceptional in transferring artwork accurately and cohesively. This combination of models proved that it could develop excellent, customized apparel products, which calls for no alteration of handmade design. This also proved to be a valuable business opportunity in terms of cost-cutting and acceleration of design processes and greater responsiveness to market trends and supportive of the whole industry because of its efforts toward sustainability where there is reduced waste made by the production of digital designs. These methods can be expanded in future studies with diverse types of products and more applications, and the outcome might be superior, innovative, and more eco-friendly design concepts for fashion industries.

ACKNOWLEDGMENT

I am deeply grateful to my advisor, Professor Sofia Francis, whose expertise and support were crucial in shaping this research. Her guidance throughout the process greatly enhanced the quality of the work. Lastly, I extend my heartfelt thanks to my family and friends for their unwavering encouragement, which provided the strength needed to complete this project.

REFERENCES

[1] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.

[2] Oh, J., Kim, S., Lee, C., Cha, J.-H., Yang, S., Im, S., ... Choi, S. (2023). Preventing Vanishing Gradient Problem of Hardware Neuromorphic System by Implementing Imidazole-Based Memristive ReLU Activation Neuron. Advanced Materials, 35.

[3] Kim, S., Lee, J., Jeong, K., Lee, J., Hong, T., & An, J. (2023). Automated door placement in architectural plans through combined deep-learning networks of ResNet-50 and Pix2Pix-GAN. Expert Systems with Applications, 244, 122932.

[4] Reimann, M., Buchheim, B., Semmo, A., Döllner, J., & Trapp, M. (2022). Controlling strokes in fast neural style transfer using content transforms. The Visual Computer, 38.

[5] Cui, Y., Liu, Q., Gao, C., & Su, Z. (2018). FashionGAN: Display your fashion design using Conditional Generative Adversarial Nets. Computer Graphics Forum, 37

[6] Choi, Y., Choi, M.-J., Kim, M., Ha, J.-W., Kim, S., & Choo, J. (2017). StarGAN: Unified Generative Adversarial Networks for Multi-domain Image-to-Image Translation. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition.

[7] Xue, G., Li, S., Hou, P., Gao, S., & Tan, R. (2023). Research on lightweight Yolo coal gangue detection algorithm based on resnet18 backbone feature network. Internet of Things, 22, 100762.

[8] Gao, R., Jie, L., & U, K. T. (2023). Complex Manga Coloring Method Based on Improved Pix2Pix Model. In *2023 International Conference on Machine Learning and Cybernetics (ICMLC)* (pp. 582-587).

[9] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2016). Image-to-Image Translation with Conditional Adversarial Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 5967-5976.

[10] Qu, Y., Chen, Y., Huang, J., & Xie, Y. (2019). Enhanced Pix2pix Dehazing Network. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

[11] Sara, U., Akter, M., & Uddin, M. S. (2019). Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. Journal of Computer and Communications.

[12] Gill, K. S., Anand, V., Gupta, R., & Pahwa, V. (2023). Invasive Ductal Carcinoma Classification Using ResNet18 Model on Transfer Learning Concepts. 2023 3rd International Conference on Intelligent Technologies (CONIT).

[13] Kazlacheva, Z., Ilieva, J., & Peneva, T. (2019). Fashion and Pattern Design of Ladies' Dresses with of Different Types of Draperies. Applied Researches in Technics, Technologies and Education.

[14] S. Nyarameni, B. Gayathri, K. Guna Sree, S. Rabiya, and S. S. Hussain, "Apparel Design Using Image Style Transfer," International Journal of Creative Research Thoughts (IJCRT), vol. 10, no. 6, pp. 2320-2882, June 2022.

[15] S. Jiang, J. Li, and Y. Fu, "Deep Learning for Fashion Style Generation," IEEE Transactions on Neural Networks and Learning Systems.

[16] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image Style Transfer Using Convolutional Neural Networks," CVPR, 2016

[17] M. Berning, K. M. Boergens, and M. Helmstaedter. SegEM: Efficient Image Analysis for High-Resolution Connec- tomics. Neuron, 87(6):1193–1206, Sept. 2015.

[18] P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," arXiv preprint arXiv:1611.07004v3, 2018.

[19] M. Cimpoi, S. Maji, and A. Vedaldi. Deep filter banks for texture recognition and segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recogni- tion, pages 3828–3836, 2015.

[20] Chandran, Prashanth & Zoss, Gaspard & Gotardo, Paulo & Gross, Markus & Bradley, Derek. (2021). Adaptive Convolutions for Structure-Aware Style Transfer. 10.1109/CVPR46437.2021.00788.

[21] Dae Young Park and Kwang HeeLee. 2019. Arbitrary Style Transfer with Style- Attentional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.IEEE,5880–5888.

[22] IEEE Transactions on Pattern Analysis and Machine Intelligence, Explicit Filterbank Learning for Neural Image Style Transfer and Image Processing July 2021, pp. 2373- 2387, vol. 43, DOI Bookmark: 10.1109/TPAMI.2020.2964205

[23] Alpana Dubey, Nitish Bhardwaj, Kumar Abhinav, Suma Mani Kuriakose, Sakshi Jain, and Veenu Arora. 2020. AI Assisted Apparel Design. In KDD '20 Workshop on AI for fashion supply chain, 24 August 2020, San Diego, USA. ACM, New York, NY,USA

[24] J. E. Kyprianidis, J. Collomosse, T. Wang, and T. Isenberg. State of the ”Art”: A Taxonomy of Artistic Stylization Tech- niques for Images and Video. Visualization and Computer Graphics, IEEE Transactions on, 19(5):866–885, 2013.

[25] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. arXiv:1409.0575 [cs], Sept. 2014. arXiv: 1409.0575.

[26] K. Simonyan, A. Vedaldi, and A. Zisserman. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. arXiv:1312.6034 [cs], Dec. 2013.