BeautifAI - Personalised Occasion-based Makeup Recommendation

Kshitij Gulati kshitij17162@iiitd.ac.in
University of Rochester

Gaurav Verma gauravv.co18@nsut.ac.in
Arizona State University

Mukesh Mohania mukesh@iiitd.ac.in
Indraprastha Institute of Information Technology, Delhi

Ashish Kundu ashkundu@cisco.com
Cisco Research

Editors: Emtiyaz Khan and Mehmet Gonen

Abstract

With the global metamorphosis of the beauty industry and the rising demand for beauty products worldwide, the need for a robust makeup recommendation system has never been more. Despite the significant advancements made towards personalised makeup recommendation, the current research still falls short of incorporating the context of occasion and integrating feedback for users. In this work, we propose *BeautifAI*, a novel recommendation system, delivering personalised occasion-oriented makeup recommendations to users. The proposed work's novel contributions, including incorporating occasion context to makeup recommendation and a region-wise method using neural embeddings, set our system apart from the current work in makeup recommendation. We also propose real-time makeup previews and continuous makeup feedback to provide a more personalised and interactive experience to users.

Keywords: Recommendation Systems; User Feedback; Style Transfer

1. Introduction

The beauty industry has seen massive growth over the past few years, given the advent of social media and targeted advertising. As we move through the world, we apply, add, and refresh our makeup throughout the day in anticipation of the different social contexts in which we will interact with others. The socio-economic conditions in India and other South-Asian countries are unique. They indicate that while a large section of the population has access to smartphones, few have the resources to regularly utilise beauty consultants' services to obtain personalised makeup beauty regimens. To ensure these services to a broader audience, we offer a personalised makeup recommendation system that is mobile deployable and easily accessible.

Various research and commercial tools have attempted to tackle various domains such as makeup recommendation Alashkar et al. (2017); Nguyen and Liu (2017); Liu et al. (2013); Ou et al. (2016), makeup product retrieval Zhang et al. (2019b) and makeup style transfer Zhang et al. (2019a); Chang et al. (2018); Li et al. (2018); Liao et al. (2017); Liu et al.

(2016). The current makeup recommendation work still falls short of providing a truly personalised experience in terms of makeup recommendation. These works lack the element of interactivity and do not capture the context of occasion for makeup recommendation. The systems Nguyen and Liu (2017); Liu et al. (2013); Alashkar et al. (2017); Ou et al. (2016) recommend a holistic makeup recommendation (providing one single makeup recommendation for the entire face) which allows limited diversity. Any attempt at tackling the makeup recommendation problem must address the questions, (1) *What kind of makeup should I wear to a particular occasion/event?* (2) *How will the chosen makeup look on my face?* (3) *Have I applied the makeup accurately with respect to my choice?* To address these questions, we propose *BeautifAI* - an end-to-end makeup recommendation system that recommends users personalised occasion-oriented makeup styles.

The proposed system requires users to provide two inputs - a picture of their face and the occasion for makeup. The system then recommends them a combination of makeup styles for the skin, eyes and lips most suited to their face and the occasion using our region-wise makeup recommendation method (4.2). The previews of these recommendations are then displayed on the picture of the users' face, allowing them to visualise how the makeup would look on their faces via our novel style-transfer method (4.3). While the users apply this recommended makeup, the system continuously provides them feedback on how accurately they have applied the recommended makeup combination (4.4). In summation, the main contributions of the proposed work are as follows:
(1) We contribute a diverse labelled makeup knowledge dataset that contextualises the occasion aspect in makeup recommendation.
(2) We propose a novel region-wise makeup recommendation method that improves upon the limited diversity in recommendation styles offered by the present work and affords greater personalisation.
(3) We contribute a region-wise makeup style transfer method based on Generative Adversarial Networks (GANs) for visualising makeup recommendation results.
(4) We propose a novel continuous feedback method for makeup evaluation that assists users in achieving their desired makeup look in real-time.
We further evaluate different aspects of the proposed system's performance by conducting a comprehensive pilot study.

2. Related Work

2.1. Makeup Recommendation

In recent works, various attempts have been made towards personalised makeup recommendation by using technologies such as those proposed by Alashkar et al. (2017); Nguyen and Liu (2017); Liu et al. (2013); Ou et al. (2016). Alashkar et al. proposes a unique methodology where an inference engine is trained on a set of predefined rules followed by a deep neural network on a labelled dataset to recommend makeup styles. Nguyen et al. employs a latent SVM model to draw relationships between makeup invariant features and makeup attributes for efficient makeup recommendation. Liu et al. (2013) propose a multiple tree-structured super-graphs model to express the complex relationships among beauty and beauty-related attributes from extracted facial and clothing features. Despite these methods achieving good results with their unique approaches for makeup recommendation, they are limited in their scope as they lack context for occasion and provide a holistic recommendation approach (providing one single makeup recommendation for the entire face).

2.2. Makeup Style Transfer

Style Transfer is defined as transferring the makeup style (fully or partially, as needed) of a person from one image to another image. Recently, GANs have been widely used for transferring style, patterns, texture from one image to another and are capable of generating high resolution realistic resultant images Zhu et al. (2017). The research proposed in Zhang et al. (2019a); Chang et al. (2018); Li et al. (2018); Chen et al. (2019) utilise GANs for makeup transfer and provide accurate and natural makeup style transfer results with some distortion. After evaluating many recently developed makeup style transfer techniques, we found Li et al. (2018); Chen et al. (2019) the best on the basis of efficiency and reproducibility of results. The user study Chen et al. (2019) found that Beautyglow fares best when compared with other makeup style transfer techniques, on the basis of quality, realism and makeup style similarity. However, these works transfer the entire makeup from one face to another and are not compatible with the results provided by our region-wise makeup recommendation method.

Table 1: Comparison of released datasets with our dataset

Dataset	**Total Images**	**Attribute Labelling**	**Occasion Labelling**
MIFS	642	No	No
MT	3834	No	No
Alashkar et al.	961	Yes	No
BeautifAI	**804**	**Yes**	**Yes**

3. Dataset

Presently, various makeup datasets have been collected for different research perspectives - YMU (YouTube Makeup) Dantcheva et al. (2012), VMU (Virtual Makeup) Dantcheva et al. (2012), MIW (Makeup in the "wild") Chen et al. (2013), MIFS (Makeup Induced Face Spoofing) Chen et al. (2017), MT Li et al. (2018) and SMU Wang and Fu (2016). However, none of these datasets has adequate labelling for makeup styles as well as occasion (see Table 1). To perform accurate occasion-oriented makeup recommendation, we contribute our high-quality dataset to integrate the aspect of occasion with makeup styles and attributes.

Social media platforms provide much user-centric data with images displaying various real-world makeup styles with natural backgrounds from users across the globe. Our dataset comprises 804 annotated images scraped from Instagram/Pinterest. Collecting the images from social media ensures that they are up-to-date with the latest makeup trends. On interviewing makeup experts, we find that the three occasions - *Casual, Office* and *Party* most accurately represent the current makeup trends. We extract images from these platforms using numerous currently trending hashtags and keywords most relevant to that occasion. We ensure that the images extracted are diverse in terms of ethnicity and balanced in terms of occasion and makeup styles. We also perform manual filtering to ensure all images adequately represent makeup knowledge and are face-frontal images. We define a set of attributes and classes, similar to that of Alashkar et al. (2017), to best represent the fashion knowledge in terms of the makeup knowledge in each image (see Appendix A). We believe makeup attribute labelling is essential as this would allow users to acquire the products from the recommended makeup styles with ease. *We will be making our dataset publicly available on publication.*

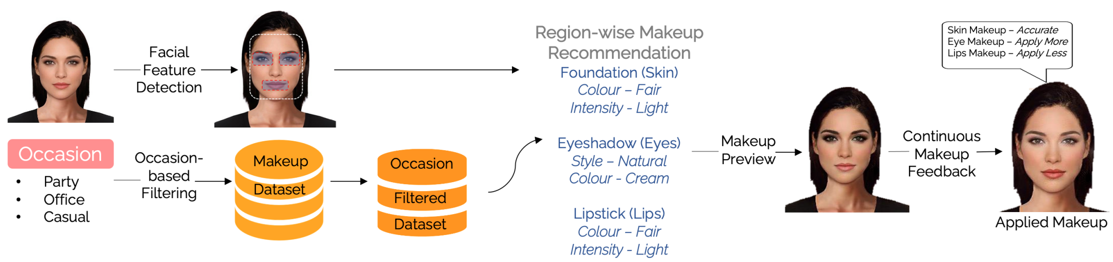

Figure 1: The proposed pipeline showing the four different methods of the proposed system

4. Methodology

The proposed system consists of four methods to provide users with the best experience while recommending a beauty regimen of makeup styles. The proposed pipeline for our system consisting of different methods is shown in Fig 1. These four methods are described in detail below. Methods 4.2, 4.3 and 4.4 are novel contributions of our research while method 4.1 is implemented by extending state of the art algorithms.

4.1. Facial Feature Extraction

During the past decade, there have been great advances in the detection and extraction of facial features for a variety of applications. We extract facial features to provide high quality makeup recommendations that are personalised for every user. Each makeup recommendation is unique to each user as it is suited to their peculiar facial structure and features. We propose a system where for a given image of the user as input, this method identifies three regions of interest i.e. the skin, the eyes and the lips (see Fig 2).

For identifying the regions of the face, we use a Faster RCNN model (with a ResNet-50 backbone) Ren et al. (2016), which is a state of the art convolutional neural network for object detection, pre-trained on the MS Coco Dataset Lin et al. (2014). We then fine-tune this model by freezing the first 50% layers and performing transfer learning on the MUCT Face Database Milborrow et al. (2010). For training, we generate the bounding boxes for each of the regions on the MUCT Database on the basis of 76 facial landmarks, detecting the regions of interest in real time. We fine-tune this model on the MUCT dataset because the presence of 76 facial landmarks allows for greater precision in finding the relevant regions of interest.

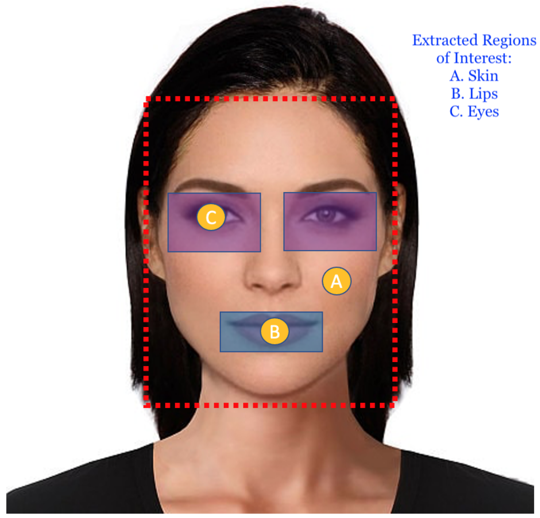

Figure 2: Extracted facial regions of interest

4.2. Occasion-oriented Region-wise Makeup Recommendation

This method's central premise is to incorporate the occasion concept for makeup and ensure greater customisability over existing works in terms of makeup recommendation. We assume that two individuals with similar facial features and similar facial structure would be suited to similar makeup styles Ou et al. (2016). Therefore, to personalise makeup recommendation, users are recommended styles from the dataset with whom their facial features match with the highest similarity.

The user picks the occasion (Casual, Professional/Office, Party) and provides a picture of their face I as an input image to the system. We then perform occasion-based filtering to filter out the makeup styles from our dataset based on the occasion and extract facial features from image I using method 4.1. The three extracted regions of interest for image I are I^{skin}, I^{eye} and I^{lip}. After extracting these facial features I^{skin}, I^{eye} and I^{lip}, we compare them with the occasion-filtered makeup images from our curated dataset - S_i, where i=$1, 2, 3......n$ and n is the number of occasion-filtered images in our dataset. We extract three regions of interest for all images S_i i.e. ${S_i}^{skin}$, ${S_i}^{eye}$ and ${S_i}^{lip}$ as bounding boxes using method 4.1.

We compare regions of interest of I with S_i (skin, eyes and lips) by building a CNN model and leveraging the concept of neural embeddings (mapping discrete variables to continous vector matrices). For the backbone of our CNN model, we use a SE-ResNeXt-50 Hu et al. (2018) pre-trained on the VGGFace2 dataset Cao et al. (2018). The Squeeze-and-Excitation (SE) blocks added to the ResNeXt-50 model, improve its representational ability by enabling it to perform dynamic channel-wise feature recalibration and ensure that the neural embeddings are robust. We fine-tune this CNN model by performing transfer learning on the extracted regions of interest (resizing to 128x128) for our entire dataset ${D_i}^{skin}$, ${D_i}^{eye}$ and ${D_i}^{lip}$, after freezing the first 75% hidden layers. After training this model, we then generate neural embeddings for ${S_i}^{skin}$, ${S_i}^{eye}$ and ${S_i}^{lip}$ as well as I^{skin}, I^{eye} and I^{lip}.

Let the neural embeddings be $S_i'^{skin}$, $S_i'^{eye}$, $S_i'^{lip}$ and I'^{skin}, I'^{eye}, I'^{lip} respectively. We then calculate the similarity between the embeddings I' and S_i' in the euclidean space. We use neural embeddings instead of classifying the regions of interest into discrete categories as it allows us to detect finer margins and compare them with greater precision.
This method recommends makeup for the regions - skin, eyes and lips in three phases. First, this method recommends skin makeup by retrieving the top ten most similar images from the dataset by comparing the neural embeddings I'^{skin} (the entire face except the eye and lip region) with the embeddings $S_i'^{skin}$. Once the user picks a skin makeup style, the method recommends the user top ten eye makeup styles followed by top ten lip makeup styles by similarly comparing the respective embeddings. The alogrithm is as follows:

for j= (skin,eyes,lips) **do** $similarities$ = list()
  for i = 1 to n **do** $distance$ =$|S_i'^j - I'^j|$; add $distance$ to $similarities$;
  end for(Return TOP 10 $S_i'^j$ with min $distance$;)
end for

The incorporation of region-wise recommendation instead of the holistic makeup recommendation (providing a single makeup recommendation for the entire face), as present in the current state of the art, ensures users a greater variety of makeup styles to choose from and greater personalisation. We believe that greater customisability is essential for personalisation as it allows users to pick only the components of a makeup style that they like and choose makeup styles on the basis of the makeup accessories they already own or can easily purchase.

4.3. Makeup Style Transfer

To show users previews of how the chosen makeup combination would look on their face, we contribute a style-transfer method in the proposed system. This method provides the user with a preview of how their face would look after every makeup stage. Method 4.2 recommends makeup styles from different regions of different facial images in our dataset. To provide previews for region-wise recommendation, we require an advanced solution for a region-wise makeup transfer. To generate real-time previews, we implement a region-wise style transfer method based on GAN technology, given it's recent success in makeup transfer based works. We initialise our model as a general pretrained GAN-based network Zhu et al. (2017). We extract the facial regions and facial features on the entire MT dataset using method 4.1. We then fine-tune this pre-trained GAN network on the MT Dataset using a Dilated ResNet (DRN) architecture after freezing the first 25% layers Chang et al. (2018). We separately train the model for each eyes, lips and skin to ensure our model is able to perform style-transfer for each region of interest independently and robustly in accordance with our region-wise makeup recommendation. The results of our model are considerably better than those of other methods because our model while utilising the DRN architecture, is trained on the much larger MT dataset independently for all three facial regions. Fig 3 displays results from our style transfer method. Here, the makeup is transferred from the reference image to the original image.

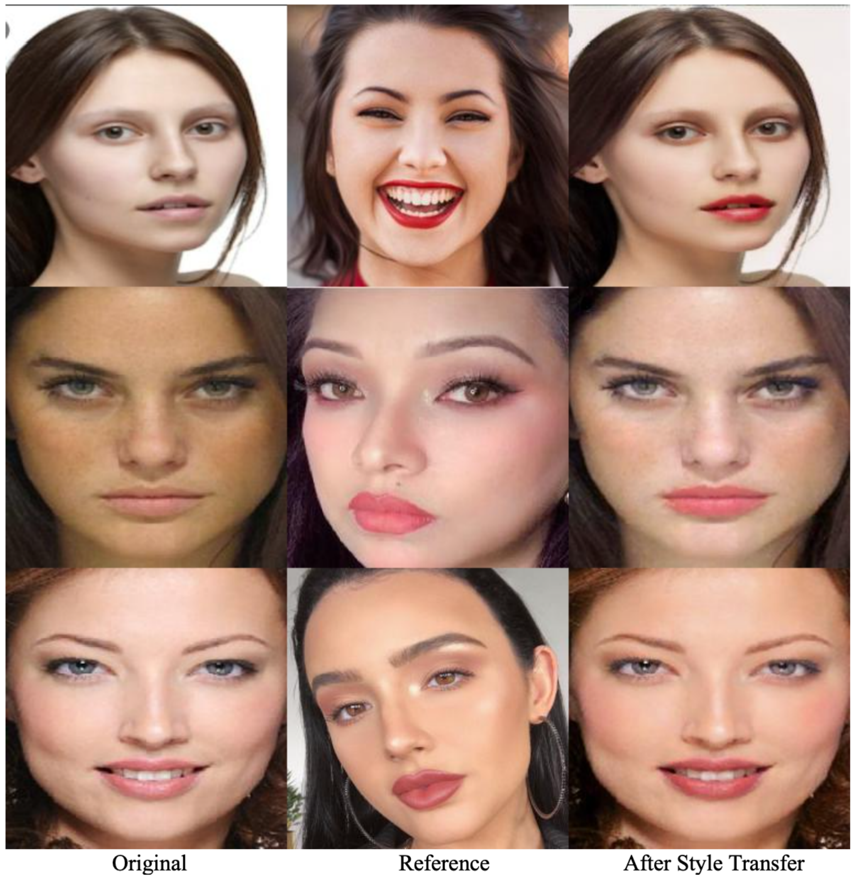

Figure 3: Previews from the Makeup Style Transfer Method

4.4. Continuous Makeup Evaluation

This method provides users with continuous feedback on how their applied makeup looks compared to the makeup recommended by the system in real-time and is a novel contribution of this work. Applying the makeup with similar precision as a system-recommended makeup image is nearly impossible for the average user. This method makes the system more interactive and provides the user with interpretive feedback to replicate their desired makeup look with the greatest precision.

To provide continuous feedback, this method requires the user to provide a picture after applying all the makeup (skin, eyes and lips). Let the image uploaded by the user with the makeup be U. This method then compares this image U with the desired makeup image X to tell the user if the makeup has been applied accurately (shown in the Continuous Evaluation step in Fig 1) with respect to each skin, eyes and lips makeup.

We compare these regions by their makeup colour intensities by calculating their Color Correlogram Huang et al. (1997) vectors. The primary benefit of using this technique is that it incorporates color adjacency and intensity and robustly tolerates significant changes in viewing positions, camera zooms, etc. The two correlogram vectors are compared by measuring their similarity and then telling the user whether the makeup (lips, eyes or skin) is either accurate or more/less based on empirically obtained thresholds. We segment the image U into three regions U^{skin}, U^{eyes} and U^{lips} via method 4.1 and compare the regions of interests with the corresponding region of interest from the final preview image X^{skin}, X^{eyes} and X^{lips} (visuals generated by style transfer). Let the colour correlogram vector of the user-applied makeup regions be ${U_C}^{skin}$, ${U_C}^{eyes}$ and ${U_C}^{lips}$ and for the preview image be ${X_C}^{skin}$, ${X_C}^{eyes}$ and ${X_C}^{lips}$. We then calculate aggregate similarity for each of these regions and provide user feedback on whether the makeup for each region is *more, less* or *accurate* as shown below.

for i= (skin,eyes,lips) **do**
  $\mathbf{Method4.1}(U)$, $\mathbf{Method4.1}(X)$= U^i,X^i;
  correlogram(U^i), correlogram(X^i)=${U^i}_C$,${X^i}_C$;
  similarity(${U^i}_C$,${X^i}_C$) = (less,accurate,more);
end for

5. Pilot Study and Evaluation

To evaluate the proposed system quantitatively and qualitatively, we conduct a pilot study. The users comprise 45 females belonging to the age group 18-45. Given our sample size of 45 and a user population of 1000, we have a confidence interval of 14.28 at a 95% confidence level. We plan to expand the sample size in future studies. Given that our system's central aim is to provide quality recommendations, improve personalisation and the user experience, we conduct a pilot study to evaluate our system thoroughly, similar to other state of the art works for makeup recommendation Alashkar et al. (2017); Li et al. (2018).

Table 2: Results of the pilot study for makeup recommendation (normalised)

Method	Diversity	Relevance
Holistic Makeup Recommendation	0.63	0.71
Region-Wise Makeup Recommendation (our proposed method)	**0.83**	**0.86**

5.1. Makeup Recommendation

We evaluate method 4.2 to demonstrate its efficacy over the traditional holistic makeup recommendation. We ask users to upload a selfie/full frontal facial image and choose the occasion for makeup. Then we provide the users with makeup styles using holistic makeup recommendations, followed by makeup styles by utilising our novel region-wise makeup recommendation. We ask users to rate each recommendation on a scale of 1-10 based on (1) relevance in terms of occasion and makeup styles. (2) diversity in the recommended makeup styles for a chosen occasion. The pilot study results (Refer Table 2) show that users rate the recommendations of our system highly in terms of diversity, affirming the

Table 3: Results of the pilot study for Makeup Style Transfer

Method	**Eyes**	**Lips**	**Skin**
BeautyGAN[Li et al. (2018)]	18%	31%	16%
Beautyglow[Chen et al. (2019)]	37%	12%	39%
BeautifAI	**45%**	**57%**	**45%**

success of our system over the holistic makeup recommendation approach. We also find that, users find recommendations with our system more relevant in terms of occasion and the makeup styles themselves than the holistic makeup recommendation. This confirms that we have been able to improve personalisation in comparison with previous works.

5.2. Makeup Style Transfer

To evaluate the quality of previews by method 4.3, we ask users to compare the results of the previews provided by the proposed method versus that provided by Li et al. (2018) and Chen et al. (2019). For every recommendation, we provide users with three different previews and ask the user to rate which visual they find most appropriate in terms of eyes, lips and skin. Table 3 depicts the percentage of users that pick each algorithm. We find that a majority of users rate the results of our system the best in terms of eyes, lips and skin.

5.3. Continuous Makeup Evaluation

We then evaluate how effective the addition of our novel continuous evaluation method is to makeup recommendation. We ask users to evaluate its efficacy in providing feedback and helping users achieve their look better in terms of each skin, eyes and lips makeup. Table 4 depicts the percentage of users that find the continuous evaluation method valuable in assisting them while applying their makeup. We ask the users to rate the continuous evaluation as 'Highly Effective', 'Moderately Effective' or 'Not Effective'. The results of this evaluation (Refer Table 4) affirm that a majority of users find the continuous evaluation method highly/moderately effective for all three regions of makeup. We also find that most users find the continuous evaluation for skin makeup the most helpful while applying makeup.

We also evaluate the effectiveness of the different image processing algorithms for continuous makeup evaluation by asking users to choose between the makeup feedback provided by the colour histogram, colour correlogram and SURF algorithms. We find that users find the feedback provided by the colour correlogram algorithm the most accurate for skin, eyes and lip makeup (Refer Table 5), affirming our choice.

6. Conclusion and Future Work

In this paper, we designed a system ensuring high-quality personalised makeup recommendations and greater interactivity. We made contributions to the domain of makeup recommendation by proposing occasion-oriented region-wise makeup recommendation, region-wise makeup transfer and continuous makeup evaluation. We successfully incorporated the

Table 4: Results of the pilot study for Continuous Evaluation

Continuous Evaluation Region	**Highly Effective**	**Moderately Effective**	**Not Effective**
Lips	23.3%	53.3%	23.3%
Eyes	56.6%	30.0%	13.3%
Skin	76.6%	6.6%	16.6%

Table 5: Comparison of Image Processing Techniques for Continuous Makeup Evaluation

Algorithm	**Eyes**	**Lips**	**Skin**
Colour Histogram	0%	19%	33%
Colour Correlogram	**68%**	**55%**	**53%**
SURF	32%	26%	14%

context of occasion and improved over existing works by proposing a region-wise makeup recommendation method. Our pilot study confirmed the efficacy of these novel additions in terms of personalisation. We conclude that users found the recommendations provided by this system highly relevant and diverse. Users also rated the results of our style transfer method highly and found the addition of the continuous makeup evaluation feature highly useful. Our dataset primarily consists of female images in the 18-45 age group, and we conducted our pilot study on a similar age-wise demographic. In the future, to further assess the performance of our system, we plan on expanding our dataset to cover a larger demographic and more occasions. We plan to conduct our pilot study on a larger sample to reduce bias and reduce error. We also believe that the system can be further improved by instilling higher-level makeup concepts like ethnicity, cultural dependence of makeup styles, time (day/night), body features, and dress, making recommendations more unique and personal.

References

Taleb Alashkar, Songyao Jiang, Shuyang Wang, and Yun Fu. Examples-rules guided deep neural network for makeup recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

Qiong Cao, Li Shen, Weidi Xie, Omkar M Parkhi, and Andrew Zisserman. Vggface2: A dataset for recognising faces across pose and age. In *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*, pages 67–74. IEEE, 2018.

Huiwen Chang, Jingwan Lu, Fisher Yu, and Adam Finkelstein. Pairedcyclegan: Asymmetric style transfer for applying and removing makeup. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 40–48, 2018.

C. Chen, A. Dantcheva, and A. Ross. Automatic facial makeup detection with application in face recognition. In *2013 International Conference on Biometrics (ICB)*, pages 1–8, 2013. doi: 10.1109/ICB.2013.6612994.

C. Chen, A. Dantcheva, T. Swearingen, and A. Ross. Spoofing faces using makeup: An investigative study. In *2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*, pages 1–8, 2017. doi: 10.1109/ISBA.2017.7947686.

Hung-Jen Chen, Ka-Ming Hui, Szu-Yu Wang, Li-Wu Tsao, Hong-Han Shuai, and Wen-Huang Cheng. Beautyglow: On-demand makeup transfer framework with reversible generative network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10042–10050, 2019.

A. Dantcheva, C. Chen, and A. Ross. Can facial cosmetics affect the matching accuracy of face recognition systems? In *2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 391–398, 2012. doi: 10.1109/BTAS.2012.6374605.

Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

Jing Huang, S Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih. Image indexing using color correlograms. In *Proceedings of IEEE computer society conference on Computer Vision and Pattern Recognition*, pages 762–768. IEEE, 1997.

Tingting Li, Ruihe Qian, Chao Dong, Si Liu, Qiong Yan, Wenwu Zhu, and Liang Lin. Beautygan: Instance-level facial makeup transfer with deep generative adversarial network. In *Proceedings of the 26th ACM international conference on Multimedia*, pages 645–653, 2018.

Jing Liao, Yuan Yao, Lu Yuan, Gang Hua, and Sing Bing Kang. Visual attribute transfer through deep image analogy. *ACM Trans. Graph.*, 36(4), July 2017. ISSN 0730-0301. doi: 10.1145/3072959.3073683. URL https://doi.org/10.1145/3072959.3073683.

Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

S. Liu, L. Liu, and S. Yan. Magic mirror: An intelligent fashion recommendation system. In *2013 2nd IAPR Asian Conference on Pattern Recognition*, pages 11–15, 2013. doi: 10.1109/ACPR.2013.212.

Si Liu, Xinyu Ou, Ruihe Qian, Wei Wang, and Xiaochun Cao. Makeup like a superstar: Deep localized makeup transfer network. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, IJCAI'16, page 2568–2575. AAAI Press, 2016. ISBN 9781577357704.

S. Milborrow, J. Morkel, and F. Nicolls. The MUCT Landmarked Face Database. *Pattern Recognition Association of South Africa*, 2010.

Tam V Nguyen and Luoqi Liu. Smart mirror: Intelligent makeup recommendation and synthesis. In *Proceedings of the 25th ACM international conference on Multimedia*, pages 1253–1254, 2017.

Xinyu Ou, Si Liu, Xiaochun Cao, and Hefei Ling. Beauty emakeup: A deep makeup transfer system. In *Proceedings of the 24th ACM International Conference on Multimedia*, MM '16, page 701–702, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450336031. doi: 10.1145/2964284.2973812. URL https://doi.org/10.1145/2964284.2973812.

Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6):1137–1149, 2016.

Shuyang Wang and Yun Fu. Face behind makeup. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

Honglun Zhang, Wenqing Chen, Hao He, and Yaohui Jin. Disentangled makeup transfer with generative adversarial network. *arXiv preprint arXiv:1907.01144*, 2019a.

Yi Zhang, Linzi Qu, Lihuo He, Wen Lu, and Xinbo Gao. Beauty aware network: An unsupervised method for makeup product retrieval. In *Proceedings of the 27th ACM International Conference on Multimedia*, MM '19, page 2558–2562, New York, NY, USA, 2019b. Association for Computing Machinery. ISBN 9781450368896. doi: 10.1145/3343031.3356075. URL https://doi.org/10.1145/3343031.3356075.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

Appendix A. Data Annotation

Fig 4 depicts the annotation scheme for our dataset. Three fashion/makeup experts annotate the images for ten different makeup attributes and three different occasions, and the inter-annotator agreement is calculated using kappa statistics. Each image is annotated by two experts and a third expert breaks the tie in case of disagreement. We find the inter-annotator agreement reasonably high with a kappa value of 0.76 for makeup attribute annotation and 0.88 for occasion annotation. Therefore, we achieve our makeup dataset, which describes each image in terms of the person, makeup, and occasion.

Makeup Attribute	Value
Foundation Colour	Light, Fair, Medium, Dark
Foundation Intensity	Light, Heavy
Blush Colour	Blanc, Pink, Plum, Beige, Bronze, Coral, Copper, Orange
Blush Style	Oval, Square, Round
Blush Intensity	Light, Heavy
Eyeshadow Style	Cut Crease, Gradient, Smoky, Cat Eye, Halo Eye, Natural Eye
Eyeshadow Colour	Brown, Cream, Blue, Warm, Smoky
Eye liner	Light, Heavy, Winged
Lipstick Colour	Pink, Red, Orange, Purple, Nude
Lip liner	Yes, No

Professional/Office	Party	Casual
293	251	260

Figure 4: Annotation scheme for our curated dataset