Improved Obstructed Facial Feature Reconstruction for Emotion Recognition with Minimal Change CycleGANs*

Tim Büchner[1 [0000−0002−6879−552X]],
Orlando Guntinas-Lichius[2 [0000−0001−9671−0784]], and
Joachim Denzler[1 [0000−0002−3193−3300]]

[1] Computer Vision Group, Friedrich Schiller University Jena, 07745 Jena, Germany
{forname.surname}@uni-jena.de
[2] Department of Otolaryngology, University Hospital Jena, 07745 Jena, Germany
{forname.surname}@med.uni-jena.de

Abstract. Comprehending facial expressions is essential for human interaction and closely linked to facial muscle understanding. Typically, muscle activation measurement involves electromyography (EMG) surface electrodes on the face. Consequently, facial regions are obscured by electrodes, posing challenges for computer vision algorithms to assess facial expressions. Conventional methods are unable to assess facial expressions with occluded features due to lack of training on such data. We demonstrate that a CycleGAN-based approach can restore occluded facial features without fine-tuning models and algorithms. By introducing the minimal change regularization term to the optimization problem for CycleGANs, we enhanced existing methods, reducing hallucinated facial features. We reached a correct emotion classification rate up to 90% for individual subjects. Furthermore, we overcome individual model limitations by training a single model for multiple individuals. This allows for the integration of EMG-based expression recognition with existing computer vision algorithms, enriching facial understanding and potentially improving the connection between muscle activity and expressions.

Keywords: Image Restoration · Facial Features · CycleGAN · Minimal Change.

1 Introduction

Facial expression recognition [12] is crucial in various research areas such as psychology, medicine, and computer vision. In computer vision, occlusion of facial features presents a challenge for existing algorithms, as most assume a fully visible face and lack consideration for occluded features in available datasets.

We focused on detecting six basic emotions (anger, disgust, fear, happiness, sadness, and surprise) defined by Ekman and Friesen [4], which are characterized

* Supported by Deutsche Forschungsgemeinschaft.

by specific facial muscle activation. Accurate measurement of muscle activation typically requires recordings detected via surface electrodes (sEMG). These electrodes and their cables cover parts of the faces. Combining EMG-based facial expression recognition with computer vision algorithms can enhance understanding of the underlying facial muscle activity. Our study involved 36 healthy subjects performing the six basic emotions four times, with recordings taken both with and without attached sEMG surface electrodes. This allowed for comparison between mimicked and targeted expressions. We employed the ResMaskNet [20] architecture for emotion detection but found that it struggled to handle occluded facial features. This resulted in only prediction of `anger` and `surprised`. Mimicked expressions, like `disgusted`, were recognized only 2 out of 528 times.

Büchner et al. [3] attempted to restore facial features by interpreting face coverage as a learnable style between uncovered and covered faces. Despite promising results, the method has drawbacks: it requires separate models for each individual, and it memorizes uncovered faces resulting in inadequate hallucination of occluded features, see Figure 3. We extended Büchner et al.'s work by introducing new regularization terms to the optimization problem, enabling a single model to be trained for multiple individuals. We increased individual accuracy from 33.8% (random guessing) up to 90% and demonstrated emotion detection on individuals not part of the training set, providing generalizability to unseen individuals. Fine-tuning the general model to specific individuals with minimal data further improved results.

We conducted ablation studies to enhance the backbone network, significantly reducing the number of parameters and computational cost while maintaining comparable results. We evaluated emotion classification accuracy and the following qualitative metrics: Frenchet Inception Distance (FID) [17], Structural Similarity Index (SSIM) [25], and Learned Perceptual Image Patch Similarity (LPIPS) [27]. These advancements are crucial for EMG-based facial expression recognition applications. The reduced parameters, improved generalization, and enhanced visual quality enable a more efficient, robust implementation. Additionally, these improvements could benefit live or therapeutic applications requiring visual feedback.

2 Related Work

Restoring occluded facial features is challenging due to their invisibility in the input images. Generative approaches, such as Generative Adversarial Networks (GANs) [6], can learn anatomically correct facial features from non-occluded faces, making them suitable for this task.

GANs have demonstrated strong results in image generation, particularly in medical applications [26]. However, they are typically trained on specific datasets and struggle to generalize to unseen data. Facial generation research [16,10,11] employs GANs to create realistic faces from specific datasets, indistinguishable from real faces [24]. These works focus on non-occluded faces, thus generating only non-occluded facial images.

Restoring hidden facial features is more challenging since GANs must learn facial features from non-occluded faces. Li et al. [15] used GANs to restore artificially altered faces; however, the unrealistic changes limit its applicability to real-world data. Moreover, the restored facial expressions do not match the original ones. Alternative methods attempt to use GANs for transferring facial attributes [18] to modify faces.

Some approaches attempted to generalize GANs to unseen data. For instance, Zhu et al. [28] demonstrated that CycleGANs can translate images between domains without paired data, allowing each generator to learn a specific domain style. CycleGANs can thus translate covered faces to uncovered faces in medical applications, treating sEMG coverage as a domain style.

Abraiam and Eklund [1] used CycleGANs to restore obfuscated faces in MRI images, focusing on patient privacy and side views with no facial expression differences. However, they did not evaluate the quality of restored faces. In contrast, our work focused on facial feature restoration and quality evaluation.

We build upon Büchner et al. [3], using the CycleGAN architecture for facial feature restoration. While their architecture remained unchanged, every individual required a separately trained model which could hallucinate facial features. We introduce a new regularization term to the optimization problem, enforcing minimal changes and also enabling a single model to be trained for multiple individuals.

3 Method

Restoring facial features can be viewed as a style transfer problem, as surface electrodes are applied consistently across individuals. Following Büchner et al. [3], we use the CycleGAN architecture for facial feature restoration, with covered faces as the source domain and uncovered faces as the target domain. The two generators and discriminators are trained adversarially, using GAN loss [6], cycle consistency loss [28], and identity loss [22]. However, the base model (Figure 3) hallucinates facial features or changes uncovered areas, affecting the original facial expression. This issue arises due to the absence of this constraint in the original training objective.

3.1 CycleGAN Architecture for Facial Feature Restoration

The CycleGAN architecture by Zhu et al. [28], consists of two generators and two discriminators, shown in Figure 1. The generators, trained adversarially [6], translate images between domains while maintaining cycle consistency [28], allowing the translated images to be reverted to the original domain. Discriminators distinguish between real and fake generated faces. To improve visual quality, generators are trained using an identity loss [22], preserving color composition between input and output. We hypothesize that the identity loss encourages generators to learn input faces' facial features, enabling better feature restoration.

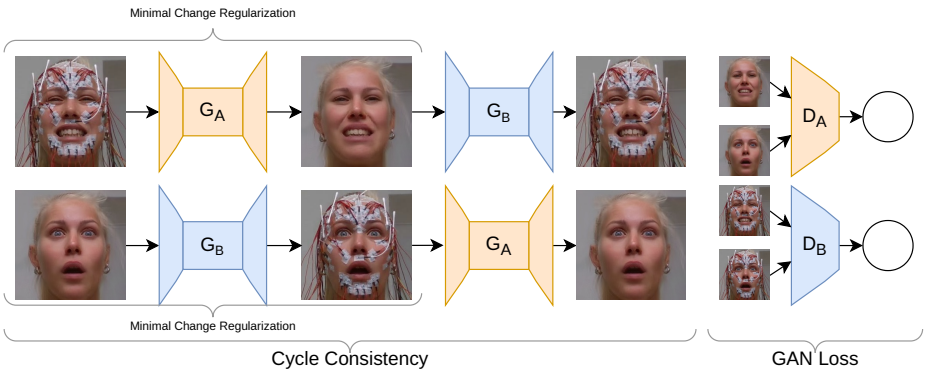

Fig. 1: The CycleGAN architecture for facial feature restoration: The generators G_A and G_B translate images between domains A and B. The minimal change regularization ensures that the generators do not modify uncovered areas.

Our generator's backbone network is a residual neural network [7]. The generator comprises two down-sampling blocks, n residual blocks, and two up-sampling blocks. In ablation studies, we explored the impact of varying residual block numbers in the generator. We found that $n = 5$ provided a balance between visual quality and computational cost and inference time. To minimize artifacts from deconvolution layers, we replace them with nearest-neighbor up-sampling layers followed by convolution layers. We use instance normalization as our normalization layer, because they are more suitable for style transfer [23]. This required a batch size of 1 during training to effectively learn the task. We maintained the discriminator architecture from [28,3], using the PatchGAN architecture [9]. We fixed the number of discriminator layers to 3, as this is sufficient for solving the task.

3.2 Optimization Problem with Minimal Change Regularization

The GAN loss (Equation 1) describes the adversarial interaction between the generator and discriminator. The generator attempts to fool the discriminator by generating uncovered faces. Cycle consistency loss (Equation 2) ensures accurate reattachment of the surfaces electrodes. The identity loss (Equation 3) enables the generator to learn the covered facial features:

$$\mathcal{L}_{GAN} = D_A(A, G_A(B)) + D_B(B, G_B(A)), \tag{1}$$

$$\mathcal{L}_{cycle} = \lambda_A \cdot \mathcal{L}_{L1}(G_B(G_A(A)), A) + \lambda_B \cdot \mathcal{L}_{L1}(G_A(G_B(B)), B), \tag{2}$$

$$\mathcal{L}_{idt} = \lambda_{idt}(\cdot\mathcal{L}_{L1}(G_B(A), A) + \cdot\mathcal{L}_{L1}(G_A(B), B)). \tag{3}$$

While the identity loss helps generators to learn facial features, it is insufficient for correct facial feature restoration. As seen in Figure 3, the generator hallucinated facial features not present in the original face, contrary to our goal. Limitations shown in [3] result from changes in uncovered face areas.

In style transfer tasks, every pixel of the input image might change according to the target domain's style. However, we wanted to enforce the generator to only change covered areas, preserving original facial features. We introduced a new regularization term (Equation 4) to the optimization problem, penalizing the generator for significant changes and reducing the reconstruction error between input and output images:

$$\mathcal{L}_{MC} = \lambda_{MC_A} \cdot \mathcal{L}_{L1}(G_B(B), B) + \lambda_{MC_B} \cdot \mathcal{L}_{L1}(G_A(A), A). \tag{4}$$

Due to the GAN loss, generators must still remove surface electrodes to deceive the discriminator. We introduced hyperparameters λ_{MC_A} and λ_{MC_B} to weight the regularization term, enabling control over each domain's regularization amount. Notably, this regularization allows for simultaneous training on multiple individuals. Since surface electrodes are positioned consistently across individuals, the model learns to preserve uncovered face areas. Consequently, the model is not limited to a single individual and can be applied to unseen individuals. The final optimization problem (Equation 5) comprises the GAN loss, cycle consistency loss, identity loss, and minimal change loss:

$$\mathcal{L} = \mathcal{L}_{GAN} + \mathcal{L}_{cycle} + \mathcal{L}_{idt} + \mathcal{L}_{MC}. \tag{5}$$

The optimization problem's individual components can be weighted using the corresponding λ hyperparameters. Setting λ_{MC_A} and λ_{MC_B} to 0 reverts to the original CycleGAN optimization problem.

4 Dataset

We evaluated our method on 36 test subjects [21] without medical conditions affecting facial expressions, such as facial paralysis. Using frontal cameras at 1280×720 pixels resolution and 30 fps, we recorded subjects mimicking the six basic emotions [4] four times each. The instruction order was randomized, and we conducted two recording sessions per subject, two weeks apart, obtaining alternating facial expressions videos with neutral expressions in between.

Each subject had two recordings without surface electrodes (baseline) and four recordings with surface electrodes to ensure accurate sEMG measurements. We used Fridlund and Cacioppo's [5] and Kuramoto et al.'s [13] schemes, applying 62 surface electrodes in total. The dataset comprised 71 recordings with electrodes and 138 without. We had 1704 emotions as ground truth for uncovered expressions and 3312 emotions requiring accurate reconstruction for electrode-covered faces. An overview of the six basic emotions with and without surface electrodes is shown in Figure 3.

5 Experiments and Results

Our experiments focused on accurate facial feature reconstruction, validated with emotion classification accuracy on restored faces and their visual quality. To assess visual quality, we used the Fréchet Inception Distance (FID) [8],

	Emotion Classification Accuracy							Visual Metrics		
sEMG	Anger	Disgusted	Fearful	Happy	Sad	Surprised	Overall	SSIM (↑)	LPIPS (↓)	FID (↓)
Without	61.6%	72.5%	28.1%	90.8%	47.1%	85.2%	64±10%	0.63±0.08	0.10±0.04	0.50±0.74
Attached	88.0%	0.3%	20.2%	21.9%	3.8%	68.8%	34±10%	0.38±0.05	0.25±0.02	10.46±2.10
Removed	66.2%	49.0%	11.9%	72.1%	42.5%	60.8%	**54±16%**	**0.66±0.09**	**0.08±0.04**	**0.35±0.48**

Table 1: Evaluation of emotion classification without (groundtruth), with, and removed surface electrodes and visual quality metrics: We achieved a higher classification rate on restored faces than on covered faces.

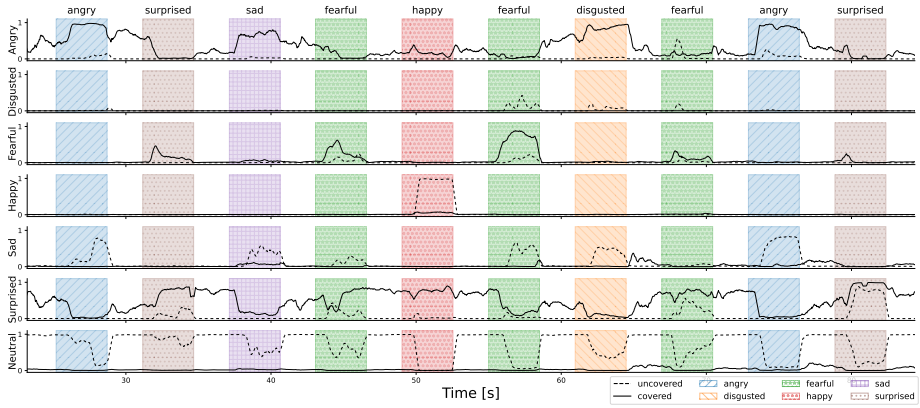

Fig. 2: Time series excerpt of emotion activations by ResMaskNet [20]: Dashed lines represent predicted emotions for uncovered faces, while solid lines represent predictions for covered faces. ResMaskNet struggles to correctly classify emotions for covered faces and predominantly activates for *angry* or *sad*. A neutral expression is never predicted for covered faces. (Best viewed digitally.)

the Structural Similarity Index (SSIM) [25], and the Learned Perceptual Image Patch Similarity (LPIPS) [27]. We employed the Residual Masking Network (ResMaskNet) [20], a ResNet-18 [7] architecture with a mask branch, for emotion classification. We did not fine-tune the model on uncovered or covered faces, ensuring unbiased performance analysis and demonstrating correct restoration. On the uncovered faces we have a mean accuracy of 64.1% as baseline. However, if test subjects fail to accurately replicate emotions, the estimation prediction may deviate from the ground truth. For unobstructed faces, one individual has a correct classification accuracy of 37.5%, while another achieves 91.5%. Thus, we expect a high variance in the emotion classification accuracy. We classified the 1704 uncovered face expressions and 3312 covered face expressions as a baseline for emotion classification accuracy, see Table 1. The emotion with the highest activation was selected as the predicted emotion, excluding the neutral expression. The accuracy for covered faces is 33.8%. The model predominantly predicts the emotions `angry` and `surprised`, as shown in Figure 2, indicating performance no better than random guessing between two classes. Emotions like `disgusted` and `sad` are predicted correctly only 23 times out of the expected 1104 times.

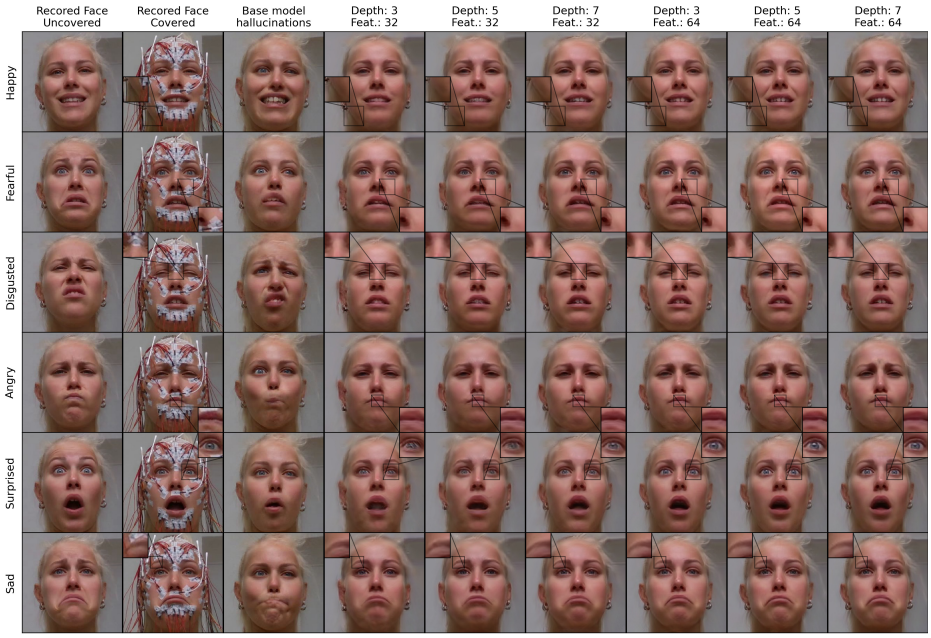

Fig. 3: Quality difference for backbone network sizes: The third column shows the base model without minimal change regularization, leading to hallucinated features. In contrast, the remaining columns incorporate minimal change regularization and tuned hyperparameters, showcasing improved restoration. (Best viewed digitally.)

Our goal was to improve emotion classification accuracy using our method without altering the ResMaskNet architecture. We aimed to restore faces through a data-centric approach, enhancing emotion classification performance.

Backbone Network Size We conducted experiments to determine an optimal backbone generator network size for achieving satisfactory results. The CycleGAN architecture was trained on various subsets for different depth and feature size combinations of the training data. Quantitative results are displayed in Table 2. Our findings indicate that emotion classification accuracy does not significantly improve with more training data. We measured a five percent point difference to the baseline accuracy and the ResMaskNet does not random guess between two classes anymore. However, the visual quality of restored faces increased with both larger amounts of training data and larger backbone networks. Fine details, highlighted in Figure 3, are better preserved with larger backbone networks and might not be measurable with these metrics. Considering the trade-off between visual quality and training time, we opted for a generators with three residual blocks and a feature dimension of 64. All further experiments utilize these model configurations.

Features	Depth	Emo. Max	Emo. Acc. (↑)	SSIM (↑)	LPIPS (↓)	FID (↓)
32	3	0.88	0.56±0.17	0.70±0.06	0.07±0.02	0.17±0.07
	5	0.83	0.56±0.15	0.70±0.07	0.07±0.02	0.25±0.16
	7	0.88	0.58±0.15	0.69±0.06	0.07±0.02	0.34±0.21
64	3	0.92	0.56±0.15	0.70±0.08	0.06±0.02	0.17±0.12
	5	0.88	0.58±0.15	0.64±0.08	0.09±0.04	0.37±0.59
	7	0.88	0.59±0.15	0.69±0.07	0.07±0.02	0.23±0.18

Table 2: Hyperparameter ablation study for the generators: The hyperparameter impact is minimal but the qualitative image generation improves.

5.1 Minimal Change Regularization

The proper restoration of faces was already shown by Büchner et al. [3] but their model might hallucinate features, see Figure 3. Thus, we investigated on the one side if the minimal change regularization improves the results and on the other side if it reduces the hallucination of features. We trained the CycleGAN architecture with different weightings of minimal change regularization for each individual test subject. Setting λ_{MC} to 0 is equivalent to the baseline model without regularization. The results in Table 3 show that the emotion classification accuracy increases with stronger minimal change regularization. This indicates that the additional loss term is no hindrance for the model to learn the task of removing the surface electrodes. We observed that increasing the λ_{MC} value increases overall performance. The restoration cannot correct an initial wrong mimicking of a facial expression. Thus, lowering the mean accuracy and impacting the measurable performance. This effect does not impact the visual metrics and they improve with higher λ values. We observed that hallucinations, such as in Figure 3, are not present anymore in the results, see Figure 4. However, an increased λ_{MC} value leads to a more blurry result but more accurate head orientation. Thus, small details might be lost in the restoration process. Thus, depending on the application, one can choose a suitable λ value to either preserve the facial features or to preserve the facial orientation.

λ_{MC}	Emo. Max	Emo. Acc.(↑)	SSIM (↑)	LPIPS (↓)	FID (↓)
0.0	0.79	0.46±0.16	0.64±0.09	0.09±0.04	0.41±0.66
0.1	0.83	0.48±0.16	0.64±0.09	0.09±0.04	0.37±0.67
0.2	0.88	0.47±0.15	0.64±0.08	0.09±0.03	0.45±0.64
0.3	0.83	0.47±0.15	0.64±0.09	0.09±0.04	0.45±0.71
0.4	0.83	0.48±0.15	0.64±0.08	0.09±0.03	0.35±0.57
0.5	0.92	0.54±0.16	0.66±0.09	0.08±0.04	0.35±0.48

Table 3: Emotion classification accuracy and FID score increase with stronger regularization: SSIM and LPIPS did not change significantly.

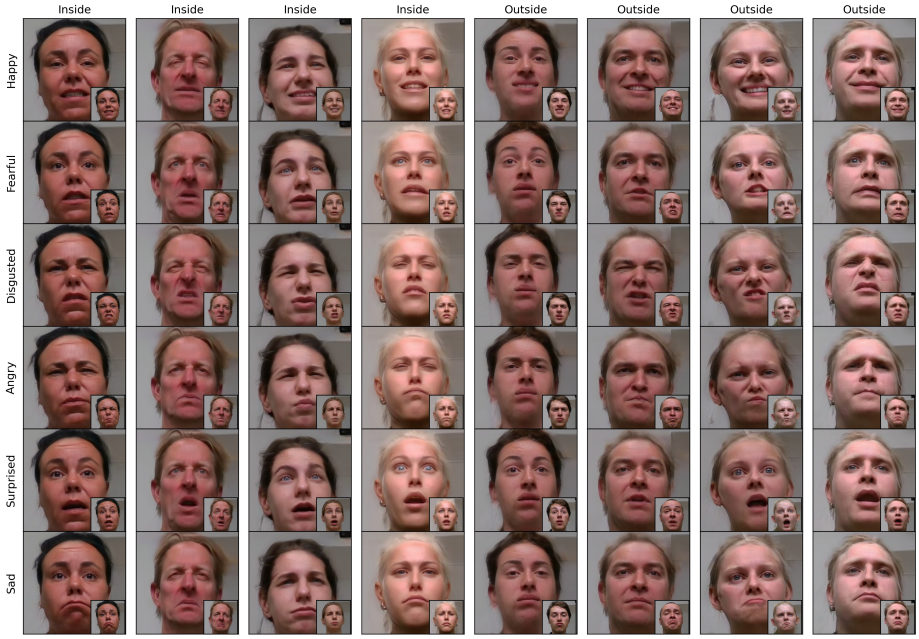

Fig. 4: Qualitative results of the generalization capabilities using the minimal change regularization: The model removes surface electrodes for individuals outside the training set. Important facial features are unchanged but the head shape is altered slightly. Uncovered expressions are shown in the lower right corner.

5.2 Generalization Capabilities

The CycleGAN architecture has the problem of not being able to generalize to unseen individuals as shown by Büchner et al. [3]. Thus, we investigate the generalization capabilities of the CycleGAN architecture with the minimal change regularization. We did a 6-fold cross validation on the 36 individuals, yielding 30 individuals for training and 6 individuals for testing. We set λ_{MC} to 0.5 for all experiments to enforce the minimal change regularization. Table 4 shows the results of the generalization experiments. The generalized model achieves a better performance on individual test subjects inside and outside the training set. Thus, we assume that the model learns to accurately remove the surface electrodes. This, in turn, leads to a better emotion classification accuracy on unseen individuals and we removed existing limitations of the work by Büchner et al. [3]. The qualitative results are shown in Figure 4 and we see that the model is able to remove the surface electrodes for unseen individuals.

Additionally, we investigated the impact of fine-tuning the model on the unseen individuals. We analyzed which combination of available training data and training epochs yields the best results. Our results in Figure 5 indicates that fine-tuning improved only the visual quality of the restored faces but not the emotion classification accuracy. We assume that important facial features

Trained On	Emo. Max	Emo. Acc. (↑)	SSIM (↑)	LPIPS (↓)	FID (↓)
False	0.83	0.53±0.15	0.60±0.06	0.12±0.03	0.53±0.84
True	0.88	0.55±0.18	0.61±0.09	0.10±0.04	0.33±0.29

Table 4: Generalization capabilities of the CycleGAN architecture with minimal change regularization: The model is able to generalize to unseen individuals.

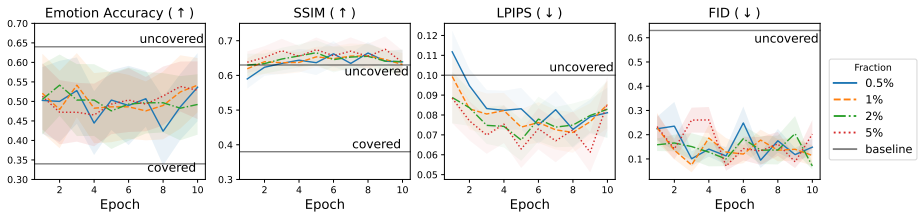

Fig. 5: Fine-tuning of the model on unseen individuals: The emotion classification accuracy does not improve but the visual quality of the restored faces does independent of the amount of training data and epochs.

are already learned by the model. Fine-tuning does not significantly enhance emotion classification accuracy; however, it improved visual quality compared to the baseline. Moreover, utilizing a smaller portion of training data and a higher number of epochs produces results similar to using larger training data fractions and fewer epochs.

6 Conclusion

Our study demonstrated that CycleGAN with minimal change regularization effectively restores individuals' faces with EMG surface electrodes, improving emotion classification accuracy. This regularization does not impede the model's ability to remove electrodes, and eliminates CycleGAN limitations like hallucinations, making the restored faces suitable for emotion classification tasks. Our approach enables direct utilization of existing methods without fine-tuning on our data. Additionally, we demonstrated its ability to generalize to unseen individuals and improve the visual quality of the restored faces without hallucinations. Here we showed that now sEMG-based measurements can be jointly used with computer vision-based techniques to enhance the comprehension of facial anatomy. The data-driven approach seamlessly integrates into existing pipelines, enabling real-time restoration of individuals' faces with surface electrodes. Thus, it can be used in applications where electrodes are used for facial muscle stimulation, such as physical therapy [2,14,19].

Acknowledgements This work has been funded by the Deutsche Forschungsgemeinschaft (DFG - German Research Foundation) project 427899908 BRIDGING THE GAP: MIMICS AND MUSCLES (DE 735/15-1 and GU 463/12-1).

References

1. Abramian, D., Eklund, A.: Refacing: Reconstructing Anonymized Facial Features Using GANS. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) pp. 1104–1108 (Apr 2019). https://doi.org/10.1109/ISBI.2019.8759515
2. Arnold, D., Thielker, J., Klingner, C.M., Puls, W.C., Misikire, W., Guntinas-Lichius, O., Volk, G.F.: Selective Surface Electrostimulation of the Denervated Zygomaticus Muscle. Diagnostics **11**(2), 188 (Feb 2021). https://doi.org/10.3390/diagnostics11020188
3. Büchner, T., Sickert, S., Volk, G.F., Anders, C., Guntinas-Lichius, O., Denzler, J.: Let's get the FACS straight - reconstructing obstructed facial features. In: International Conference on Computer Vision Theory and Applications (VISAPP). pp. 727–736. SciTePress (2023). https://doi.org/10.5220/0011619900003417
4. Ekman, P., Friesen, W.V.: Facial Action Coding System. Consulting Psychologists Press (1978)
5. Fridlund, A.J., Cacioppo, J.T.: Guidelines for human electromyographic research. Psychophysiology **23**(5), 567–589 (Sep 1986). https://doi.org/10.1111/j.1469-8986.1986.tb00676.x
6. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative Adversarial Networks. Advances in neural information processing systems **27** (Jun 2014). https://doi.org/10.48550/arXiv.1406.2661
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition pp. 770–778 (2016). https://doi.org/10.48550/arXiv.1512.03385
8. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. Advances in neural information processing systems **30** (2017). https://doi.org/10.48550/arXiv.1706.08500
9. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-Image Translation with Conditional Adversarial Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition pp. 1125–1134 (2017). https://doi.org/10.48550/arXiv.1611.07004
10. Kammoun, A., Slama, R., Tabia, H., Ouni, T., Abid, M.: Generative Adversarial Networks for Face Generation: A Survey. ACM Computing Surveys **55**(5), 1–37 (May 2023)
11. Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., Aila, T.: Analyzing and Improving the Image Quality of StyleGAN. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) pp. 8107–8116 (Jun 2020). https://doi.org/10.1109/CVPR42600.2020.00813
12. Klingner, C.M., Guntinas-Lichius, O.: Mimik und Emotion. Laryngo-Rhino-Otologie **102**(S 01), S115–S125 (May 2023). https://doi.org/10.1055/a-2003-5687
13. Kuramoto, E., Yoshinaga, S., Nakao, H., Nemoto, S., Ishida, Y.: Characteristics of facial muscle activity during voluntary facial expressions: Imaging analysis of facial expressions based on myogenic potential data. Neuropsychopharmacology Reports **39**(3), 183–193 (Sep 2019). https://doi.org/10.1002/npr2.12059
14. Kurz, A., Volk, G.F., Arnold, D., Schneider-Stickler, B., Mayr, W., Guntinas-Lichius, O.: Selective Electrical Surface Stimulation to Support Functional Recovery in the Early Phase After Unilateral Acute Facial Nerve or Vocal Fold Paralysis. Frontiers in Neurology **13** (2022)

15. Li, Y., Liu, S., Yang, J., Yang, M.H.: Generative Face Completion. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3911–3919 (2017)
16. Liu, M., Li, Q., Qin, Z., Zhang, G., Wan, P., Zheng, W.: BlendGAN: Implicitly GAN Blending for Arbitrary Stylized Face Generation. In: Neural Information Processing Systems (Oct 2021)
17. Liu, S., Wei, Y., Lu, J., Zhou, J.: An Improved Evaluation Framework for Generative Adversarial Networks. CoRR (Jul 2018). https://doi.org/10.48550/arXiv.1803.07474
18. Liu, Y., Li, Q., Sun, Z., Tan, T.: A ^{3} GAN: An Attribute-Aware Attentive Generative Adversarial Network for Face Aging. IEEE Transactions on Information Forensics and Security **16**, 2776–2790 (2021). https://doi.org/10.1109/TIFS.2021.3065499
19. Loyo, M., McReynold, M., Mace, J.C., Cameron, M.: Protocol for randomized controlled trial of electric stimulation with high-volt twin peak versus placebo for facial functional recovery from acute Bell's palsy in patients with poor prognostic factors. Journal of Rehabilitation and Assistive Technologies Engineering **7**, 2055668320964142 (Dec 2020). https://doi.org/10.1177/2055668320964142
20. Luan, P., Huynh, V., Tuan Anh, T.: Facial expression recognition using residual masking network. In: IEEE 25th International Conference on Pattern Recognition. pp. 4513–4519 (2020)
21. Mueller, N., Trentzsch, V., Grassme, R., Guntinas-Lichius, O., Volk, G.F., Anders, C.: High-resolution surface electromyographic activities of facial muscles during mimic movements in healthy adults: A prospective observational study. Frontiers in Human Neuroscience **16** (2022)
22. Taigman, Y., Polyak, A., Wolf, L.: Unsupervised Cross-Domain Image Generation. CoRR (Nov 2016). https://doi.org/10.48550/arXiv.1611.02200
23. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Instance Normalization: The Missing Ingredient for Fast Stylization. CoRR (Nov 2017). https://doi.org/10.48550/arXiv.1607.08022
24. Wang, X., Guo, H., Hu, S., Chang, M.C., Lyu, S.: GAN-generated Faces Detection: A Survey and New Perspectives. CoRR (Feb 2022)
25. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Transactions on Image Processing **13**(4), 600–612 (Apr 2004). https://doi.org/10.1109/TIP.2003.819861
26. Yi, X., Walia, E., Babyn, P.: Generative adversarial network in medical imaging: A review. Medical Image Analysis **58**, 101552 (Dec 2019). https://doi.org/10.1016/j.media.2019.101552
27. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition pp. 586–595 (Apr 2018). https://doi.org/10.48550/arXiv.1801.03924
28. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition pp. 2223–2232 (2017). https://doi.org/10.48550/arXiv.1703.10593