3D Face Data Augmentation Based on Gravitational Shape Morphing for Intra-Class Richness

Emna Ghorbel[a] and Faouzi Ghorbel
CRISTAL Laboratory, GRIFT Research Group ENSI, La Manouba University 2010, La Manouba, Tunisia

Keywords: Low Size Dataset, 3D Face Data Augmentation, PointNet++, Surface Morphing.

Abstract: This paper introduces the 3D Face Gravitational Morphing to elevate the performance of Deep Learning models in the realm of 3D facial classification. Addressing the constraints imposed by small-scale datasets, our approach amplifies intra-class variability while maintaining the semantic fidelity of 3D models. This is accomplished by generating shapes within the proximity of the original models in the context of shape space, facilitated by a curvature-based correspondence. The integration of Face Gravitational Morphing into the architecture is demonstrated through its application to the BU3DFE dataset for classification purposes. A comparative analysis reveals the method's relative performance, representing an initial step towards mitigating limitations in facial classification. Ongoing investigations are underway to refine and extend these promising results.

1 INTRODUCTION

In the field of computer vision, Convolutional Neural Networks (CNNs) have made significant strides in the recognition and classification of facial images. In fact, CNNs have demonstrated their effectiveness in a wide range of applications within facial analysis. Nevertheless, the performance of CNNs can decline when confronted with the challenges of small-scale datasets. In fact, the learning phase of neural network models demands copious data for convergence, and such datasets, in practical applications, often fall short. To address this limitation, several data augmentation methods have been proposed (Summers and Dinneen, 2019; Inoue, 2018; Kang et al., 2017; Zhong et al., 2020; Gatys et al., 2015; Konno and Iwazume, 2018; Bowles et al., 2018; Su et al., 2019; El-Sawy et al., 2016; Patel et al., 2019; Ciregan et al., 2012; Sato et al., 2015; Patel et al., 2019; Yin et al., 2019; Paulin et al., 2014; Chatfield et al., 2014; Xiao and Wachs, 2021; Blanz and Vetter, 1999; Tan et al., 2018; Cheng et al., 2019). These techniques can be classified into three distinct categories. The first one regroups geometric transformations such as rotation, scaling, and translation. These transformations were instrumental in introducing variability to training datasets, aiding models in learning invariant features across different orientations and scales. As advancements in geometric data augmentation transformations, some works have proposed non-uniform scaling, shearing, and perspective transformations.

In another hand, Data augmentations based on the integration of jittering and noise injection into point clouds (Xiao and Wachs, 2021) have proven effective in enhancing the robustness of models to noisy input.

Finally, Generative models, such as morphable models (Blanz and Vetter, 1999), Variational Autoencoders (VAEs) (Tan et al., 2018) and Generative Adversarial Networks (GANs) (Cheng et al., 2019), have been employed to generate new samples.

Despite their contributions, adapting these augmentation methods for facial classification can be especially challenging, given the nuanced complexities involved in this task. Often, these methods prove inadequate in capturing the nuances of intra-class variations, which may lead to the loss of meaning in the process.

In this paper, we introduce a novel data augmentation technique meticulously crafted to enhance Deep learning performance in 3D facial classification. Our method aims to augment intra-class variability while preserving the semantic integrity of 3D models by generating shapes within the neighborhood of the original models in terms of shape space. Therefore, We present the 3D Face data augmentation based on gravitational shape morphing and curvature-

[a] https://orcid.org/0000-0002-6179-1358

Ghorbel, E. and Ghorbel, F.
3D Face Data Augmentation Based on Gravitational Shape Morphing for Intra-Class Richness.
DOI: 10.5220/0012466700003636

In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICAART 2024) - Volume 3*, pages 1294-1299
ISBN: 978-989-758-680-4; ISSN: 2184-433X

based correspondence. In that context, we explore the theoretical foundations, integration of Face Gravitational Morphing into architecture, and its application to a relative small-scale 3D facial dataset namely BU3DFE (Yin et al., 2006).

2 DATA AUGMENTATION BASED ON 3D GRAVITATIONAL MORPHING

3D Facial Morphing, often employed in computer vision, is a technique that generate in-between faces from a source and a target one. Therefore, we propose a Morphing technique adapted to the case of 3D face data augmentation.

We present in the following the key components of our proposed 3D Face Morphing data augmentation method, which include the following steps; (1) the curvature-based sorting for correspondence between a pair of cloud points models belonging to a same class, (2) the Interpolation between the source and the target object while selecting only shapes in the gravitational shape space area of input data, and (3) a Data Cleaning post-processing based on the DBSCAN algorithm (Schubert et al., 2017) which is a non-linear machine learning clustering method.

2.1 Curvature-Based Sorting

As a first step for Morphing Shape, we propose a correspondence between 3D model vertices based on curvature measures. The idea consists in sorting the point clouds according to the distance between the point with the highest curvature value and each vertex. Let $\mathcal{S} = \mathbf{R}^{3N}/\mathbf{SE}(3) \times \mathbf{R}^{*}+$ be the Shape Space of 3D Surfaces where N are the number of vertices, $\mathbf{SE}(3)$ is the Special Euclidean Group in Three Dimensions and $\mathbf{R}^{*}+$ the multiplicative group of nonzero positive real numbers associated with scaling transformations. Given a point cloud $V \in \mathcal{S}$ represented as $\mathbf{v}_i = [x_i, y_i, z_i]$ with $\{i = 1, 2, ..., N\}$. The curvature K_i computation using the normal vector n_i would be given by,

$$\mathrm{K}_i = \frac{\|\mathbf{n}_i \times (\mathbf{v}_i - \mathbf{v}_0)\|}{\|\mathbf{v}_i - \mathbf{v}_0\|^2}$$

where $\mathbf{v}_0$ is the position vector of a reference point. The curvature values are then used to identify the point with the highest curvature, denoted as the new reference point,

$$\hat{\mathbf{v}}_0 = arg \max_{\mathbf{v} \in V} \|\mathrm{K}_i(\mathbf{v}_i)\|$$

Therefore, the curvature-based sorting is expressed by the distance between the new reference point with the highest curvature value and all other points of the cloud as follows,

$$\mathrm{d}_i = \|\mathbf{v}_i - \hat{\mathbf{v}}_0\|$$

The point cloud V is subsequently transformed based on this distance metric, resulting in an ordered arrangement that enhances the discernment of geometric features. Since, our work focuses on interpolating 3D faces having close characteristics, we judge this approach as a valuable preprocessing step for corresponding 3D Faces vertices.

Figure 1 illustrates an example of an original face model vertices from BU3DFE before sorting and after curvature-based sorting.

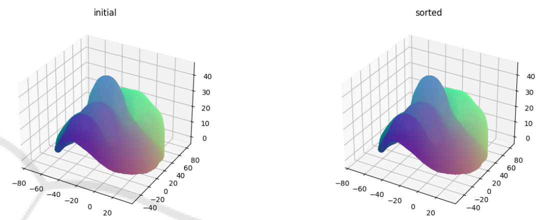

Figure 1: (a) Initial model vertices, (b) Model vertices after curvature-based sorting.

2.2 Gravitational 3D Face Morphing

We propose, in this part, to select pairs of 3D models denoted respectively A and B from the same class of a dataset $D = \{Class_1, ...Class_k\}$. Let $V = \{x^i, y^i, z^i | i \in \{1, ..., N\}\}$ be the normalized and curvature-sorted vertices associated to a model in the dataset. Therefore, the interpolation is applied on the two corresponding point clouds V^A and V^B with $t \in [0, 1]$ in order to obtain the in-between 3D clouds as follows,

$$V_{AB}(t) = (1 - t) \cdot V^A + t \cdot V^B$$

where each $V_{AB}(t)$ is a 3D generated point cloud representing a face model at time t.

Figure 2 illustrates two examples of the obtained face interpolation from a source and a target models from BU3DFE dataset to highlight the performance of the curvature-based sorting. In fact, when applying the sorting method, the in-between shapes relatively preserve the global aspect of input surfaces.

However, there are instances where shapes obtained do not strictly belong to the expected class, especially when dealing with complex shapes such as 3D faces. In response to this, we propose a novel approach termed "Gravitational morphing", where only the generated shapes within a ε-neighborhood, in the shape space, of the input elements are retained. This

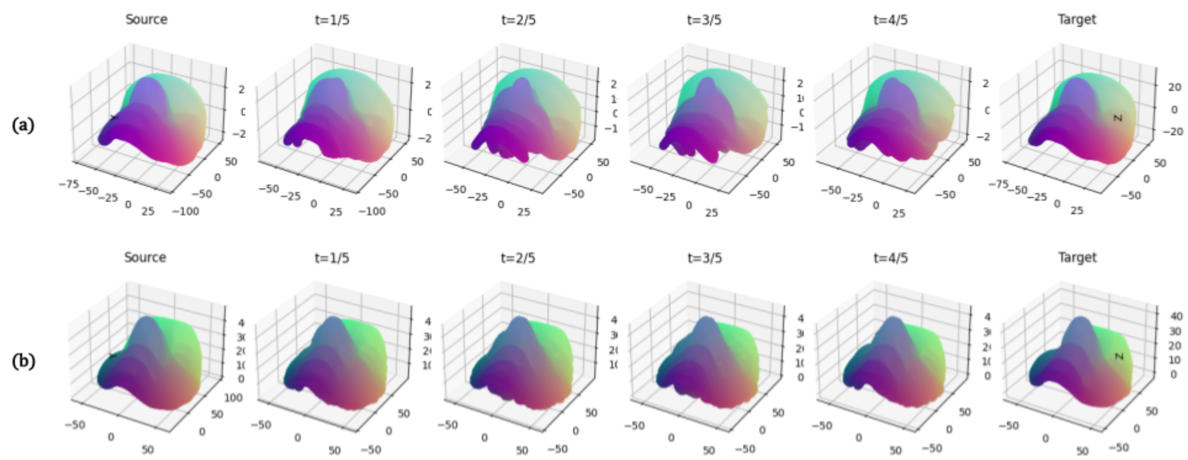

Figure 2: Two example of a morphing sequence (interpolation) between two point clouds (Faces from BU3DFE) Belonging to a same class : (a) before curvature-based sorting (b) with curvature-based sorting.

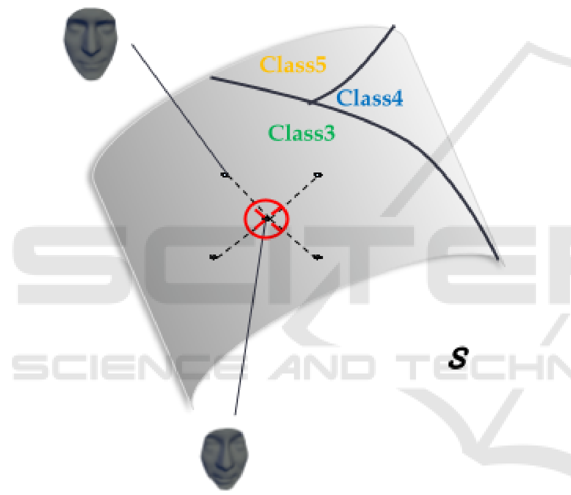

Figure 3: Overview of the 3D surface shape space $\mathcal{S}$: interpolation between Face Shape for Gravitational intra-class covering. The selected in-between shapes are within the red limitation, which correspond to the ε-neighborhood of a Face shape.

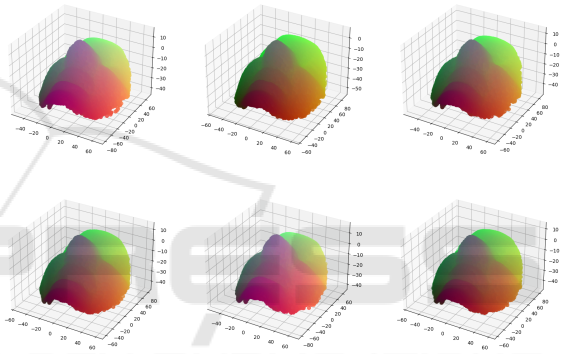

Figure 4: Examples of selected point clouds with the gravitational morphing method.

technique provides a more refined handling of complex 3D shapes, particularly in the case of data augmentation. In Figure 3, an overview of the proposed approach is illustrated, where only the intermediate shapes belonging to the ε-neighborhood (red circle) of input objects are selected. Therefore, we select the intermediate shapes, denoted as $V_{AB}(t)$ with $t \in [0,1]$, by finding those in the vicinity of the source and target shapes. The selection is determined by their proximity to the source and target within a ε-neighborhood as follows,

$$\mathbf{E}_f = \{V_{AB} | \|V_{AB} - V_A\| < \varepsilon \text{ or } \|V_{AB} - V_B\| < \varepsilon\}$$

with $\mathbf{E}_f$ is the set of selected faces. These selected intermediate shapes are then stored for data augmentation purpose. Figure 4 shows examples of selected point clouds with the gravitational morphing method. Consequently, a post-processing for cleaning the generated object is carried out in order to ensure the integrity and consistency of our augmented 3D face dataset. In fact, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Schubert et al., 2017) is applied on in-between clouds as an outlier eliminator. Note that the DBSCAN parameters are manually fixed for this study, with the intention of conducting a more in-depth investigation in future work.

The 3D Face Blending data augmentation pipeline is illustrated in Figure 5, showcasing the stages of our approach.

In the following, we propose to validate the proposed method qualitatively and quantitatively through PointNet++ (Qi et al., 2017) model in the case of a relative low-size 3D Face Dataset.

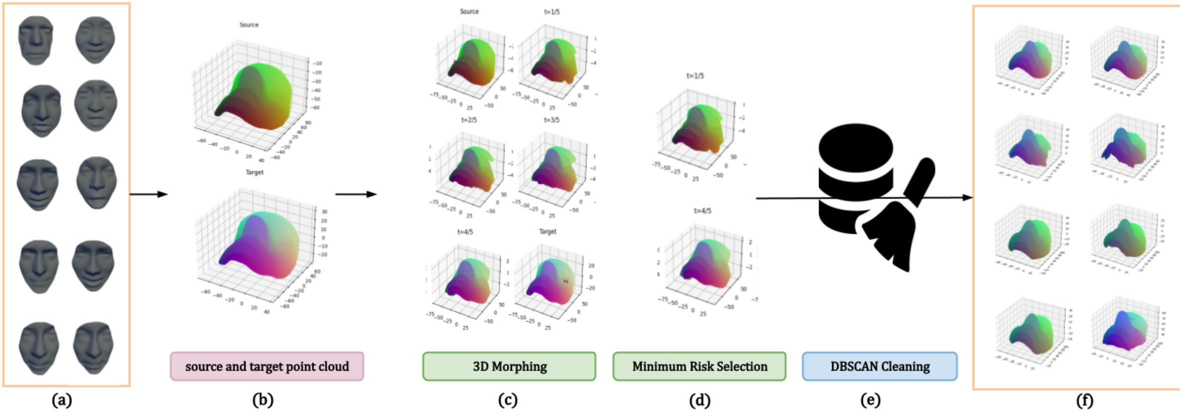

Figure 5: 3D Morphing Face Data Augmentation pipeline; (a) Original 3D model dataset, (b) Selection of a pair of model point clouds, (c) 3D point cloud morphing, (d) Selection of the closed generated shapes for Minimum Risk, (e) Application of DBSCAN algorithm on obtained cloud points as a cleaning process for eliminating outlier vertices, and (f) Augmented dataset.

3 EXPERIMENTS

In this part, qualitative and quantitative results are presented in order to validate the proposed method for enhancing face classification.

3.1 Datasets

The BU3DFE Dataset (Yin et al., 2006) encompasses a diverse group of 100 subjects, comprising 56% female and 44% male participants, with ages ranging from 18 to 70 years. The dataset reflects a broad spectrum of ethnic and racial backgrounds, including White, Black, East-Asian, Middle-East Asian, Indian, and Hispanic Latino. During the data collection process, each subject was recorded while performing seven distinct facial expressions in front of a 3D face scanner. Excluding the neutral expression, the six prototypic expressions (happiness, disgust, fear, anger, surprise, and sadness) were captured at four different intensity levels. In total, the dataset comprises 2,500 3D facial expression models, offering a rich resource for experiments and investigations in the field of facial expression analysis. For augmenting the BU3DFE with the Gravitational Morphing method, we propose to blend pairs of shapes belonging to a same subject with different level of a same expression.

3.2 Implementation Settings

We uniformly sample 2,048 points on the cloud faces and normalize them to be contained in a unit sphere, which is a standard setting (Qi et al., 2017). When performing the cloud morphing, we disregard elements from the point set having the highest cardinality. We use Python and implement PointNet++ (Qi et al., 2017) model and the gravitational morphing method using the TensorFlow and Keras framework. The model is trained for 20 epochs with a batch size of 32. The model is training on a single T4 GPU. For the training phase, we use the following configuration: (1) Loss Function: Sparse Categorical Cross entropy, (2) Optimizer: Adam with a learning rate of 10^{-3}, and (3) Metric: Sparse Categorical Accuracy.

3.3 Qualitative Results

Figure 6 illustrates various examples of the obtained face interpolation from a source and a target models from BU3DFE dataset. We observe that the obtained data conserve the meaning of the source and target shapes.

3.4 Quantitative Results

Since our goal is to validate the proposed approach, we conduct a simple comparative analysis between the ”Gravitational-Morphing-PointNet++” model, noise injection (Xiao and Wachs, 2021) data augmentation and the standard PointNet++ model (Xiang and Zhu, 2017) according to Sparse Categorical Accuracy and several metrics of which the weighted average (w.a.) Precision, the w.a. Recall, the w.a. F1 score. In the case of the low-size BU3DFE dataset, we observe in Table 1 that the model trained without any augmentation (No-aug) exhibits an accuracy of 29.53%, indicating its struggle to effectively learn from the dataset. The precision, recall, and F1

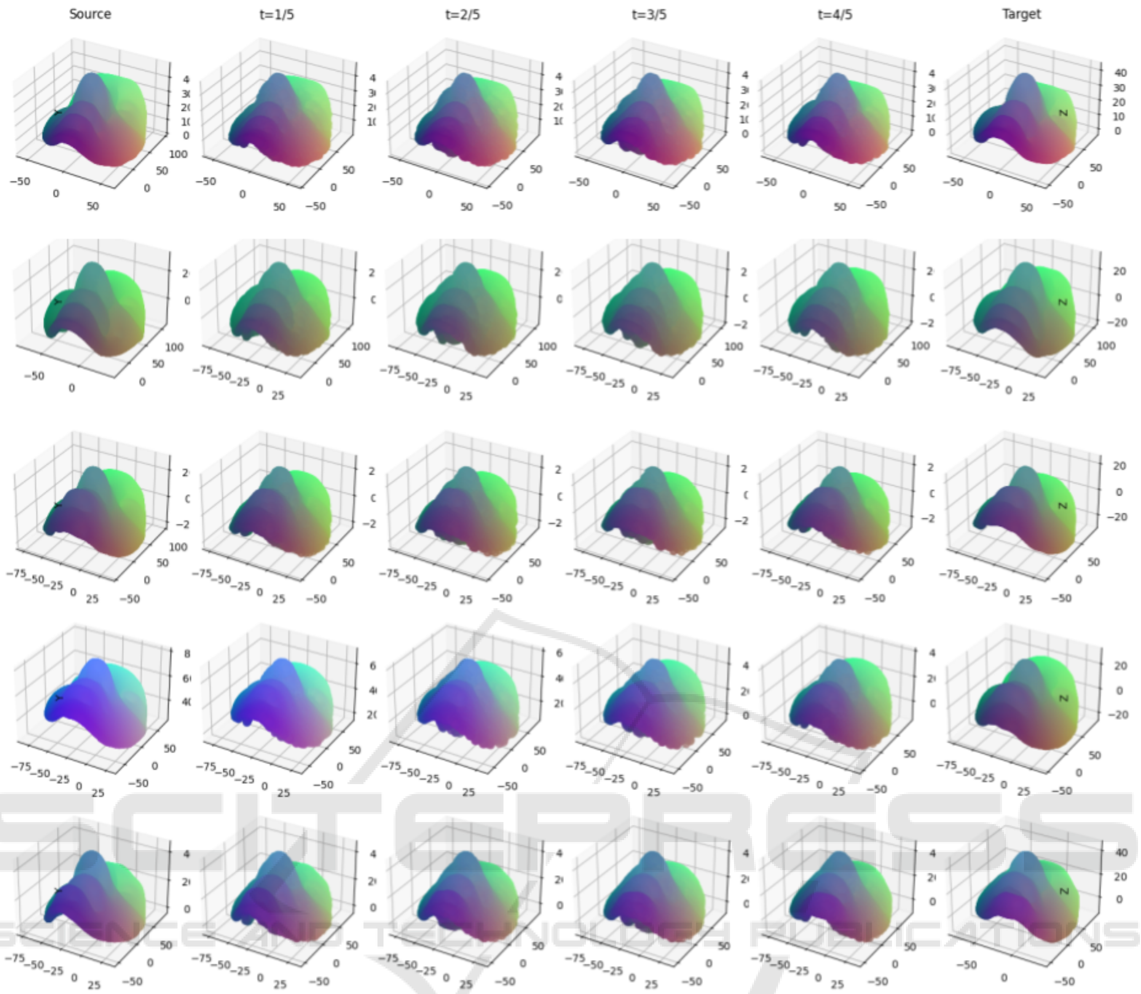

Figure 6: Examples of morphing sequence (interpolation) between two point clouds (Faces from BU3DFE) Belonging to a same class.

Table 1: Comparison of data augmentation methods with the PointNet++ model trained on the 3D Face dataset BU3DFE according to various performance metrics (20 epochs).

Method	Sp. C. Acc.(%)	Precision	Recall	F1 score
No-augmentation (Qi et al., 2017)	29.53	0.1922	0.0875	0.0801
Noise and jitter Injection (Xiao and Wachs, 2021)	33.65	0.1229	0.1370	0.1271
Gravitational-Morphing (ours)	**62.79**	**0.5511**	**0.4258**	**0.5279**

score are notably low, emphasizing the challenges in distinguishing between different facial expressions. Applying the noise injection method yields an improvement in accuracy (33.65%) compared to No-aug, but the values are still relatively low. The precision, recall, and F1 score show a modest increase, suggesting that the noise injection aids the model in capturing more nuanced patterns. Our proposed Gravitational-Morphing (GM) method achieves a significantly higher accuracy of 62.79%, surpassing both No-aug and Noise Injection. The precision, recall, and F1 score are notably enhanced, indicating the effectiveness of the gravitational morphing technique in improving the model's performance in 3D facial expression classification when dealing with low-size datasets. Nevertheless, it is crucial to emphasize that this work represents an initial exploration, and further scientific investigations are underway to build upon these preliminary and modest results.

Nevertheless, it is crucial to emphasize that this work represents an initial exploration, and further scientific investigations are underway to build upon these preliminary and modest results.

4 CONCLUSION

3D Face Gravitational Morphing emerges as an attractive solution to the challenges posed by low-size dataset in 3D facial classification. By prioritizing the augmentation of intra-class variability while preserving semantic integrity, the approach showcases promising results in enhancing the performance of Deep Learning models. The integration of Face Gravitational Morphing into the classification architecture, demonstrated through its application to the BU3DFE dataset, signifies a meaningful advancement in addressing the intricacies of 3D facial cloud point classification tasks. Our comparative analysis underscores the relative performance of the proposed method, establishing a foundation for further refinement and extension of these encouraging outcomes.

REFERENCES

Blanz, V. and Vetter, T. (1999). A morphable model for the synthesis of 3d faces; siggraph'99 proceedings of the 26th annual conference on computer graphics and interactive techniques.

Bowles, C., Chen, L., Guerrero, R., Bentley, P., Gunn, R., Hammers, A., Dickie, D. A., Hernández, M. V., Wardlaw, J., and Rueckert, D. (2018). Gan augmentation: Augmenting training data using generative adversarial networks. *arXiv preprint arXiv:1810.10863*.

Chatfield, K., Simonyan, K., Vedaldi, A., and Zisserman, A. (2014). Return of the devil in the details: Delving deep into convolutional nets. *arXiv preprint arXiv:1405.3531*.

Cheng, S., Bronstein, M., Zhou, Y., Kotsia, I., Pantic, M., and Zafeiriou, S. (2019). Meshgan: Nonlinear 3d morphable models of faces. *arXiv preprint arXiv:1903.10384*.

Ciregan, D., Meier, U., and Schmidhuber, J. (2012). Multicolumn deep neural networks for image classification. In *2012 IEEE conference on computer vision and pattern recognition*, pages 3642–3649. IEEE.

El-Sawy, A., Hazem, E.-B., and Loey, M. (2016). Cnn for handwritten arabic digits recognition based on lenet-5. In *International conference on advanced intelligent systems and informatics*, pages 566–575. Springer.

Gatys, L. A., Ecker, A. S., and Bethge, M. (2015). A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576*.

Inoue, H. (2018). Data augmentation by pairing samples for images classification. *arXiv preprint arXiv:1801.02929*.

Kang, G., Dong, X., Zheng, L., and Yang, Y. (2017). Patchshuffle regularization. *arXiv preprint arXiv:1707.07103*.

Konno, T. and Iwazume, M. (2018). Icing on the cake: An easy and quick post-learnig method you can try after deep learning. *arXiv preprint arXiv:1807.06540*.

Patel, V., Mujumdar, N., Balasubramanian, P., Marvaniya, S., and Mittal, A. (2019). Data augmentation using part analysis for shape classification. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1223–1232.

Paulin, M., Revaud, J., Harchaoui, Z., Perronnin, F., and Schmid, C. (2014). Transformation pursuit for image classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3646–3653.

Qi, C. R., Yi, L., Su, H., and Guibas, L. J. (2017). Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in neural information processing systems*, 30.

Sato, I., Nishimura, H., and Yokoi, K. (2015). Apac: Augmented pattern classification with neural networks. *arXiv preprint arXiv:1505.03229*.

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., and Xu, X. (2017). Dbscan revisited, revisited: why and how you should (still) use dbscan. *ACM Transactions on Database Systems (TODS)*, 42(3):1–21.

Su, J., Vargas, D. V., and Sakurai, K. (2019). One pixel attack for fooling deep neural networks. *IEEE Transactions on Evolutionary Computation*, 23(5):828–841.

Summers, C. and Dinneen, M. J. (2019). Improved mixed-example data augmentation. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1262–1270. IEEE.

Tan, Q., Gao, L., Lai, Y.-K., and Xia, S. (2018). Variational autoencoders for deforming 3d mesh models. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5841–5850.

Xiang, J. and Zhu, G. (2017). Joint face detection and facial expression recognition with mtcnn. In *2017 4th international conference on information science and control engineering (ICISCE)*, pages 424–427. IEEE.

Xiao, C. and Wachs, J. (2021). Triangle-net: Towards robustness in point cloud learning. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 826–835.

Yin, D., Lopes, R. G., Shlens, J., Cubuk, E. D., and Gilmer, J. (2019). A fourier perspective on model robustness in computer vision. *arXiv preprint arXiv:1906.08988*.

Yin, L., Wei, X., Sun, Y., Wang, J., and Rosato, M. J. (2006). A 3d facial expression database for facial behavior research. In *7th international conference on automatic face and gesture recognition (FGR06)*, pages 211–216. IEEE.

Zhong, Z., Zheng, L., Kang, G., Li, S., and Yang, Y. (2020). Random erasing data augmentation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 13001–13008.