Supplementary for Enhancing Artwork Style Clustering via Neural Representation Re-Alignment

Abhishek Dangeti, Pavan Gajula, Vikram Jamwal, Vivek Srivastava
TCS Research
INDIA
{abhishek.dangeti, pavanbhargav.gajula, vikram.jamwal, srivastava.vivek2}@tcs.com

A. Summary

This supplementary material is organized as follows:

- In Section B, we discuss the different design objectives based on which we design our feedback approaches.
- In Section C, we discuss and showcase the different ways a human clusters a small set of artworks.
- In Section D, we present additional results for the experiments performed in the Section 7.

B. Design Objectives for Feedback Approaches

We design our feedback approaches with the following objectives to ease the process of providing feedback and capture the stylistic similarity between the artworks:

B.1. Simple visual feedback mechanism

The style of an artwork is a subjective property and it is hard to define a mathematical or ML model for it. The metadata information that is available for an artwork such as art periods and genres might not be enough to understand the artistic style and provide feedback. In applications such as topic-based document clustering [2], the documents contain text that would include information that might help the user to categorize based on topic. For mobile applications clustering, for example [1], build a knowledge graph out of the category labels associated with each app helps the user categorize them. For artworks, such category labels are restricted to art movements and artists and it is difficult to build a fine-granular style-specific knowledge graph based on this type of categorical labels. Moreover, in most scenarios, the artworks collections actually lack consistent meta data information. Given these challenges, the objective is to facilitate preference feedback based just on the user's visual interpretation without requiring any additional knowledge about artworks.

B.2. Feedback on very few examples

It is cognitively difficult for a user to evaluate a cluster with large number of images. Displaying all the artworks present in a cluster is also not feasible for larger data sets as the user would have to compare across each artwork. Existing methods such as [7] which focuses on clustering similar faces displays 10 images per cluster to the user and the user is asked to merge the clusters with similar faces. However, the approach does not allow the user to remove any outliers that might be present in a cluster. Approaches like [6] utilize clustering and feedback for improving object-based image retrieval. In their approach, for a given user query image, their method retrieves 30 images that are similar to query image and asks the user to re-rank the retrievals based on relevance. Some earlier works on clustering large-scale datasets like [4, 8] proposed the sampling of representative subsets from these large datasets, their main focus was to tackle the non-linear time complexities of traditional clustering algorithms. However, in our case, the objective is to ease the process of giving preference feedback by taking it on just a small representative subset of a large dataset and then be able to cluster the original dataset utilizing the feedback.

B.3. Minimal iterative feedback

The earlier works on preference feedback for clustering require multiple iterations of feedback to produce the best results. In [3, 5, 7], the method requires the user to provide preference feedback for 5-6 iterations to obtain the best results. However, providing preference feedback for multiple iterations may prove cumbersome to the user. A few existing methods like [1] and [9] are able to cluster data points only after a single iteration but these clustering methods focus on specific applications that have readily available text-based metadata information. Since we primarily depend upon human feedback for style, our objective is to limit the number of preference feedback iterations to receive high quality input for achieving stylistic clustering of artworks.

C. Subjectivity of Artistic Style

The style associated with an artwork is dependent on the person analyzing the specific artwork. As style is subjective in nature, if we provide a few people with a collection of artworks and ask them to cluster the artworks, the clustering will most likely not be the same. This can be seen in Figure 1 where we present the clustering performed by two participants as well as the art movement ground truth on 72 samples picked from the WikiArt dataset. We observe that the clustering done by the two participants is different when compared to the art movement ground truth. Among the two participants, the two clusterings have a few clusters that are similar in style but the rest of the clusters are different. This further reinforces that artistic style is subjective in nature.

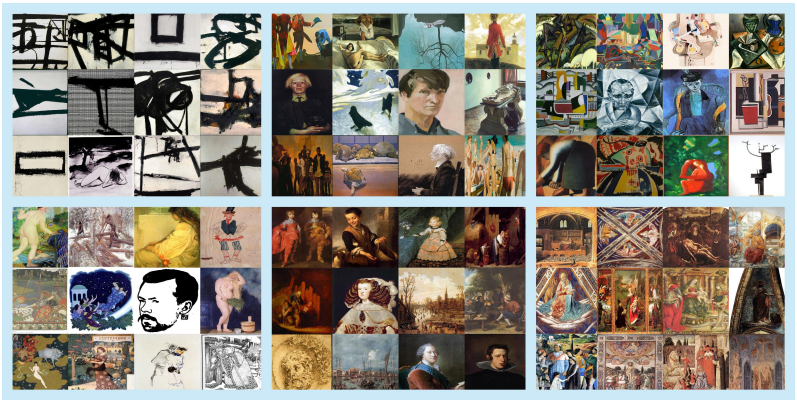
D. Additional Results

D.1. Normalized Mutual Information for all Approaches

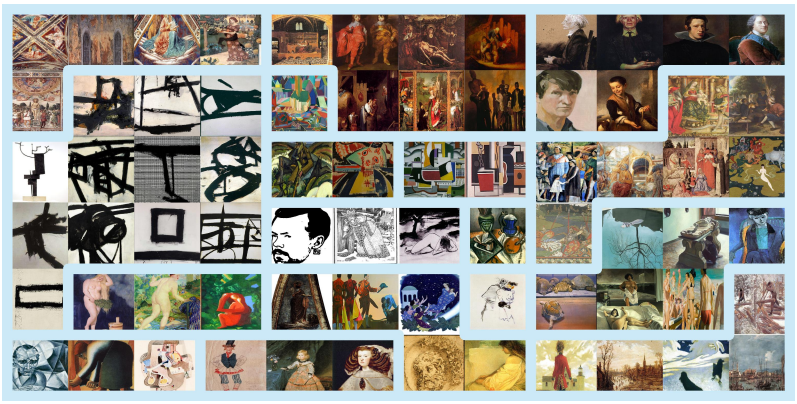
In Figure 2 and 3, we present the NMI scores for all the features on both the dataset which shows a similar trend as compared to the ARI metric across iterations.

D.2. Centroid Projection Approach with Accumulative Feedback and Encoder Training - Clustering Metrics (SC and CHI scores)

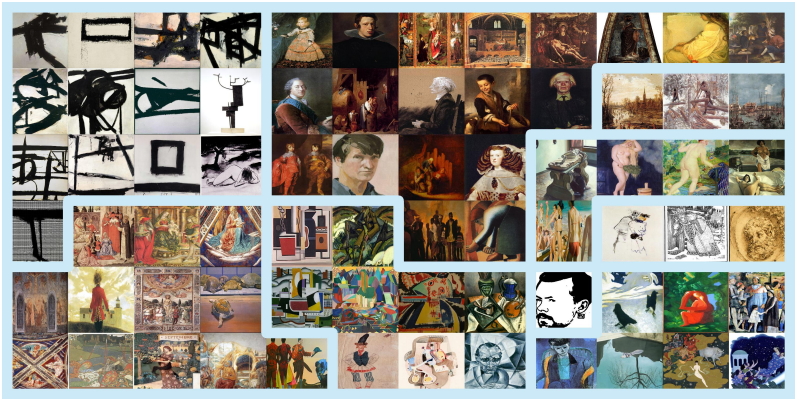
We report the SC and CHI scores in the Figures 4 and 5. Applying feedback on different features for both the datasets shows a drastic improvement in the metric scores after first iteration, with the same trend being seen for a few more subsequent iterations in a few features. However, after a point the scores fluctuate for all the features indicating that the clusters that are getting formed are not well dense and sparse. This is due to the feedback influencing the style capturing ability of features and not their cluster formation ability.



(a) Ground Truth clusters based on Art Movement labels in the WikiArt Dataset. Samples shown in these six clusters belong to the following Art Movements: *Action Painting*,*Contemporary Realism*, *Cubism*, *Art Nouveau Modern*, *Baroque*, *Early Renaissance* (Order: left to right, top to bottom)

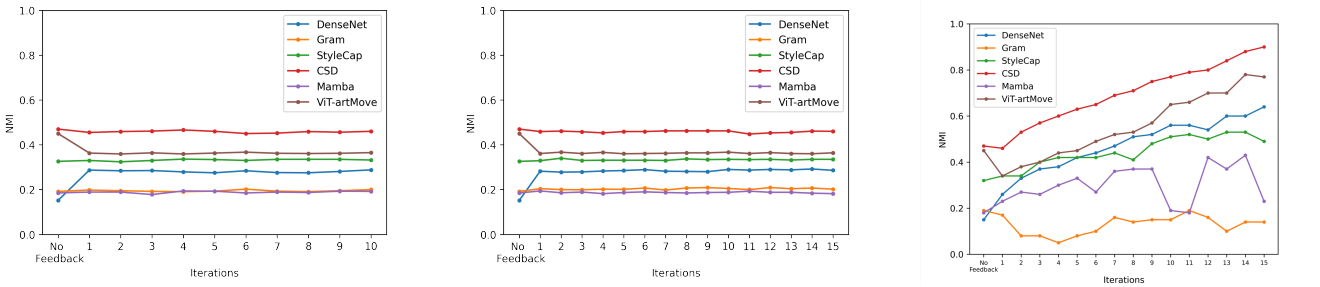


(b) Clustering done by Participant 1



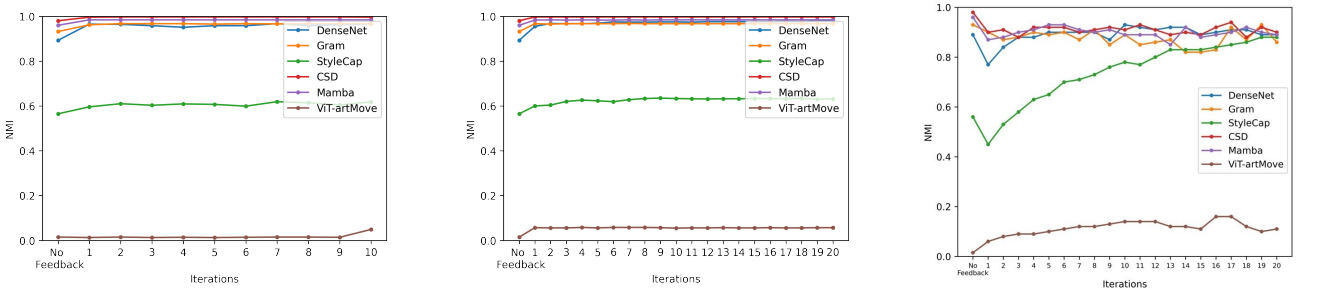
(c) Clustering done by Participant 2

Figure 1. Collages showing three different clusterings based on style: Art Movement Ground Truth, Participant 1 and Participant 2, highlighting the differences in the definitions and the subjectivity involved in human interpretations of artistic styles. 72 artworks were picked from the WikiArt dataset and shown to each participant.



(a) Simple Feedback (b) Accumulative Feedback (c) Accumulative Feedback and Encoder Training

Figure 2. Normalized Mutual Information results for the three approaches on a set of initial clusters obtained through kmeans for the **WikiArt-4k** dataset.



(a) Simple Feedback (b) Accumulative Feedback (c) Accumulative Feedback and Encoder Training

Figure 3. Normalized Mutual Information results for the three approaches on a set of initial clusters obtained through kmeans for the **Curated-4k** dataset.

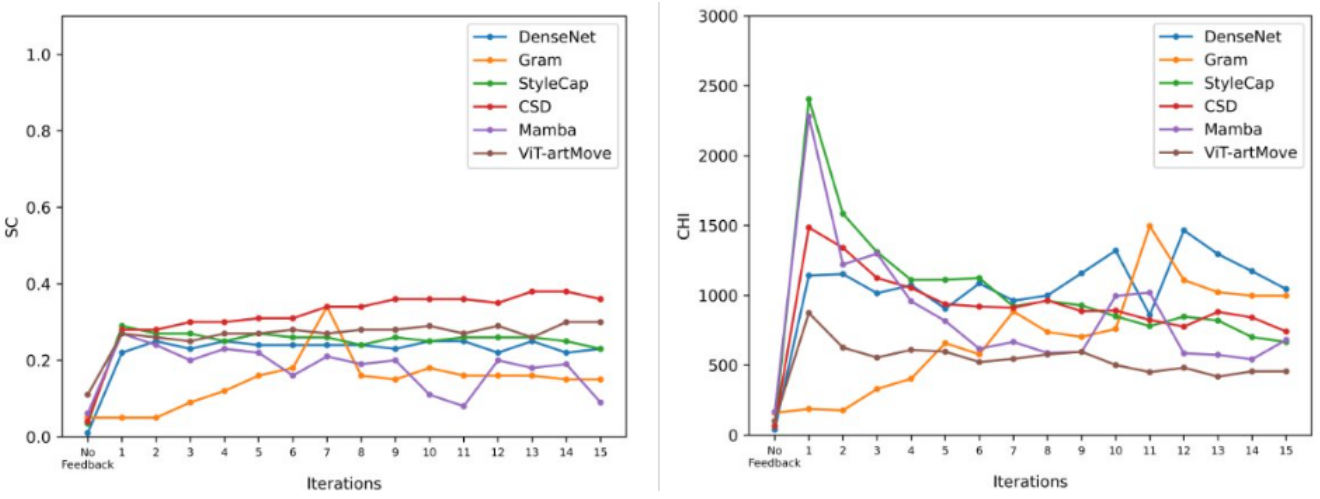


Figure 4. Quantitative clustering-metric results (SC and CHI) for **Centroid Projection with Accumulative Feedback and Encoder Training** for different features on **WikiArt-4k** dataset. The metric scores were tested mainly on the third approach as we are modifying the representations, the cluster formation ability of these representations might vary.

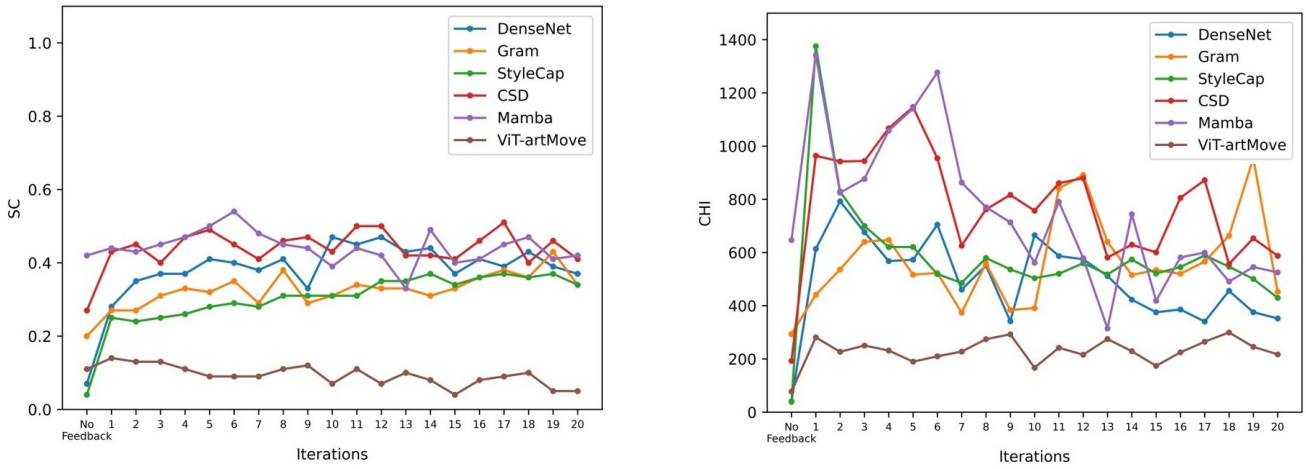


Figure 5. Quantitative clustering-metric results (SC and CHI) for **Centroid Projection with Accumulative Feedback and Encoder Training** for different features on **Curated-4k** dataset. The metric scores were tested mainly on the third approach as we are modifying the representations, the cluster formation ability of these representations might vary.

References

[1] Shuo Chang, Peng Dai, Lichan Hong, Cheng Sheng, Tianjiao Zhang, and Ed Chi. Appgrouper: Knowledge-based interactive clustering tool for app search results. pages 348–358, 2016.

[2] Douglass R. Cutting, David R. Karger, Jan O. Pedersen, and John W. Tukey. Scatter/gather: a cluster-based approach to browsing large document collections. In *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, page 318–329, New York, NY, USA, 1992. Association for Computing Machinery.

[3] Mennatallah El-Assady, Rita Sevastjanova, Fabian Sperrle, Daniel A. Keim, and Christopher M. Collins. Progressive learning of topic modeling parameters: A visual analytics framework. *IEEE Transactions on Visualization and Computer Graphics*, 24:382–391, 2018.

[4] Sudipto Guha, Rajeev Rastogi, and Kyuseok Shim. Cure: An efficient clustering algorithm for large databases. *ACM Sigmod record*, 27(2):73–84, 1998.

[5] Chengcui Zhang, Liping Zhou, Wen Wan, Jeffrey Birch, and Wei-Bang Chen. An image clustering and feedback-based retrieval framework. *IJMDEM*, 1:55–74, 2010.

[6] Chengcui Zhang, Liping Zhou, Wen Wan, Jeffrey Birch, and Wei-Bang Chen. An image clustering and feedback-based retrieval framework. *IJMDEM*, 1:55–74, 2010.

[7] Liyan Zhang, Dmitri Kalashnikov, and Sharad Mehrotra. Context-assisted face clustering framework with human-in-the-loop. *International Journal of Multimedia Information Retrieval*, 3:69–88, 2014.

[8] Tian Zhang, Raghu Ramakrishnan, and Miron Livny. Birch: an efficient data clustering method for very large databases. *ACM sigmod record*, 25(2):103–114, 1996.

[9] İnanç Arın, Mert Kemal Erpam, and Yücel Saygın. I-twec: Interactive clustering tool for twitter. *Expert Systems with Applications*, 96:1–13, 2018.