CSE515 - Multimedia and Web Databases - Fall 2023

Project Report - Phase 3
Group No. - 5

Kaushik Ravishankar	Madhura Bhalchandra Wani	Mohankrishna Chandrashekar
Niraj Sonje	Pavan Rathnakar Shetty	Pranav Rameshwar Borikar

ABSTRACT

In the third phase of our Multimedia and Web Databases project, we will be concentrating on the in-depth analysis and utilization of the Caltech 101 image dataset. Task 0a and Task 0b set the foundation by computing the "inherent dimensionality" for even-numbered images and their unique labels. This initial examination uncovers the underlying structural characteristics of the dataset, paving the way for more thorough analyses. Utilizing this understanding, Task 1 uses latent semantic analysis to derive k latent semantics for each unique label in even-numbered images. Thereafter, predictions for odd-numbered images are made using distances calculated under label-specific latent semantics. Precision, recall, F1-score, and overall accuracy metrics are extensively reported, offering a comprehensive evaluation of the classification performance.

Following that, Task 2 introduces clustering through the DBScan algorithm, which identifies the c most significant clusters that are linked to each unique label in even-numbered Caltech 101 images. These clusters are visualized both in a 2-dimensional MDS space and as groups of image thumbnails, offering an intuitive representation of the dataset's inherent grouping. The label-specific clusters are then used to predict odd-numbered images, and performance metrics are reported for each label, contributing to a nuanced understanding of the clustering impact on classification accuracy. Task 3 deepens the exploration into classification techniques, implementing m-NN, decision-tree, and PPR classifiers for even-numbered images. The selected classifier is then used to make subsequent predictions for odd-numbered images, and detailed performance metrics are exhibited. This task offers valuable insights into the efficacy of different classifiers, aiding in the selection of appropriate models for real-world applications.

Tasks 4a and 4b introduce a Locality Sensitive Hashing (LSH) tool and a corresponding image search algorithm. For even-numbered Caltech101 images, the LSH tool constructs an in-memory index structure. The image search algorithm allows users to visualize the t most

similar images to a given query image, with unique and overall numbers of images considered reported. Tasks 4a and 4b introduce a Locality Sensitive Hashing (LSH) tool and a corresponding image search algorithm. For even-numbered Caltech101 images, the LSH tool constructs an in-memory index structure based on Euclidean distance. The image search algorithm allows users to visualize the t most similar images to a given query image, with unique and overall numbers of images considered reported. Ultimately, Task 5 integrates relevance feedback systems, including an SVM-based system and a probabilistic system. These systems enable users to tag and filter results from the image search, enhancing the project's interactive and adaptive nature.

KEYWORDS

Caltech101 Dataset, Inherent Dimensionality, Latent Semantic Analysis, Clustering, Classification, Locality Sensitive Hashing, SVM-Based Relevance Feedback, Probabilistic Relevance Feedback, Precision, Recall, F1-Score, DBScan Algorithm, Decision-Tree Classifier, Personalized PageRank, Image Search Algorithm, User Interaction, Data Visualization.

INTRODUCTION

In the third and final phase of the project, we explore further into the Caltech101 image dataset. The tasks encompass fundamental investigations into the dataset's inherent dimensionality, analyses of latent semantics, an unconventional approach to clustering, diverse classification methodologies, implementation of locality sensitive hashing. The introduction of relevance feedback systems further enhances user interaction. Collectively, these endeavors contribute to a refined understanding of the dataset, fostering the development of an adaptive and robust image recommendation system.

Key Terminology:

- *Inherent Dimensionality*: The inherent structural characteristics defining the complexity and dimensions of the Caltech101 image dataset.
- *Latent Semantic Analysis* (LSA): A technique revealing underlying relationships and patterns within the dataset by analyzing the latent semantics associated with unique labels.
- *DBSCAN Algorithm*: Density-based spatial clustering of applications with noise, utilized for identifying significant clusters in the dataset based on spatial proximity.
- *Classification Metrics*: Precision, Recall, and F1-Score, providing quantitative measures of classification performance.
- *Locality Sensitive Hashing (LSH)*: A method for approximate nearest neighbor search, facilitating efficient retrieval of similar images based on Euclidean distance.
- *Relevance Feedback Systems*: Interactive systems allowing users to tag and refine search results, enhancing the adaptability and precision of image retrieval.

- *Multidimensional Scaling (MDS)*: Visualization technique that maps high-dimensional data points onto a lower-dimensional space for interpretation purposes.
- *Feature Space*: The abstract space where images are represented based on selected features, crucial for classification and clustering algorithms.
- *Decision-Tree Classifier*: A model for classification based on a tree-like graph of decisions, aiding in the identification of image labels.
- *PPR (Personalized PageRank) Classifier*: A classifier utilizing the personalized PageRank algorithm for personalized and context-aware image classification.
- *Locality Sensitive Hashing (LSH) Tool*: An indexing structure for efficient approximate nearest neighbor search in high-dimensional spaces.
- *SVM-Based Relevance Feedback*: Support Vector Machine-based system for refining image search results through user feedback.
- *Probabilistic Relevance Feedback*: A system incorporating probabilistic models for enhanced precision in refining image search results.
- *Image Search Algorithm*: An algorithm facilitating the retrieval of similar images based on the LSH index structure and a user-specified visual model.
- *User Interaction*: The dynamic engagement between users and the multimedia and web databases system, involving tagging, feedback, and adaptive refinement of search results.

Goal Description:

The primary objective of this phase is to improve our understanding of the Caltech101 image dataset and the nuances of image retrieval through advanced techniques, including latent semantic analysis, clustering, and classification. Leveraging advanced techniques such as Locality Sensitive Hashing (LSH) makes image retrieval more efficient, while the incorporation of relevance feedback systems betters user interaction and refines search results. The overarching goal is to contribute to the development of a robust image recommendation system, equipped with advanced analytical tools and adaptive features, fostering a deeper understanding of complex image datasets.

Task 0a: Inherent Dimensionality Computation:

- Compute and print the inherent dimensionality associated with even-numbered Caltech101 images.

Task 0b: Label-specific Inherent Dimensionality:

- Compute and print the inherent dimensionality associated with each unique label of the even-numbered Caltech101 images.

Task 1: Latent Semantic Analysis and Classification:

- Derive k latent semantics for each unique label in even-numbered Caltech101 images. Predict labels for odd-numbered images using label-specific latent semantics. Output per-label precision, recall, F1-score, and overall accuracy values.

Task 2: Clustering and Visualization:

- Compute c most significant clusters associated with each unique label in even-numbered Caltech101 images using DBScan. Visualize clusters in a

2-dimensional MDS space and as groups of image thumbnails. Predict labels for odd-numbered images using label-specific clusters. Output per-label precision, recall, F1-score, and overall accuracy values.

Task 3: Classifier Implementation:

- Create m-NN, decision-tree, and PPR classifiers for even-numbered Caltech101 images. Predict labels for odd-numbered images using the selected classifier. Output per-label precision, recall, F1-score, and overall accuracy values.

Task 4a: Locality Sensitive Hashing (LSH) Tool:

- Implement an LSH tool (for Euclidean distance) to create an in-memory index structure for even-numbered Caltech101 images.

Task 4b: Image Search Algorithm using LSH:

- Implement an image search algorithm using the LSH index structure. Visualize the t most similar images to a given query image. Output the numbers of unique and overall images considered during the process.

Task 5: Relevance Feedback Systems:

- Implement SVM-based and probabilistic relevance feedback systems. Enable users to tag and refine image search results. Return a new set of ranked results based on user feedback.

Assumptions:

In undertaking the tasks of Phase 3, certain assumptions guide our approach:

- We assume that the Caltech101 image dataset exhibits inherent structural characteristics resembling the data labels, which can be well-represented by latent semantic analysis and clustering techniques.
- Additionally, we assume that the DBScan algorithm is appropriate for identifying meaningful clusters in the dataset, contributing to improved classification.
- The efficacy of classification models, including m-NN, decision-tree, and PPR classifiers, is assumed to be dependent on the feature space that is selected.
- Furthermore, the successful implementation of Locality Sensitive Hashing (LSH) is predicated on the assumption that Euclidean distance provides a meaningful metric for image retrieval.
- Relevance feedback systems are assumed to enhance search precision, relying on truthful and accurate user interactions for effective refinement.

PROPOSED SOLUTION/IMPLEMENTATION

Task 0

Specification

The program outputs the inherent dimensionality according to the requirement specified in the problem statement. In task 0a, the inherent dimensionality for all the even index images is calculated and in task 0b, the inherent dimensionality for all the unique labels present in the database is calculated.

Description

1. The feature vectors of the images of which the inherent dimensionality is to be calculated is given as input.
2. To calculate the inherent dimensionality, the following procedure is followed:
    a. The mean of the data is calculated and the data is centered by subtracting the mean from the data
    b. The next step is to normalize the data by dividing the data by the standard deviation
3. If the dimension of the data is more than three, the data is converted to two dimensions as for calculating the dot product the shape of the data should be 2.
4. Next, the covariance matrix is calculated by taking the dot product of the transpose of the normalized data and the normalized data. It is divided by reshaped_normalized_data.shape[0] - 1 in order to bring the values in the range of 0 to 1
5. Further, the eigenvalues and eigenvectors are computed
6. The significant eigenvalues are calculated by comparing the eigenvalues which are greater than the threshold.
7. The inherent dimensionality is the total number of significant eigenvalues.

Design

PCA:
In this task, PCA is used to compute the inherent dimensionality. Here, PCA is used as it works by extracting the significant bits of information from all features in the original dataset and creates lesser numbers of new features.

Threshold:

The threshold selected for computing the inherent dimensionalities is 0.5, threshold values from 0.1 were tried, and 0.5 gave a more appropriate inherent dimensionality. The threshold values from 0.1 to 0.4 included almost all the eigenvalues, which means that all the eigenvalues contribute most significantly to the data's structure.

Feature Space:
We selected the feature space obtained by the FC-layer of resnet50 as it provided more accurate results based on the observations made in the previous phases.

Output

Task 0a with threshold=0.5:

```
Inherent dimensionality associated with the even numbered images:  260
```

Task 0b with threshold=0.5:

```
Inherent dimensionality associated with label 0 : 140
Inherent dimensionality associated with label 1 : 134
Inherent dimensionality associated with label 2 : 75
Inherent dimensionality associated with label 3 : 182
Inherent dimensionality associated with label 4 : 27
Inherent dimensionality associated with label 5 : 186
Inherent dimensionality associated with label 6 : 20
Inherent dimensionality associated with label 7 : 20
Inherent dimensionality associated with label 8 : 22
Inherent dimensionality associated with label 9 : 26
Inherent dimensionality associated with label 10 : 22
Inherent dimensionality associated with label 11 : 16
Inherent dimensionality associated with label 12 : 63
Inherent dimensionality associated with label 13 : 48
Inherent dimensionality associated with label 14 : 20
Inherent dimensionality associated with label 15 : 42
Inherent dimensionality associated with label 16 : 44
Inherent dimensionality associated with label 17 : 24
Inherent dimensionality associated with label 18 : 21
Inherent dimensionality associated with label 19 : 60
Inherent dimensionality associated with label 20 : 23
Inherent dimensionality associated with label 21 : 28
Inherent dimensionality associated with label 22 : 30
Inherent dimensionality associated with label 23 : 53
Inherent dimensionality associated with label 24 : 22
Inherent dimensionality associated with label 25 : 34
Inherent dimensionality associated with label 26 : 35
Inherent dimensionality associated with label 27 : 34
Inherent dimensionality associated with label 28 : 24
Inherent dimensionality associated with label 29 : 25
Inherent dimensionality associated with label 30 : 27
Inherent dimensionality associated with label 31 : 33
Inherent dimensionality associated with label 32 : 25
Inherent dimensionality associated with label 33 : 31
Inherent dimensionality associated with label 34 : 33
Inherent dimensionality associated with label 35 : 37
Inherent dimensionality associated with label 36 : 31
Inherent dimensionality associated with label 37 : 25
Inherent dimensionality associated with label 38 : 31
Inherent dimensionality associated with label 39 : 42
Inherent dimensionality associated with label 40 : 32
Inherent dimensionality associated with label 41 : 33
Inherent dimensionality associated with label 42 : 21
Inherent dimensionality associated with label 43 : 16
Inherent dimensionality associated with label 44 : 16
Inherent dimensionality associated with label 45 : 25
Inherent dimensionality associated with label 46 : 48
Inherent dimensionality associated with label 47 : 49
Inherent dimensionality associated with label 48 : 20
Inherent dimensionality associated with label 49 : 26
Inherent dimensionality associated with label 50 : 43
Inherent dimensionality associated with label 51 : 39
Inherent dimensionality associated with label 52 : 15
Inherent dimensionality associated with label 53 : 31
Inherent dimensionality associated with label 54 : 42
Inherent dimensionality associated with label 55 : 56
Inherent dimensionality associated with label 56 : 29
Inherent dimensionality associated with label 57 : 40
Inherent dimensionality associated with label 58 : 38
Inherent dimensionality associated with label 59 : 19
Inherent dimensionality associated with label 60 : 32
Inherent dimensionality associated with label 61 : 21
Inherent dimensionality associated with label 62 : 19
Inherent dimensionality associated with label 63 : 41
Inherent dimensionality associated with label 64 : 15
Inherent dimensionality associated with label 65 : 37
Inherent dimensionality associated with label 66 : 27
Inherent dimensionality associated with label 67 : 16
Inherent dimensionality associated with label 68 : 19
Inherent dimensionality associated with label 69 : 22
Inherent dimensionality associated with label 70 : 18
Inherent dimensionality associated with label 71 : 22
Inherent dimensionality associated with label 72 : 25
Inherent dimensionality associated with label 73 : 16
Inherent dimensionality associated with label 74 : 28
Inherent dimensionality associated with label 75 : 39
Inherent dimensionality associated with label 76 : 28
Inherent dimensionality associated with label 77 : 24
Inherent dimensionality associated with label 78 : 19
Inherent dimensionality associated with label 79 : 30
Inherent dimensionality associated with label 80 : 19
Inherent dimensionality associated with label 81 : 41
Inherent dimensionality associated with label 82 : 27
Inherent dimensionality associated with label 83 : 17
Inherent dimensionality associated with label 84 : 31
Inherent dimensionality associated with label 85 : 21
Inherent dimensionality associated with label 86 : 42
Inherent dimensionality associated with label 87 : 29
Inherent dimensionality associated with label 88 : 31
Inherent dimensionality associated with label 89 : 16
Inherent dimensionality associated with label 90 : 42
Inherent dimensionality associated with label 91 : 23
Inherent dimensionality associated with label 92 : 42
Inherent dimensionality associated with label 93 : 37
Inherent dimensionality associated with label 94 : 111
Inherent dimensionality associated with label 95 : 18
Inherent dimensionality associated with label 96 : 28
Inherent dimensionality associated with label 97 : 16
Inherent dimensionality associated with label 98 : 27
Inherent dimensionality associated with label 99 : 19
Inherent dimensionality associated with label 100 : 29
```

Interpretation of Output:
The inherent dimensionality of all the even-numbered images in the Caltech101 dataset taken together is 260. This value suggests that a moderate number of features are necessary to capture the essential information for classifying and identifying the objects in the images.
A higher inherent dimensionality would suggest a good amount of complexity. Conversely, if it was lower, it would imply that the feature space is simplistic or that the feature extraction process is not capturing the full range of information present in the feature space.
The inherent dimensionality for each label can be similarly interpreted as well.

Task 1

Specification

In this task, given a value k:

- We calculate the label-specific latent semantics of even-numbered images from the dataset.
- We use these latent semantics to find the most likely labels for the odd numbered images in our dataset.

Description

Extracting Label Specific Semantics:

1. The user provides a value k which determines the dimensions of latent semantics to be extracted.
2. We calculate the label vectors based on image features (Fully Connected layer of resnet50 in our case)
3. These label vectors are used to calculate the label specific latent semantics using dimension reduction technique (SVD in our case)
4. We get the label-semantic, semantic-core and semantic-feature matrices

Finding most likely Labels:

1. For odd numbered images we use the semantic feature matrix and get a vector of dimension k for that image
2. We compare this vector with the label specific semantics and calculate euclidean distance between them.
3. The distance is sorted ascendingly to find the closest matching label.
4. This process is repeated for all images

Design

Selecting feature model:

- We have selected the FC-Layer of Resnet as our feature model as it provides features formed by deep neural networks and extract the best characteristics within an image.
- We also tried with avgpool-layer and color moments to verify our understanding but the results were not as good.

Selecting Dimension Reduction Technique:

- We tried multiple dimension reduction techniques to increase our accuracy and SVD proved to be the best.
- SVD also stores the semantic-feature matrix which makes it easy to compare new images with existing semantics.

Comparing Odd images with existing label semantics:

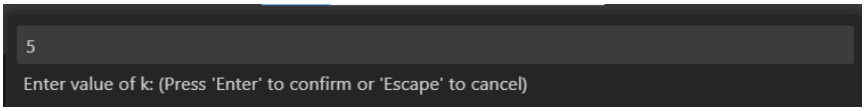
- For an odd image we do a matrix multiplication of the image features (m) and semantic-feature(mXk) matrices stored in our latent semantics and convert the feature dimensions(m) to K for that image.
- This reduced dimension (K) of the image features is used for comparison with the label-semantic matrix.

Calculating F1 score, Recall, Precision and Accuracy:

- If predicted label matches true label of image it is added in true positive else if the do not match its added in false negative.
- If the predicted label does not match the true label then the predicted label is added to the false positive. True negatives are calculated by subtracting summation of true positives, false negatives and false positives of a label.
- Once we have these values for all labels, we can calculate F1, Precision Recall and Accuracy using their respective formulas.

Output

K = 5




```
Precision for Label 0: 0.6666666666666666
Recall for Label 0: 0.4608294930875576
F1 Score for Label 0: 0.544959128065395

Precision for Label 1: 0.604
Recall for Label 1: 0.6926605504587156
F1 Score for Label 1: 0.6452991452991452

Precision for Label 2: 0.6764705882352942
Recall for Label 2: 0.23
F1 Score for Label 2: 0.3432835820895523

Precision for Label 3: 0.9027027027027027
Recall for Label 3: 0.41854636591478694
F1 Score for Label 3: 0.5719178082191781

Precision for Label 4: 0.4
Recall for Label 4: 0.2222222222222222
F1 Score for Label 4: 0.2857142857142857

Precision for Label 5: 0.9180327868852459
Recall for Label 5: 0.56
F1 Score for Label 5: 0.6956521739130435
...
Precision for Label 100: 0.3584905660377358
Recall for Label 100: 0.6333333333333333
F1 Score for Label 100: 0.45783132530120485
Overall Accuracy: 0.3727524204702628
```

For K=5
Overall accuracy = 37.28%.

```
Precision for Label 0: 0.9203980099502488
Recall for Label 0: 0.8525345622119815
F1 Score for Label 0: 0.8851674641148325

Precision for Label 1: 0.8805309734513275
Recall for Label 1: 0.9128440366972477
F1 Score for Label 1: 0.8963963963963965

Precision for Label 2: 0.9900990099009901
Recall for Label 2: 1.0
F1 Score for Label 2: 0.9950248756218906

Precision for Label 3: 1.0
Recall for Label 3: 0.9799498746867168
F1 Score for Label 3: 0.9898734177215189

Precision for Label 4: 0.9629629629629629
Recall for Label 4: 0.9629629629629629
F1 Score for Label 4: 0.9629629629629629

Precision for Label 5: 1.0
Recall for Label 5: 0.95
F1 Score for Label 5: 0.9743589743589743
...
Precision for Label 100: 0.8928571428571429
Recall for Label 100: 0.8333333333333334
F1 Score for Label 100: 0.8620689655172413
Overall Accuracy: 0.8978792070078377
```

For K = 150
Accuracy - 89.78%

Interpretation of Output:
The accuracy is low when we take a very small value of k which is nowhere close to the inherent dimensionality of the feature space.
With a larger value of k closer to the inherent dimensionality we get an accuracy of ~90%. This in turn benefits our label classification.

Task 2

Specification

This task involves clustering the even-numbered images in each label using the DBSCAN algorithm to get a specific number of clusters. Then, these clusters are visualized in a 2-D MDS space, and image thumbnails are also shown for these clusters. Finally, labels are predicted for every odd-numbered image using these clusters, per-label metrics are calculated as well as an overall accuracy score.

Description

1. First, the desired number of clusters ‘c’ is taken as input.

2. The feature space selected for this task is the output of the fully-connected (ResNet-FC-1000) layer of the pre-trained ResNet50 model. Euclidean distance measure is used.

Clustering:

1. For every label, DBSCAN clustering is done with heuristically estimated parameters.
2. The DBSCAN algorithm is implemented:
    a. For each point, check if it is a core point or noise point based on the number of neighbors in its region.
    b. For every core point, find all its connected components, i.e. neighbors in the region and assign them to the cluster of the original core point.
3. In order to get exactly c clusters, the parameters for DBSCAN have to be chosen appropriately.
4. The MinPts and Epsilon parameters are estimated by a grid search-like method.
    a. To find the range for this grid search, first we consider MinPts. We iterate over the range of all possible MinPts starting from twice the value of c, or the number of images in the given label, whichever is smaller, and decrease till we approach 2.
    b. For each value of MinPts that we try, we find the average and minimum k-nearest-neighbor distance for the label's data, and divide this range into 100 equal demarcation points, and use these as epsilon values.
    c. For each of these combinations, we check for the desired number of clusters and minimal noise points.
    d. For each label, the combination of parameters and the clustering that gives us the closest to the desired number of clusters and minimum noise points is stored.

Visualization:

1. For the stored clusterings for each label, we project this space onto a 2-dimensional MDS space using the classical MDS approach, or Torgerson scaling [2].
2. Clusters are visualized as differently colored point clouds in this space.
3. For the image grouping, this space is scaled up, and one image per cluster is shown at the centroid of the cluster in the MDS space.

Prediction and metrics:

1. First, all the core points across all the cluster sets are aggregated into one list.
2. For each image to be classified, its distance from all the core points is calculated, and the closest c core points are polled and the majority label is taken as the label predicted for the image.
3. The precision, recall, F1-score metrics for each label, as well as the overall accuracy score are calculated.

Design

Since clustering is done per-label, we decided not to go for dimensionality reduction as that would shrink the already small dataset even further and possibly increase the error probability. Also, we decided on the avgpool-layer feature model with Euclidean distance measure as it was a good compromise between performance and accuracy. ResNet-SoftMax and HOG feature spaces require KL-divergence and cosine distance, which are not metric measures, and as such, are not suitable for DBSCAN, which is a neighborhood-based algorithm. Other feature models that we tried, such as FC-layer or color moments did not give better results.

Clustering:

1. By its definition, DBSCAN is density-based, and as such, direct control over the number of clusters is not possible. To circumvent this, we try several combinations of values for the parameters and find the best fitting model. Even so, the clustering is not as effective since we prioritize a target number of clusters
2. Based on the knowledge we have of the dataset, each label has roughly around 20-60 images, with a few outlier labels that are much larger. Based on this, we consider the range of MinPts as mentioned above, to adjust for the inherent dimensionality of the label's data, while also considering a minimum cluster size to optimize for the given c value. This approach can be replaced with better estimates in other settings depending on knowledge about the dataset's dimensionality.
3. One approach to calculate epsilon involves plotting the KNN distances in a sorted order and finding the knee point, where k≈MinPts. However this works as an optimal value for the original DBSCAN to get maximal density, whereas in this task, we prioritize a specific number of clusters. Our approach is to take the mean and minimum KNN distances where k≈MinPts, and try 100 equally spaced values in this range, and then searching for the optimal combination based on cluster count and noise points count.
4. Even exhaustively trying combinations of parameters, we are unable to obtain exactly c clusters for some labels, especially the smaller ones with lower inherent dimensionality. This is most likely because such labels, when clustered, form a few large clusters and the rest are classified as noise points. For larger labels, the approach is somewhat effective but slower as the object-object distance matrix grows quadratically in size. In fact, some labels only form a single trivial cluster.
5. We also observe a high level of noise for almost all the labels, which can be attributed to the same reasons as before: cluster count control and low inherent dimensionality.

Visualization:

1. The classical MDS approach was used to visualize the clusters in a 2-D space. However, metric MDS may be a better approach. This is because, while classical MDS preserves rank-ordering of distances, metric MDS preserves actual distances. This is desirable in a clustering setting. However, metric MDS is computationally more expensive than classical MDS. Moreover, in this setting, the space that we

project to, usually has significantly fewer dimensions than the inherent dimensionality of the label's data. Since accuracy is lost anyway, we preferred a computationally cheaper approach. In retrospect, metric MDS would have shown slightly better tightness and separation of clusters.

2. Our interpretation of "groups of image thumbnails" was to display one image per cluster at the cluster centroid in the MDS space. Note that these images usually overlap a lot. Thus, interactive plots are generated to better display overlays of images at runtime.

Prediction and metrics:

1. Since the feature space does not differ across the labels, the approach we used for prediction involves taking all the core points of all the clusters, across all the labels, together. Then we compare each target image with every core point, and poll the closest few core points to get the majority label. This approach is computationally expensive, in fact it is several times slower than the clustering itself. The reasoning for this was the fact not only are we using the DBSCAN algorithm in a way not suited for the algorithm, but also the fact that noise points were large in number, so our intuition was to not rely on the clusters as a whole, but the core points that make up the clusters. Statistical evaluation of the results are satisfactory. There may be other approaches that utilize cluster boundaries and are more performant.

2. The formulae for the precision, recall and F1-score metrics are given as:

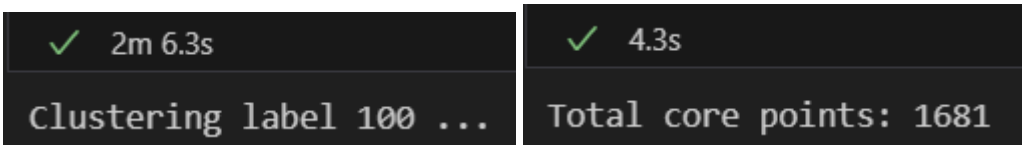
$$Precision = \frac{TP}{TP+FP}, \quad Recall = \frac{TP}{TP+FN}, \quad F1\,score = \frac{2*Precision*Recall}{Precision+Recall}$$

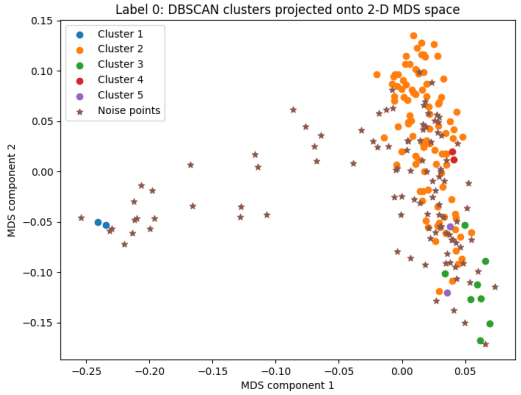
where, TP=True Positives, FP=False Positives, FN=False Negatives. These metrics are calculated per label.

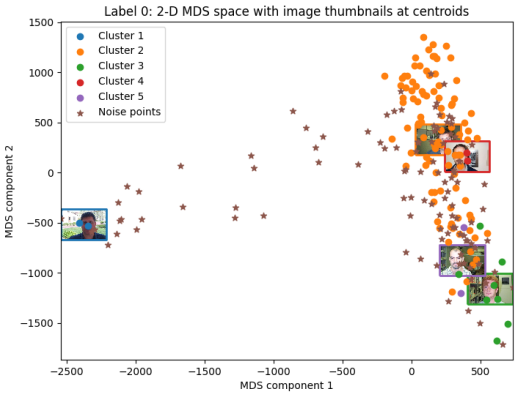
3. Overall accuracy is calculated as the fraction of images where the predicted label matches the actual label, taken across all the labels together.

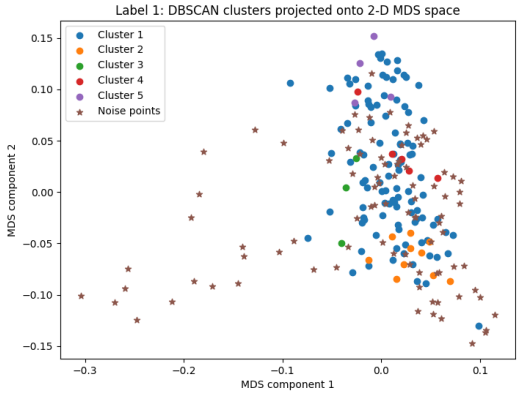
Output

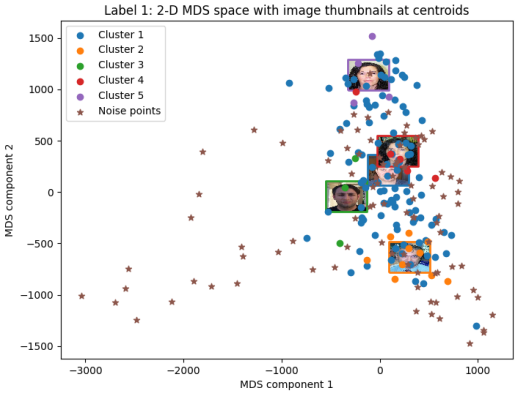
c=5

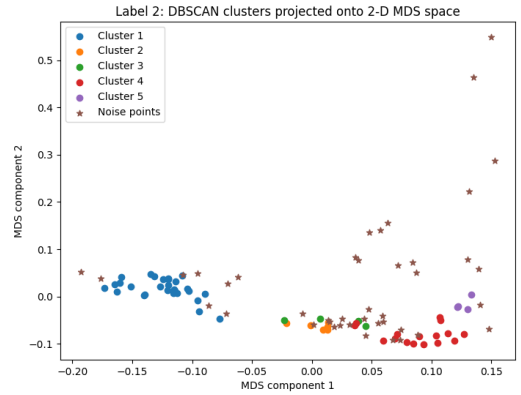


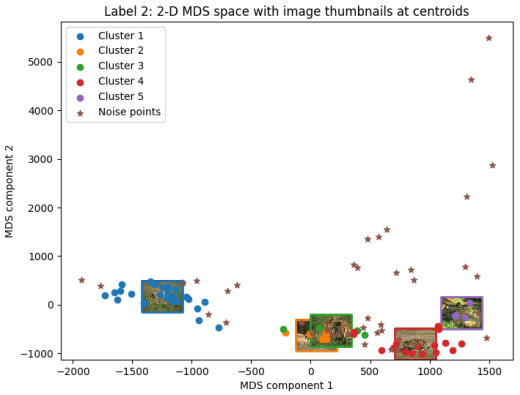

Label 0: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

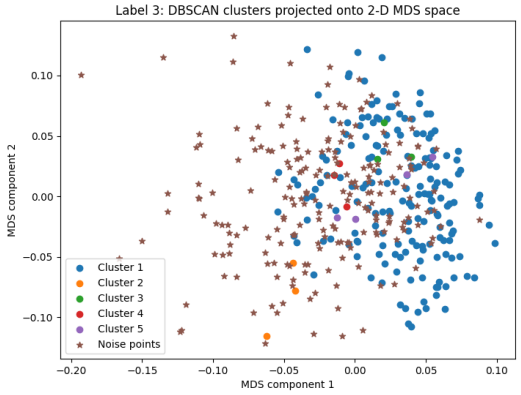

Label 0: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

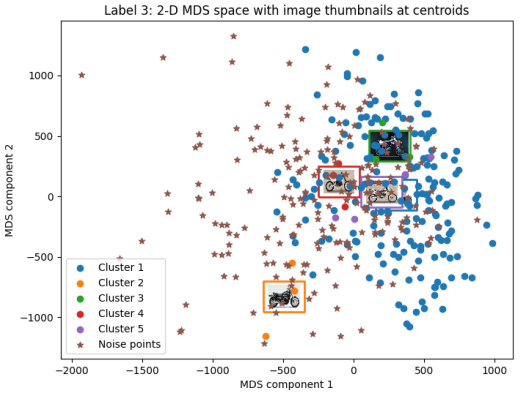

Label 1: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

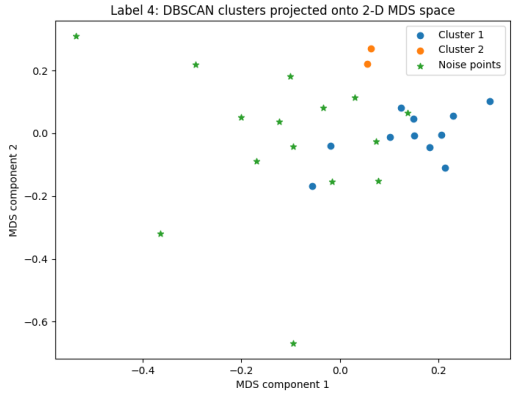

Label 1: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

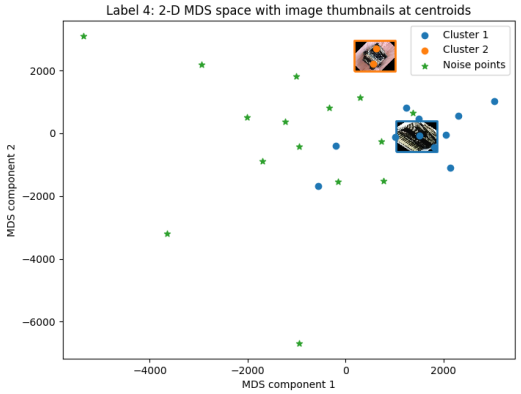

Label 2: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

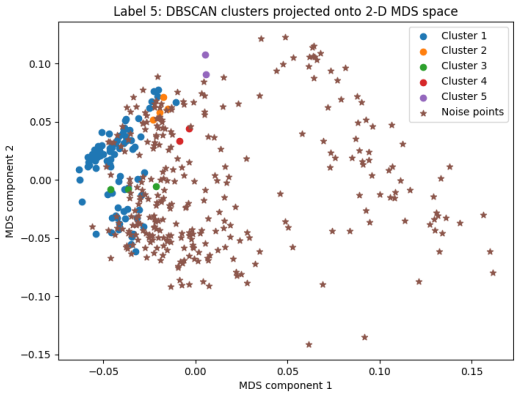

Label 2: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

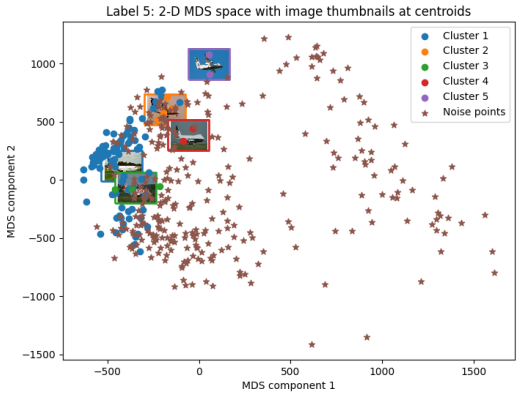

Label 3: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

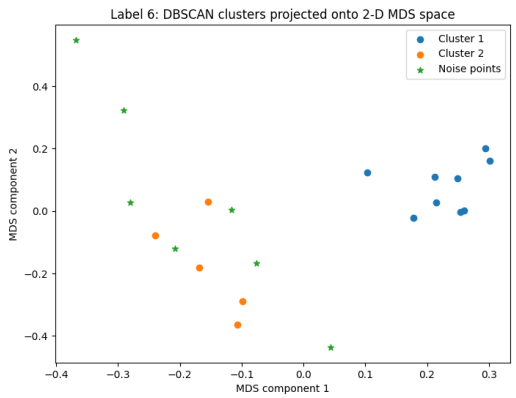

Label 3: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

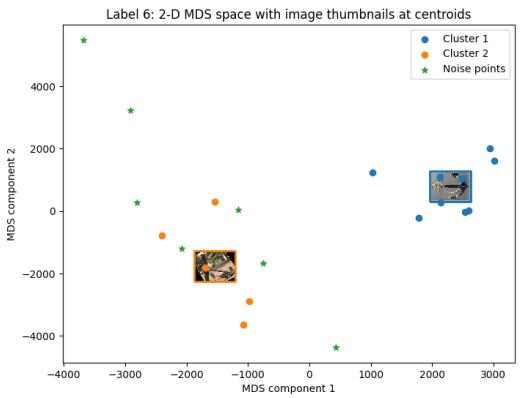

Label 4: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

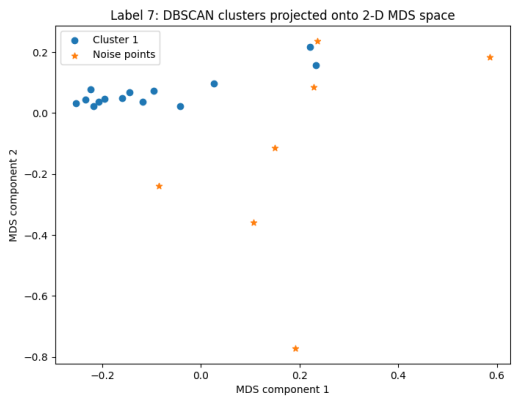

Label 4: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

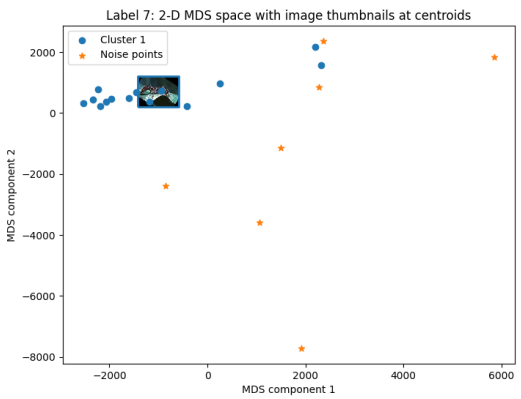

Label 5: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

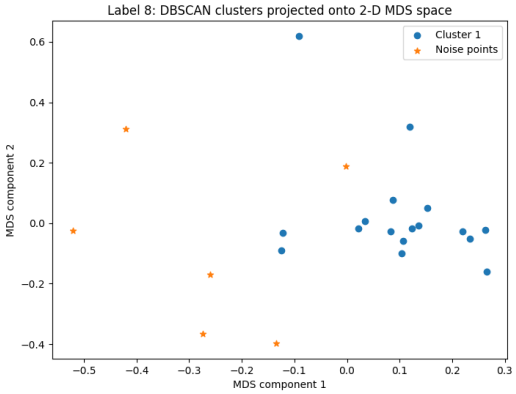

Label 5: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

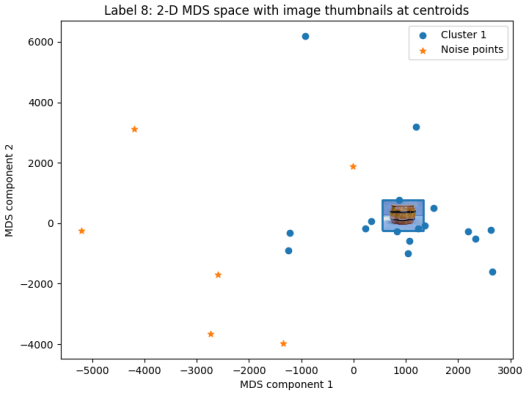

Label 6: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

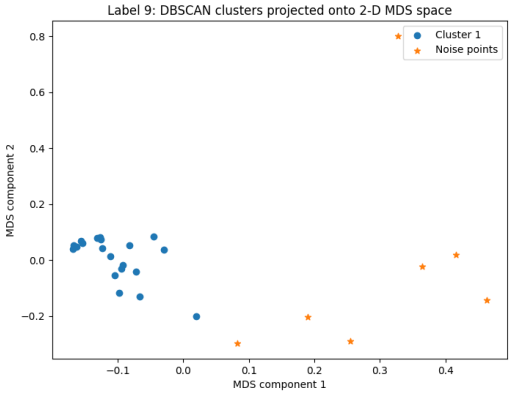

Label 6: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

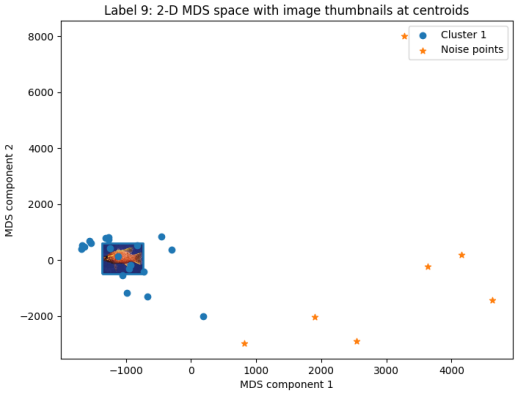

Label 7: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

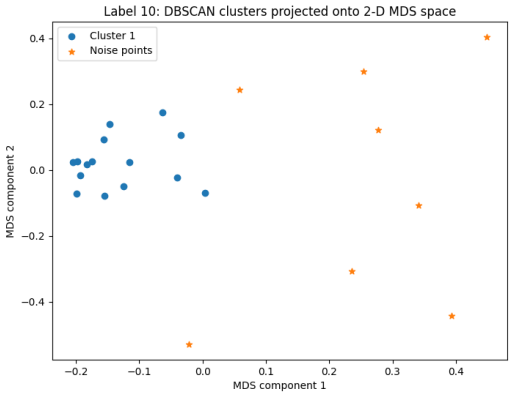

Label 7: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

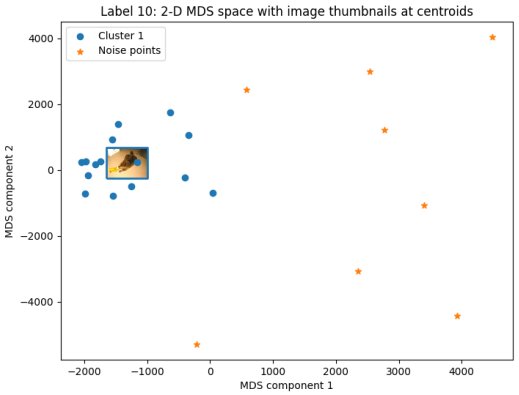

Label 8: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

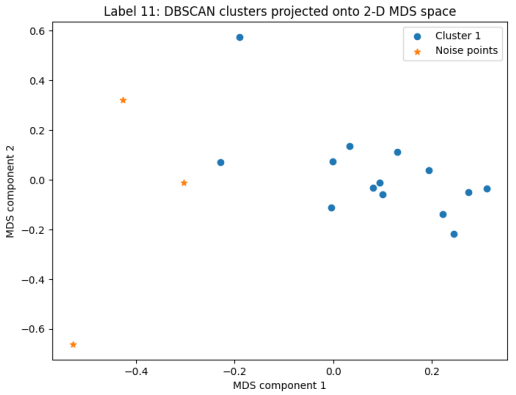

Label 8: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

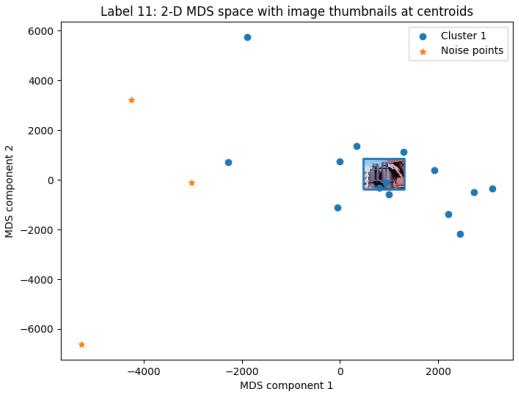

Label 9: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

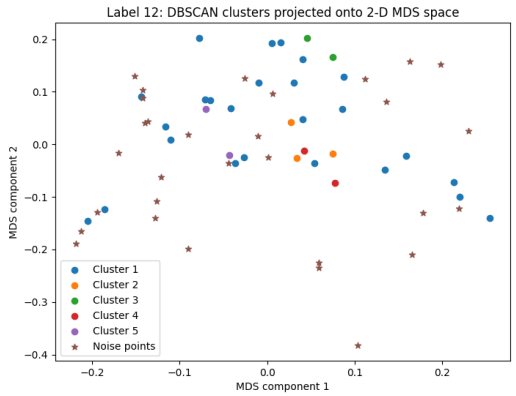

Label 9: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

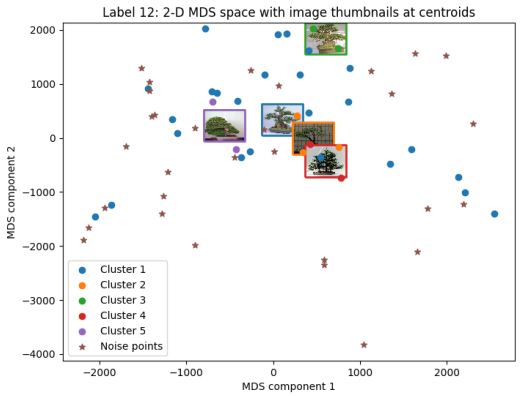

Label 10: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

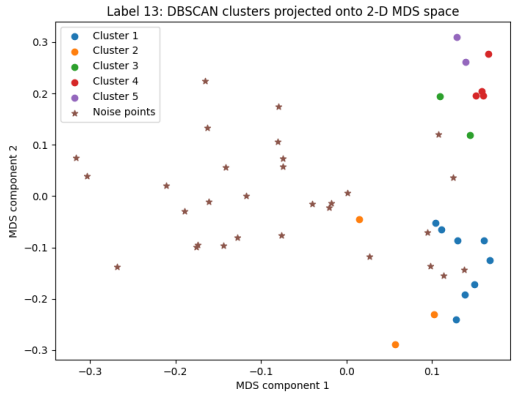

Label 10: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

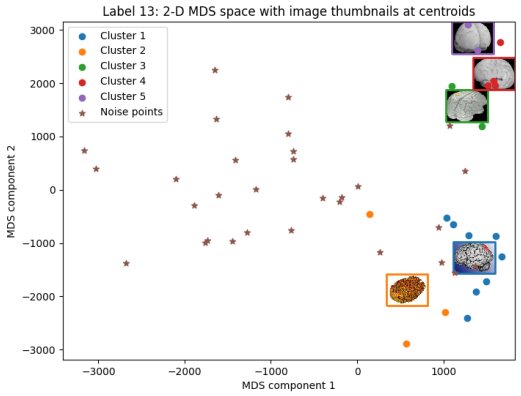

Label 11: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

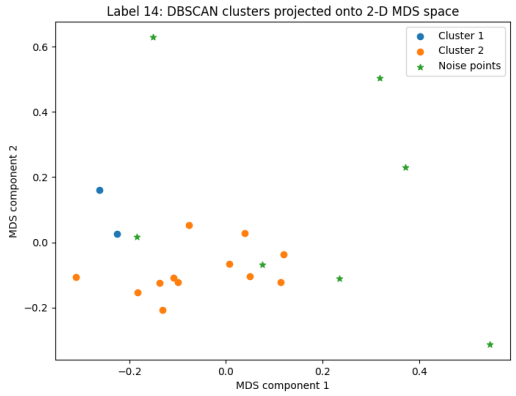

Label 11: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

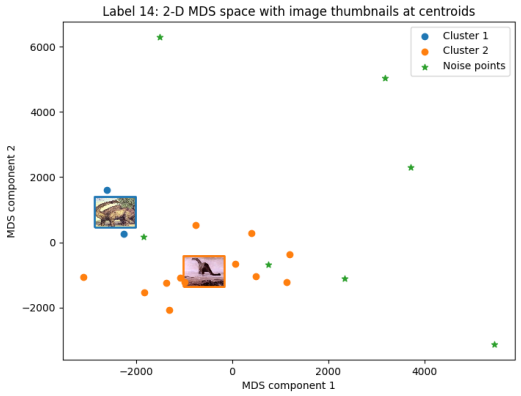

Label 12: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points

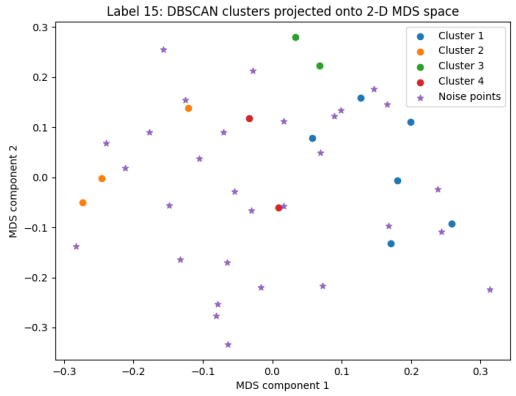

Label 12: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points

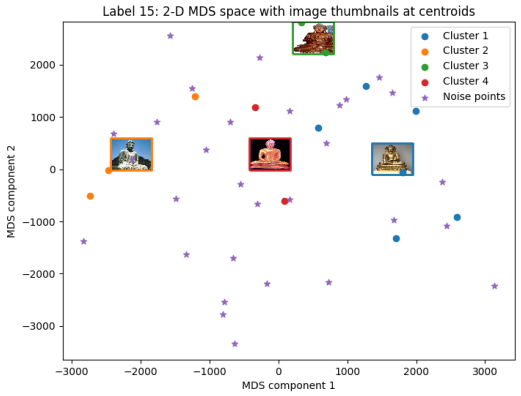

Label 13: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points

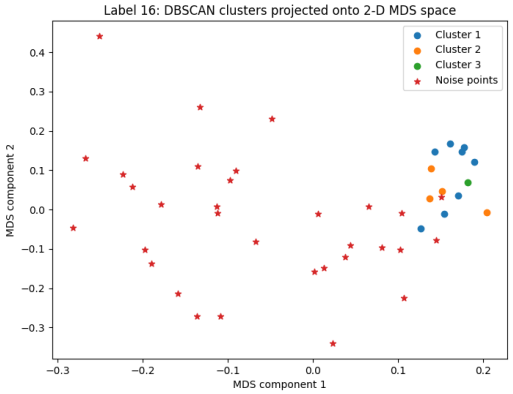

Label 13: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points

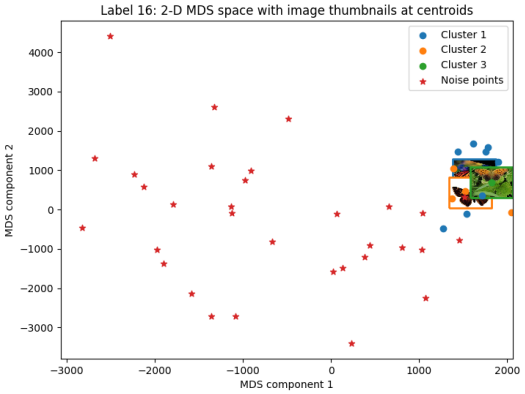

Label 14: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Noise points

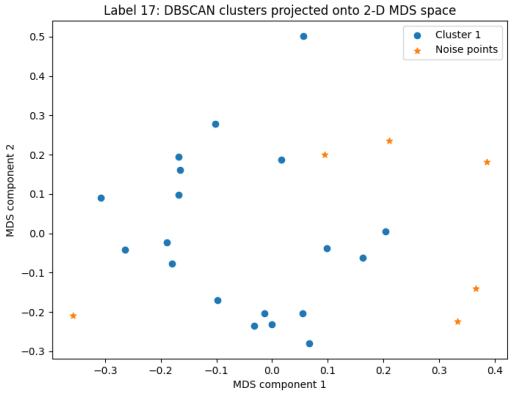

Label 14: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Noise points

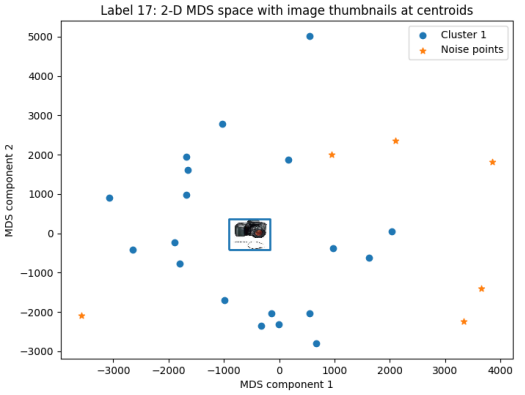

Label 15: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points

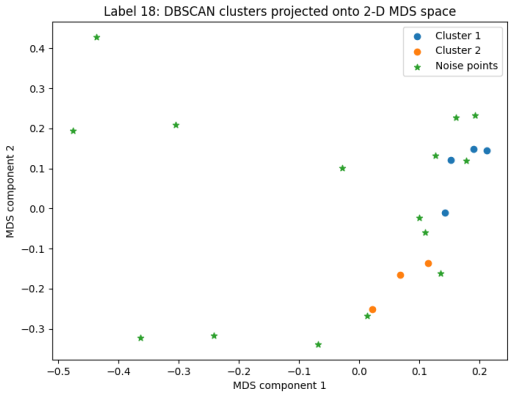

Label 15: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points

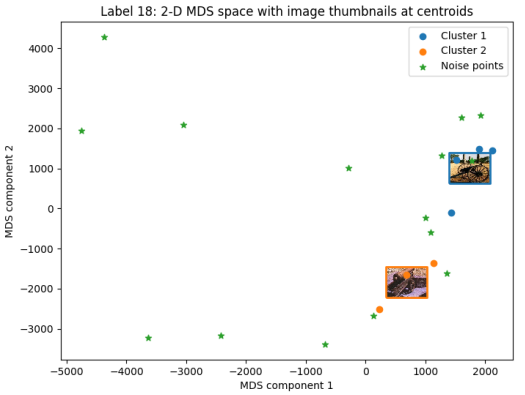

Label 16: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

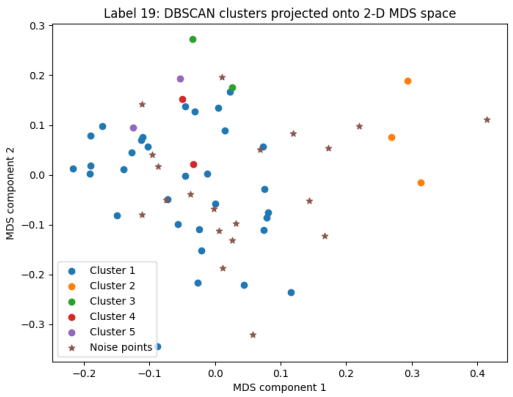

Label 16: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

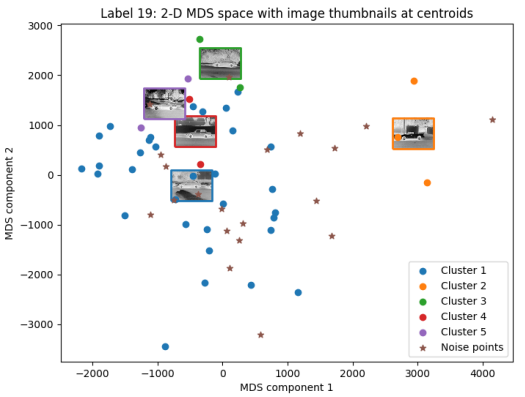

Label 17: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

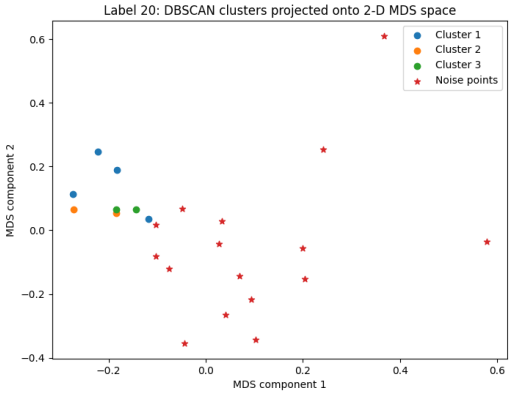

Label 17: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

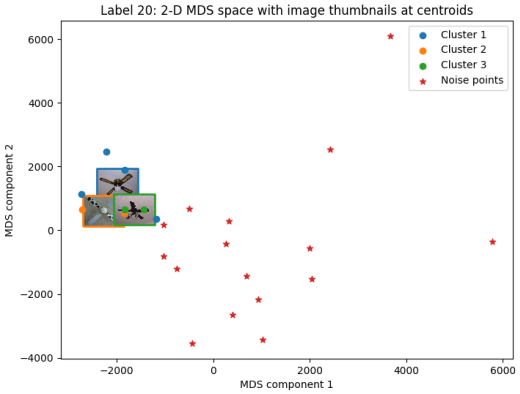

Label 18: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

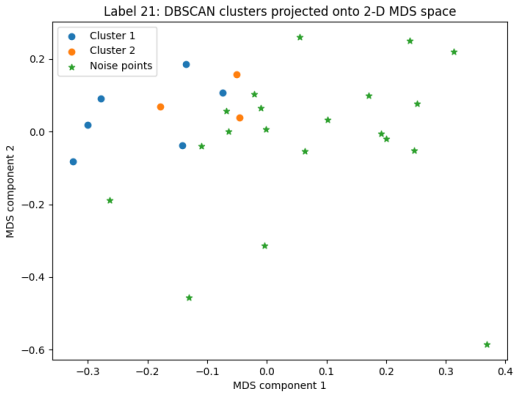

Label 18: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

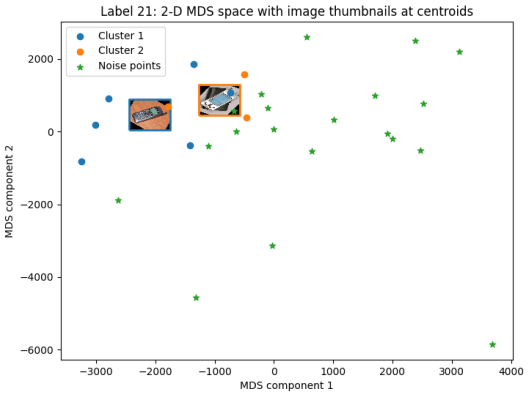

Label 19: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

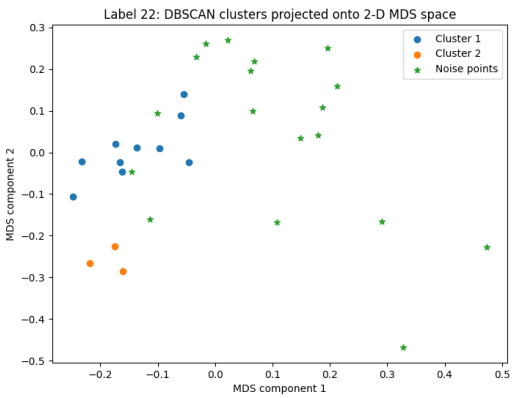

Label 19: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

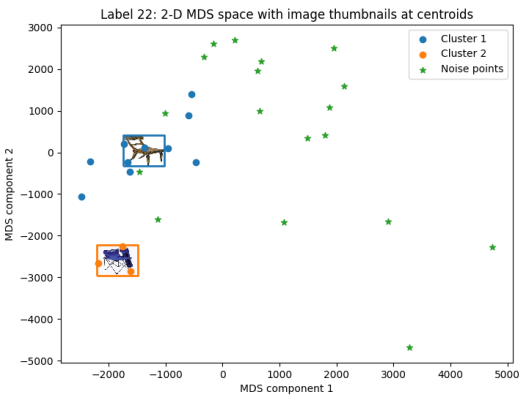

Label 20: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

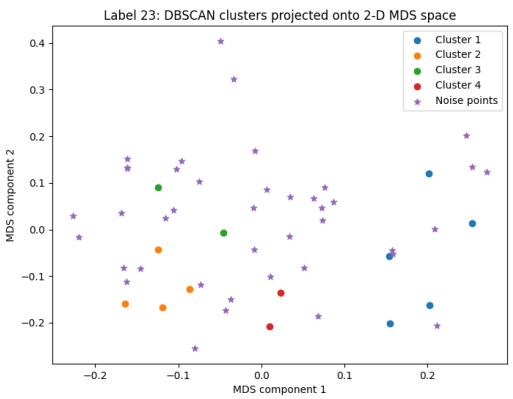

Label 20: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

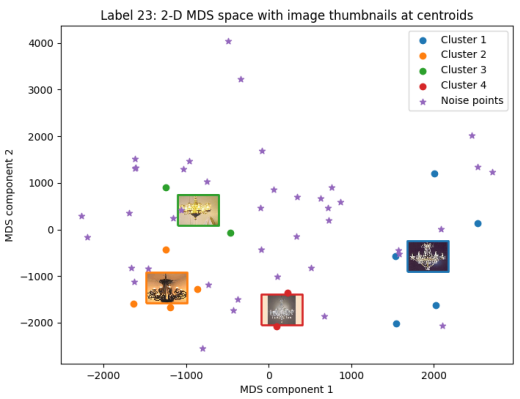

Label 21: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

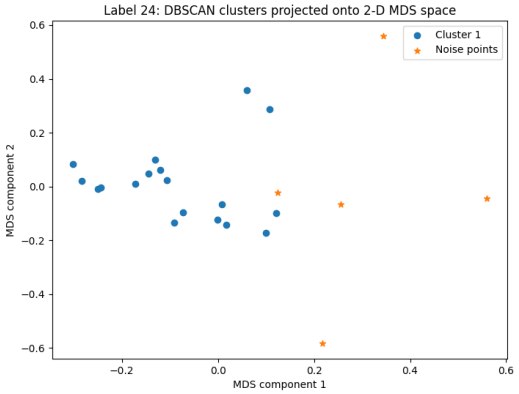

Label 21: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

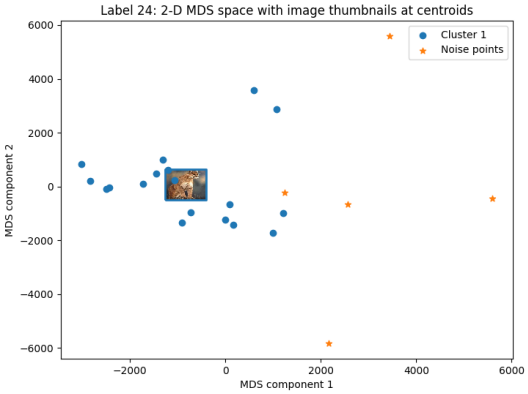

Label 22: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

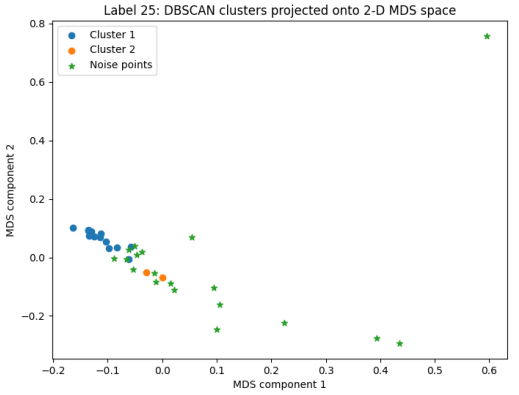

Label 22: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

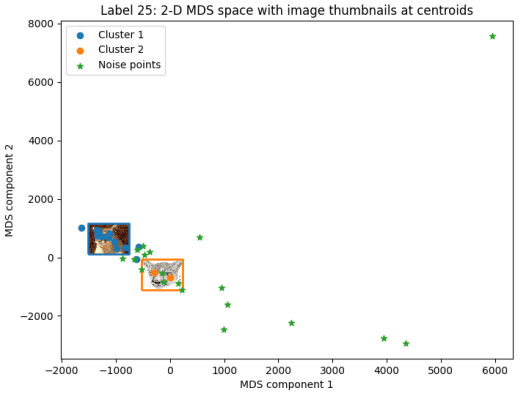

Label 23: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

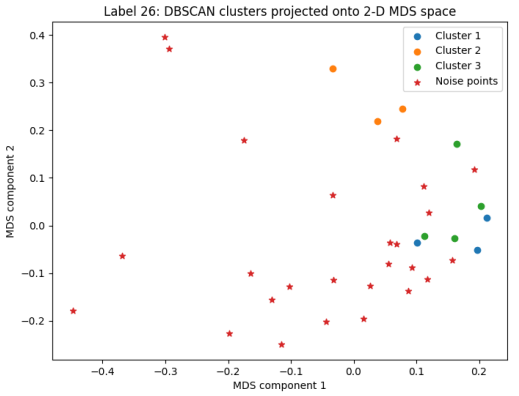

Label 23: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

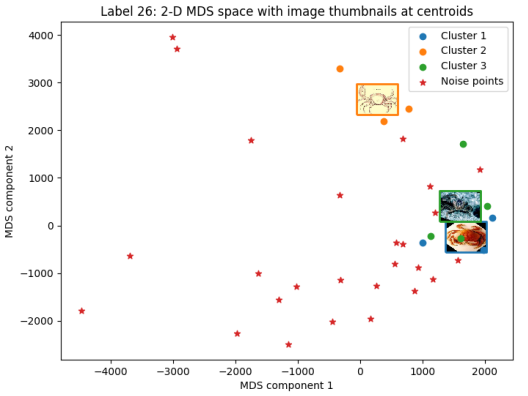

Label 24: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

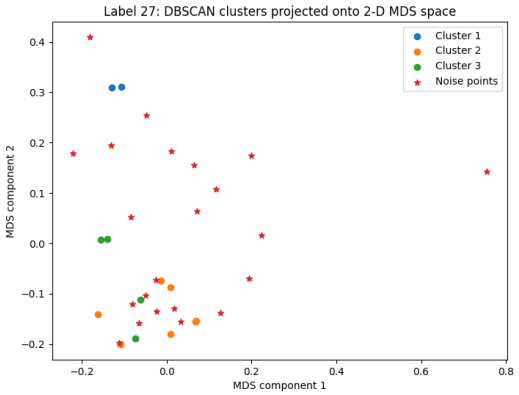

Label 24: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

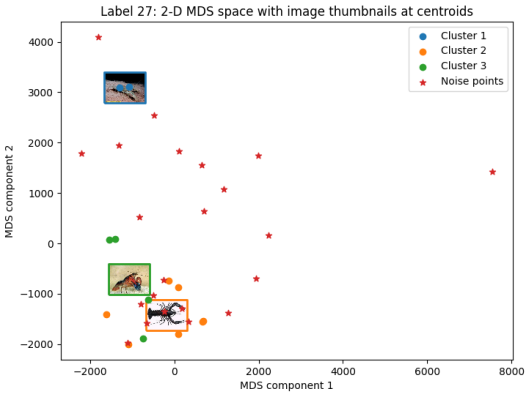

Label 25: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

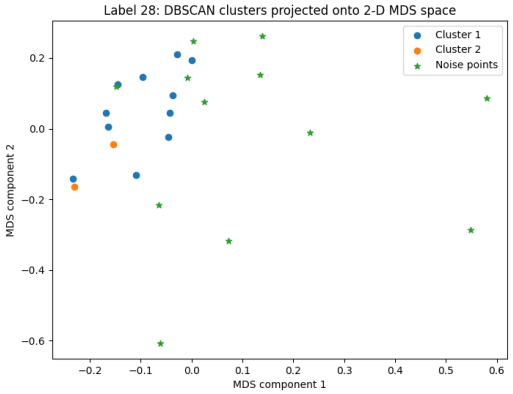

Label 25: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

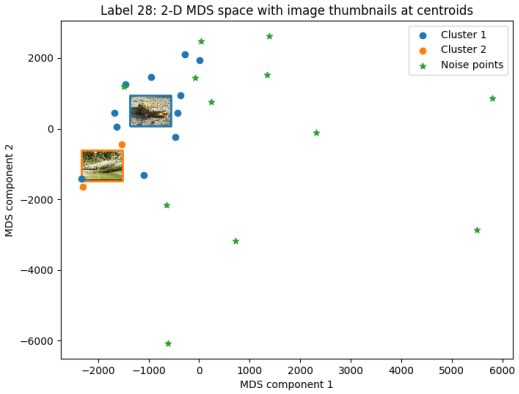

Label 26: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

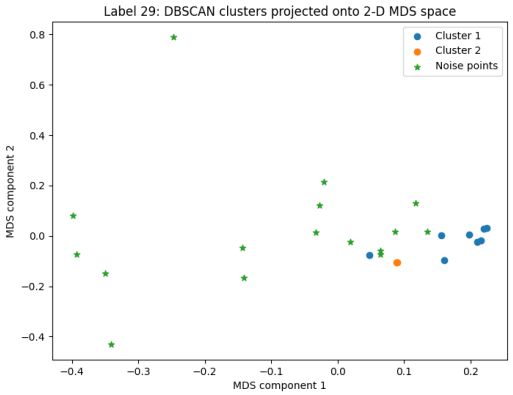

Label 26: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

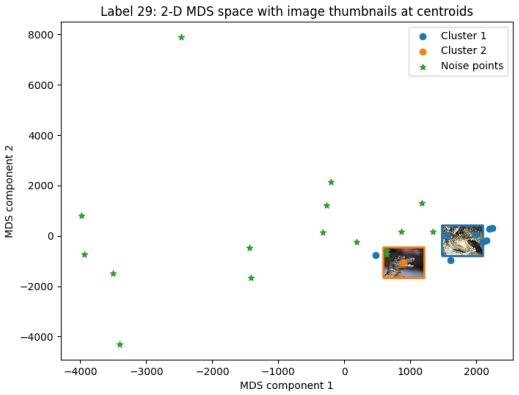

Label 27: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

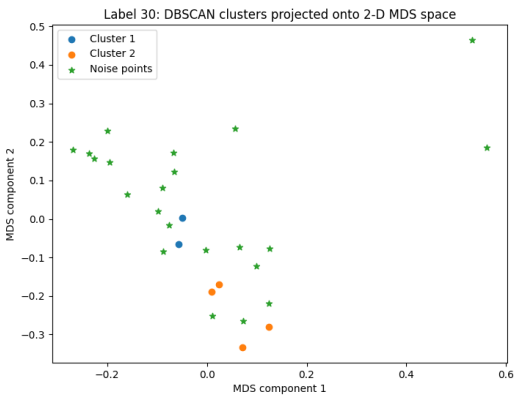

Label 27: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

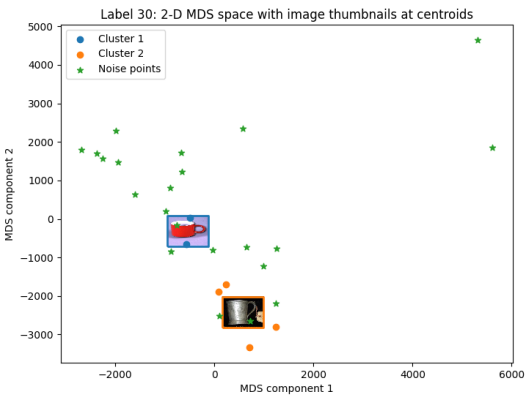

Label 28: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

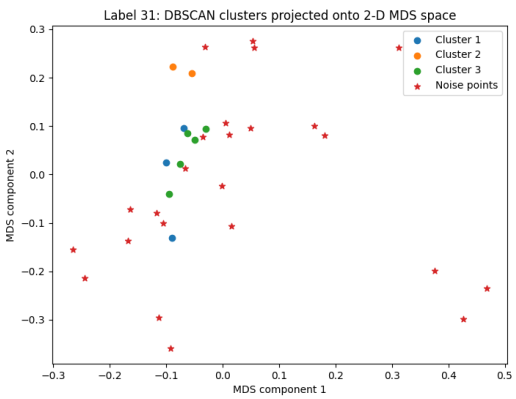

Label 28: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

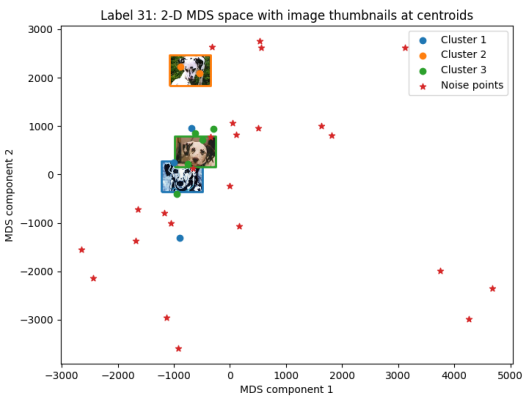

Label 29: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

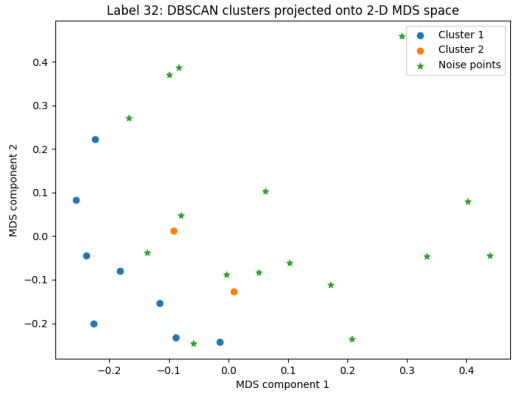

Label 29: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

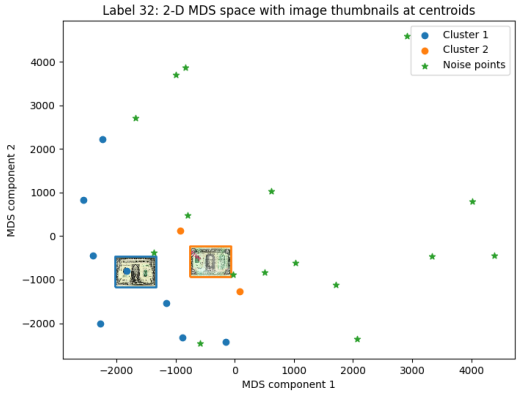

Label 30: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

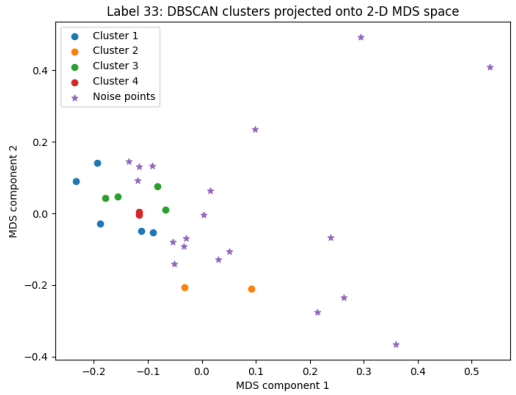

Label 30: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

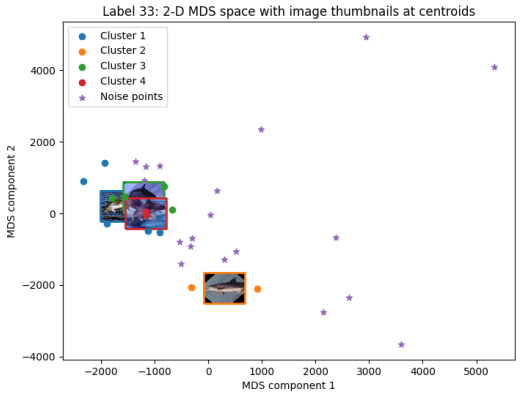

Label 31: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

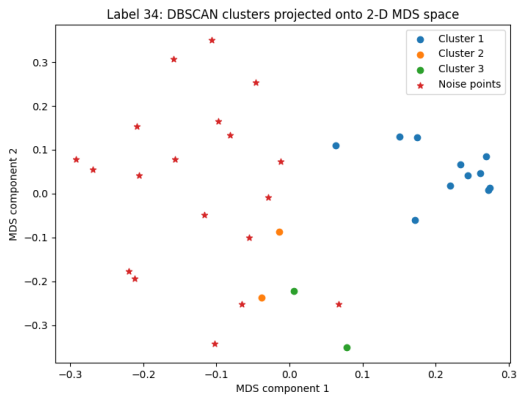

Label 31: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

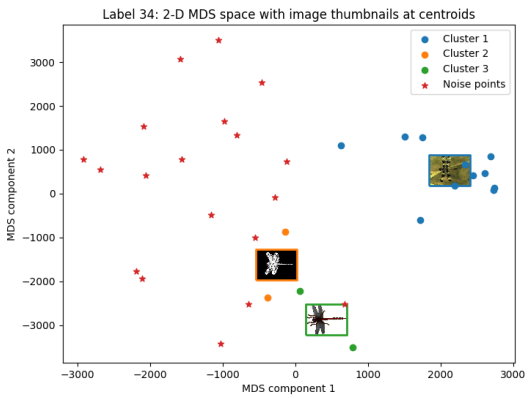

Label 32: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

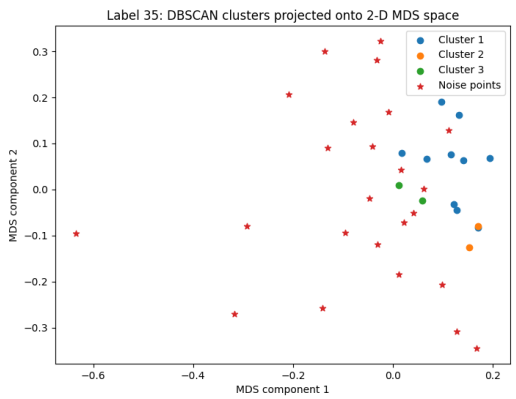

Label 32: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

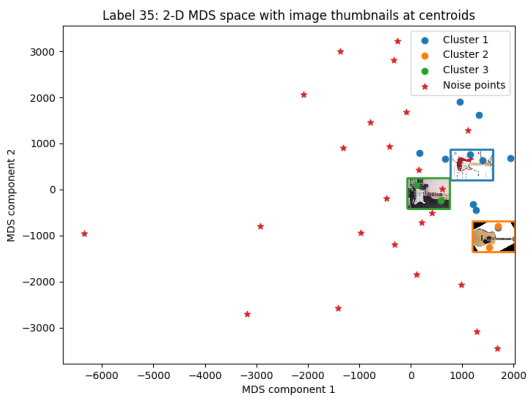

Label 33: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

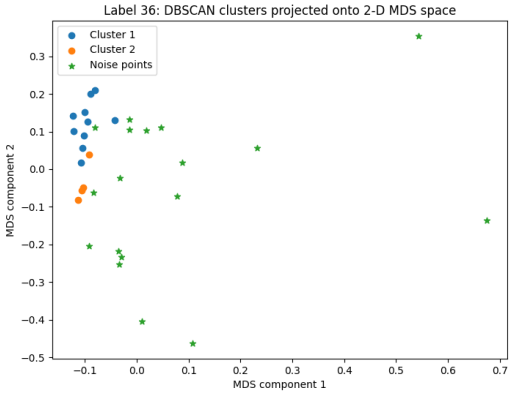

Label 33: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

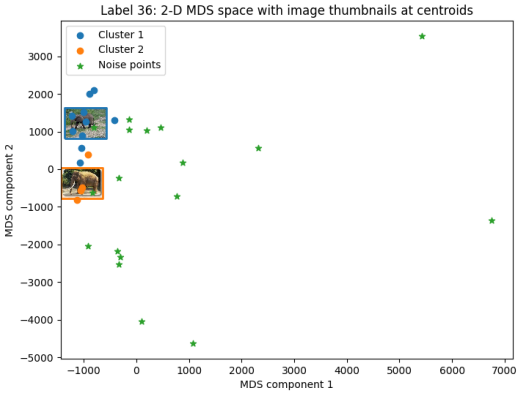

Label 34: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

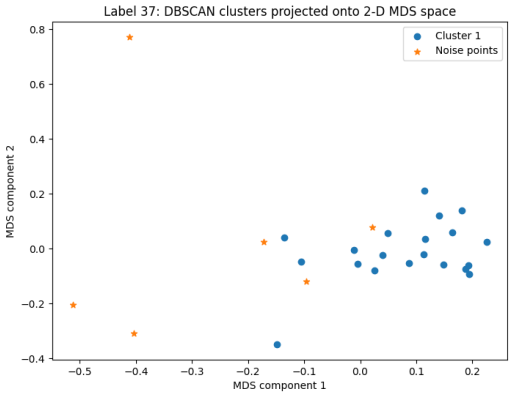

Label 34: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

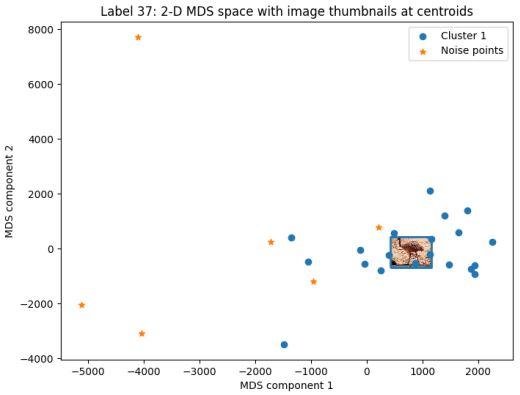

Label 35: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

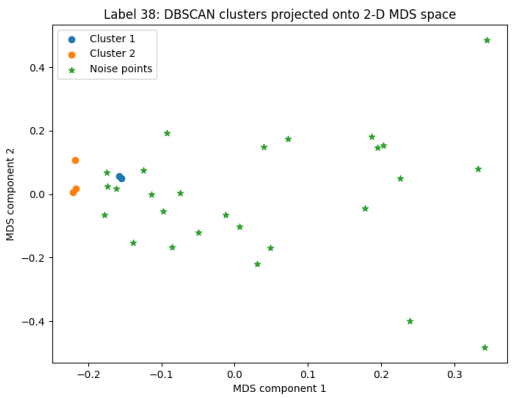

Label 35: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

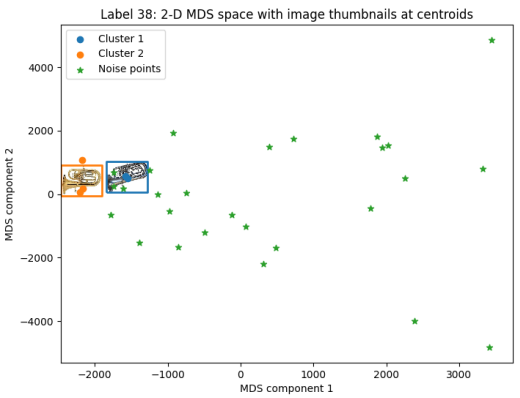

Label 36: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

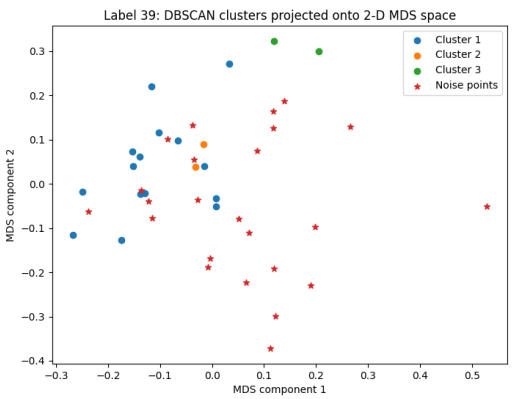

Label 36: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

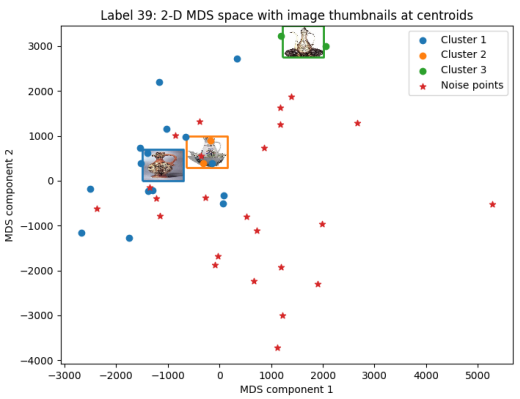

Label 37: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

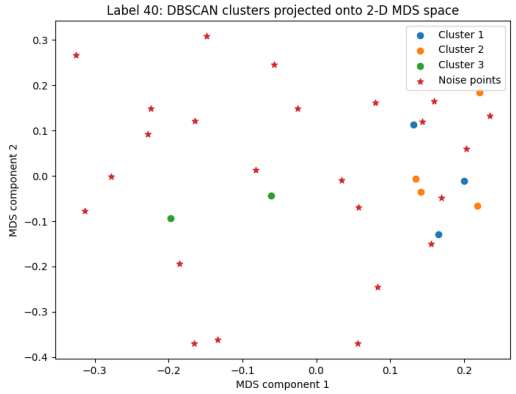

Label 37: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

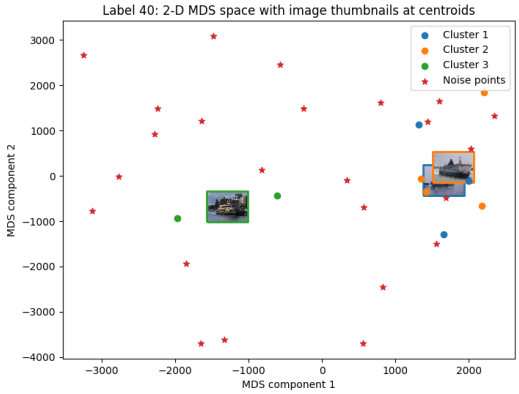

Label 38: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

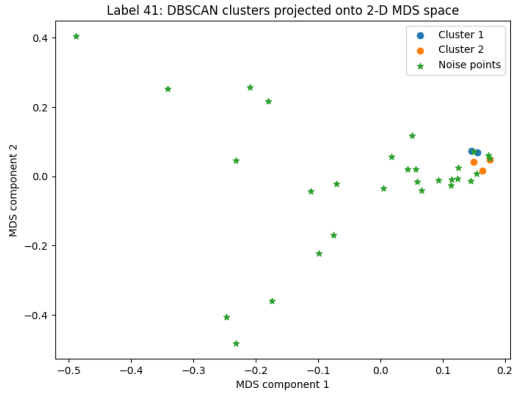

Label 38: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

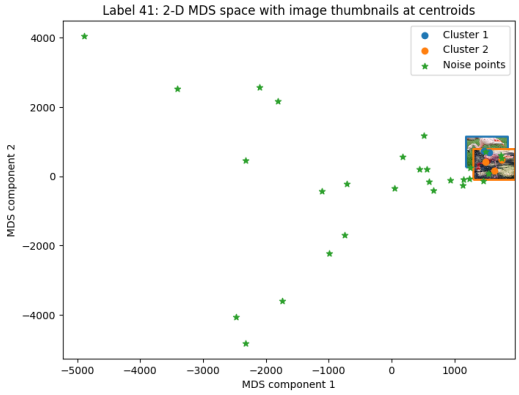

Label 39: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

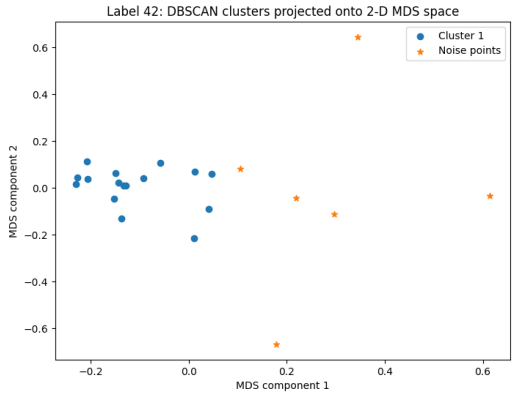

Label 39: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

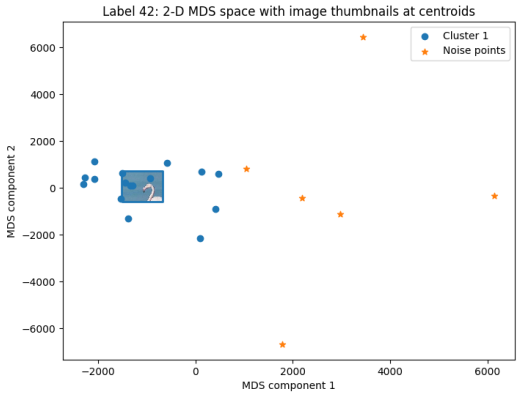

Label 40: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

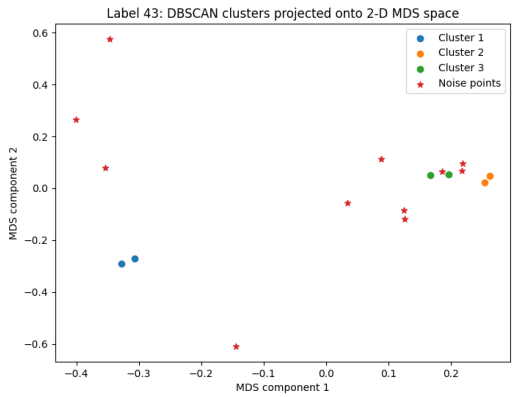

Label 40: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

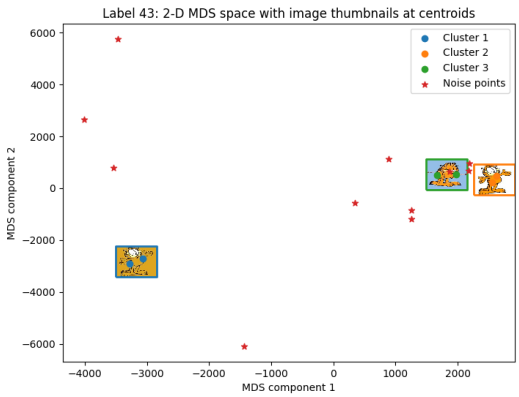

Label 41: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

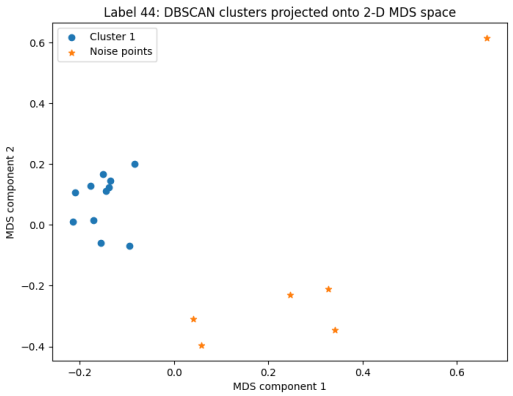

Label 41: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

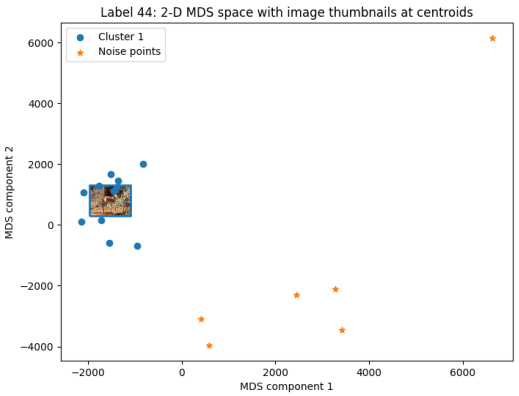

Label 42: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

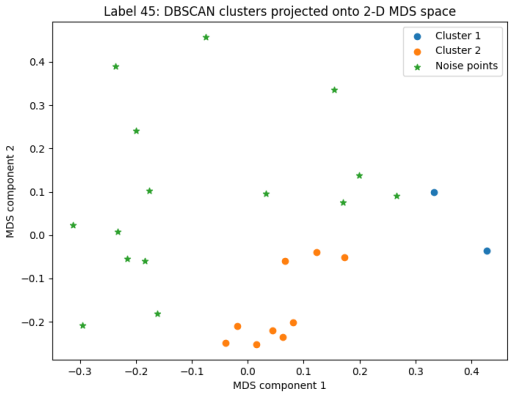

Label 42: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

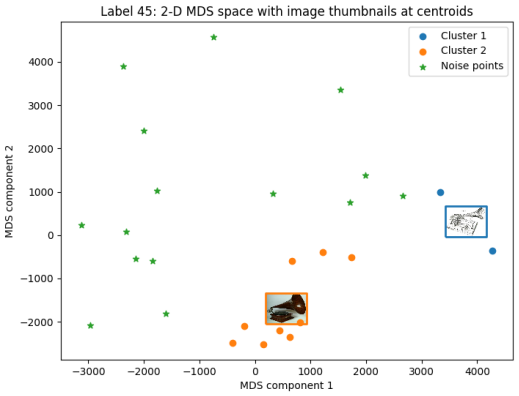

Label 43: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

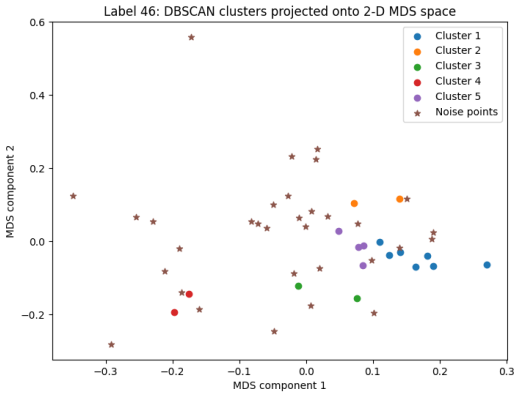

Label 43: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

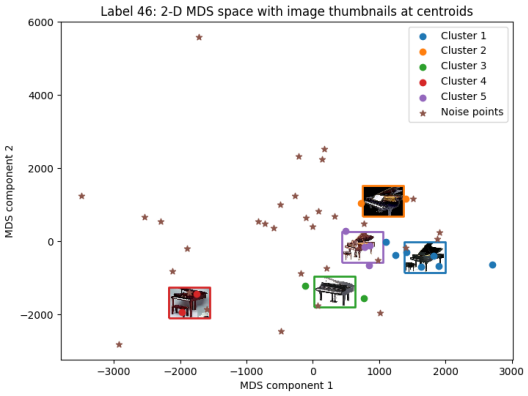

Label 44: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

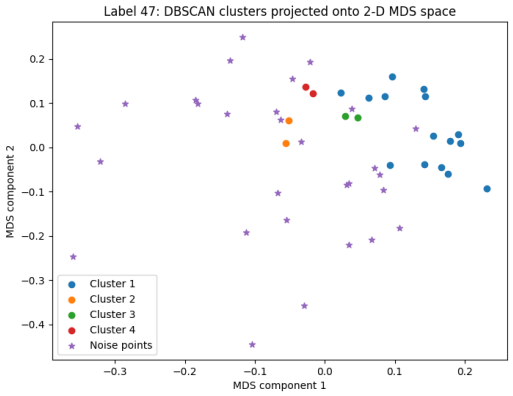

Label 44: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

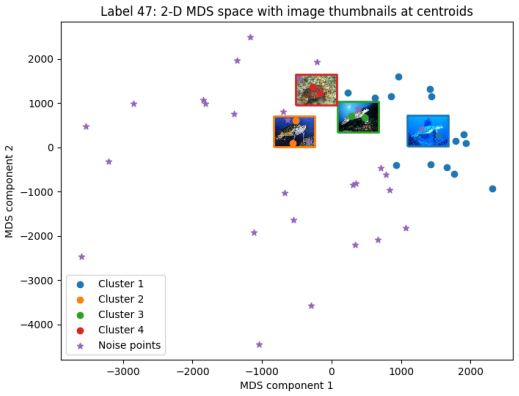

Label 45: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

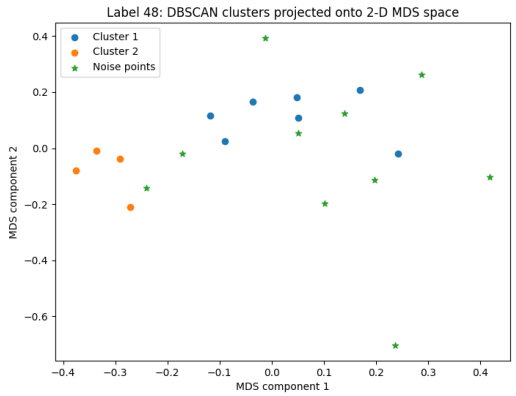

Label 45: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

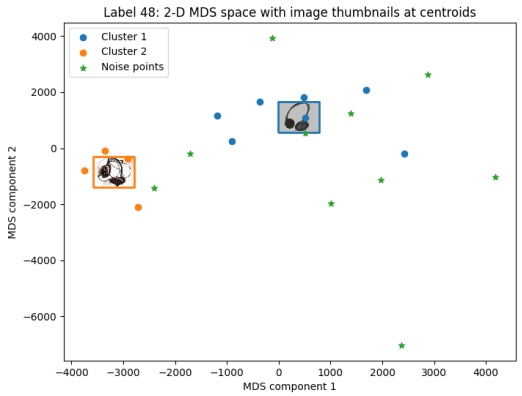

Label 46: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

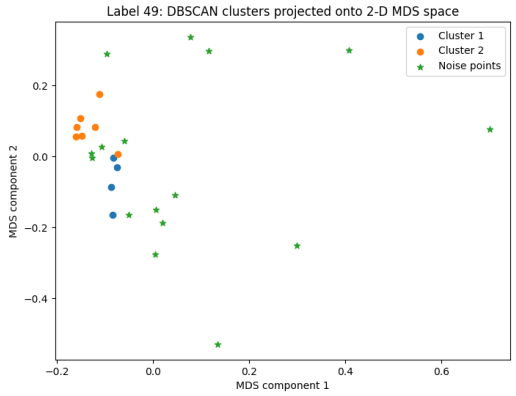

Label 46: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

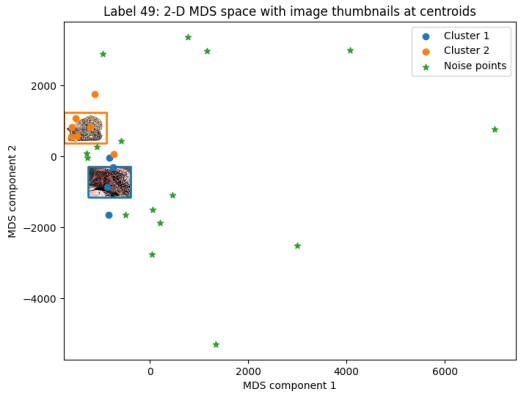

Label 47: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

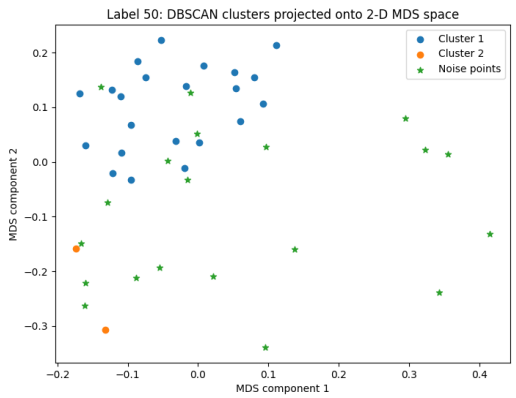

Label 47: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

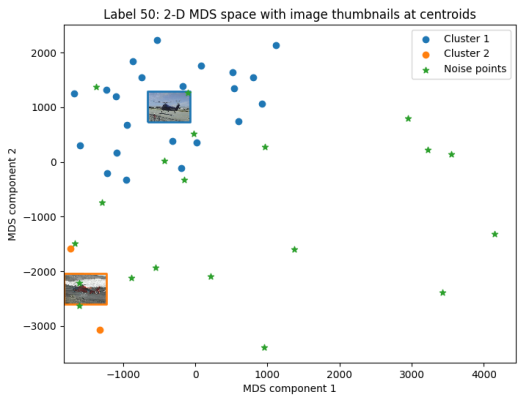

Label 48: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

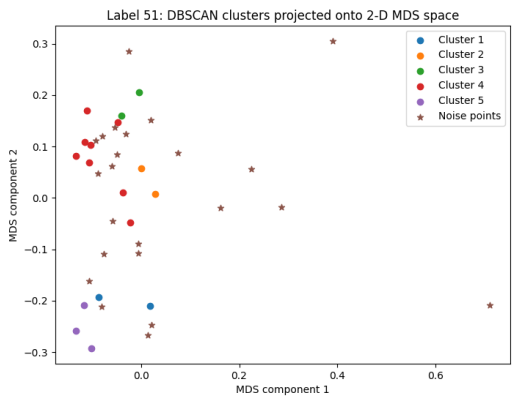

Label 48: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

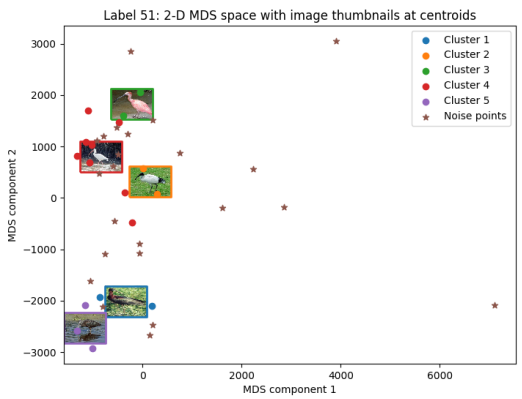

Label 49: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

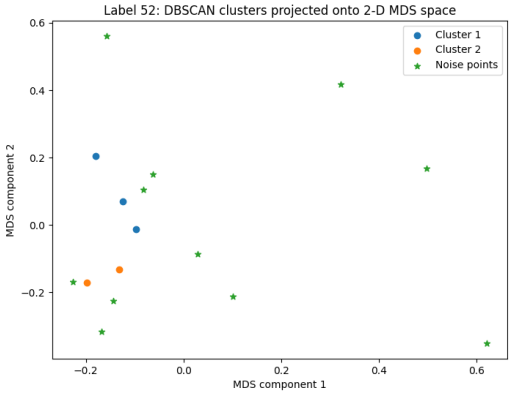

Label 49: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

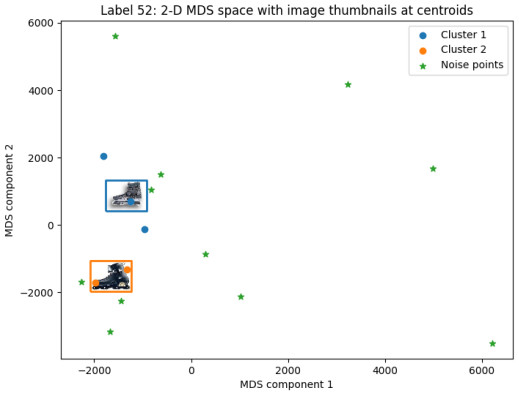

Label 50: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

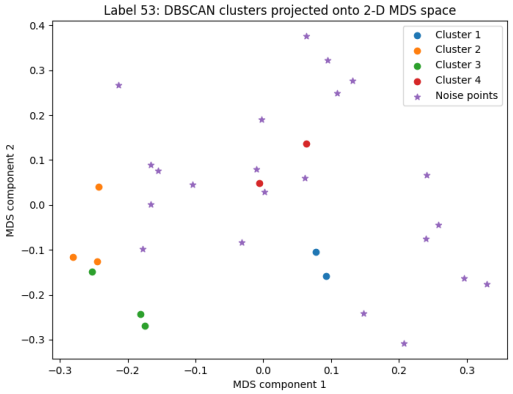

Label 50: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

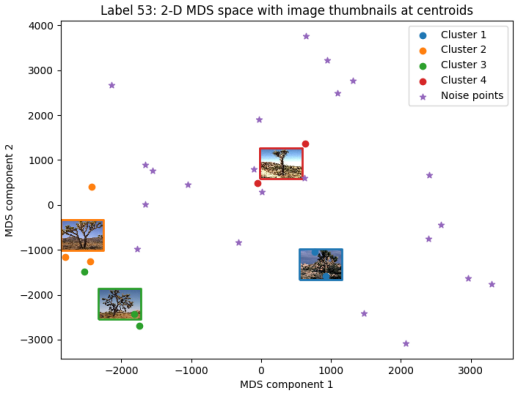

Label 51: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

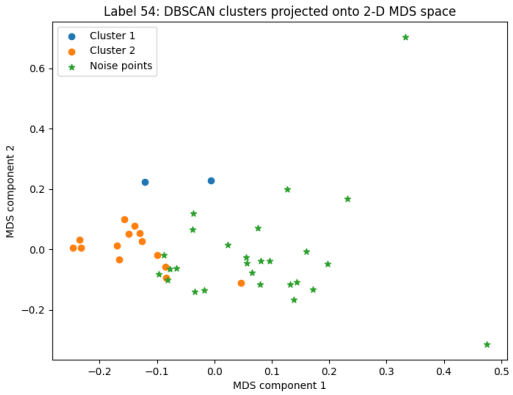

Label 51: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

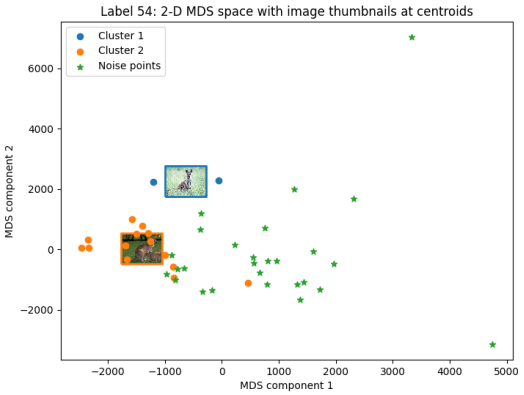

Label 52: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

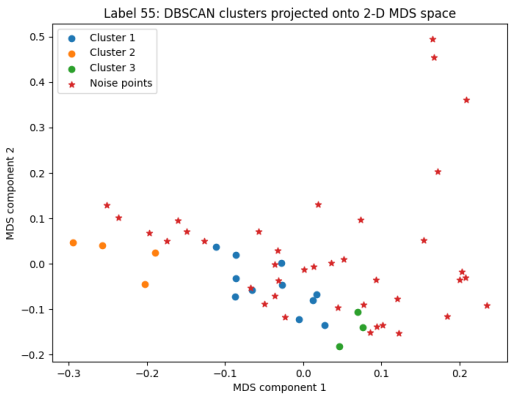

Label 52: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

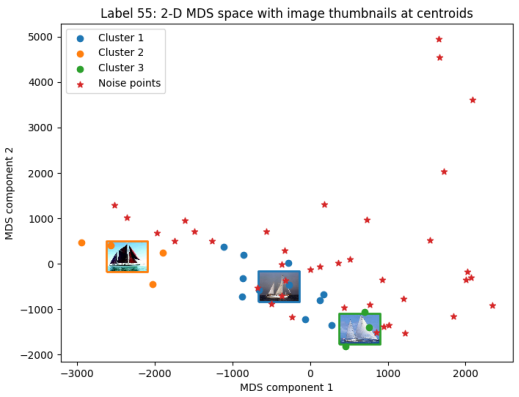

Label 53: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

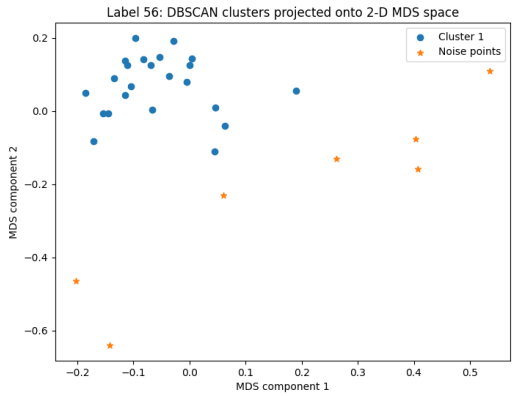

Label 53: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

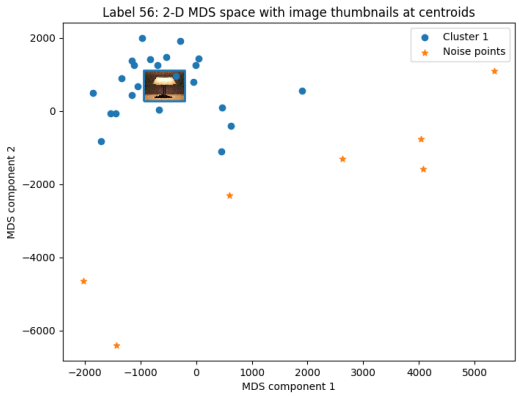

Label 54: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

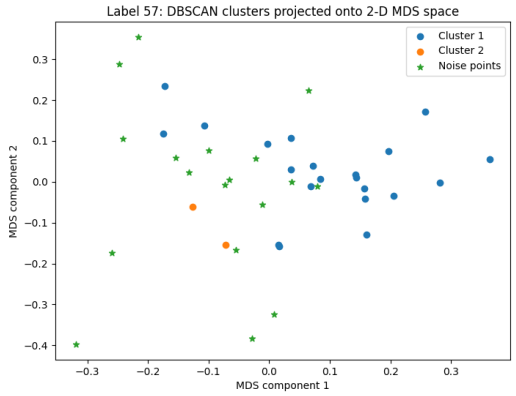

Label 54: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

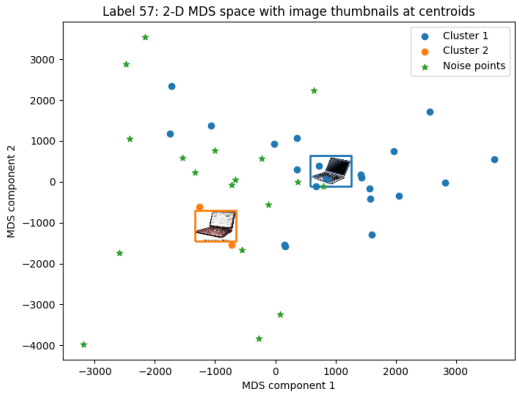

Label 55: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

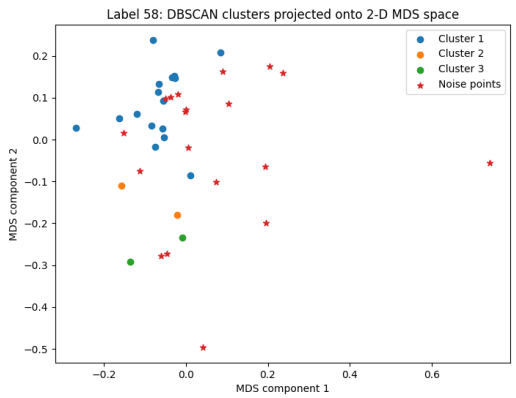

Label 55: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

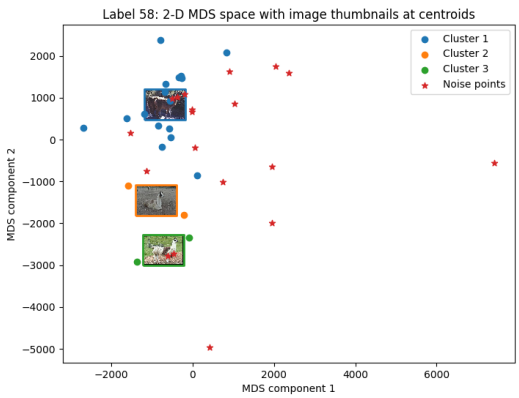

Label 56: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

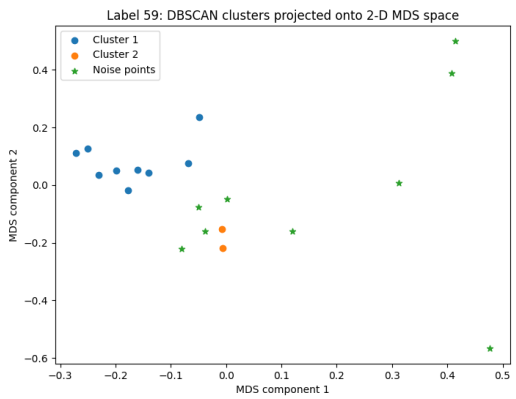

Label 56: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

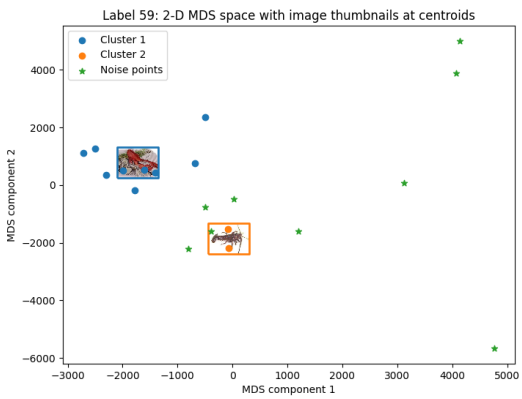

Label 57: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

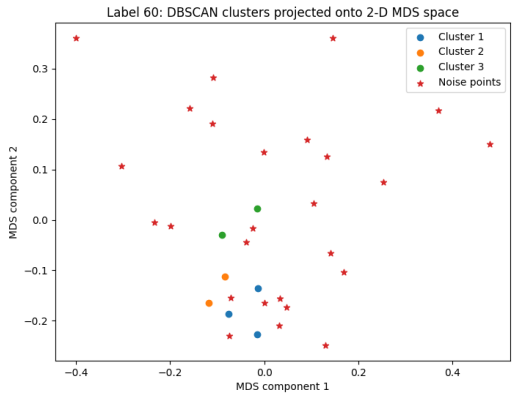

Label 57: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

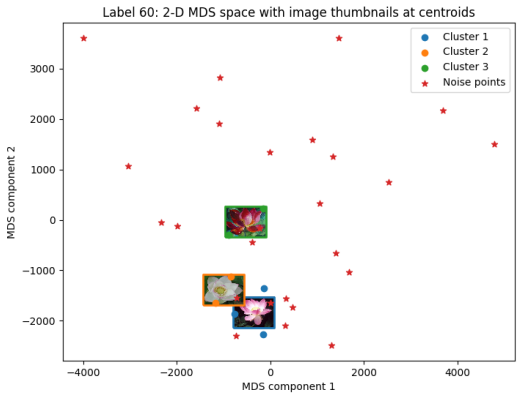

Label 58: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

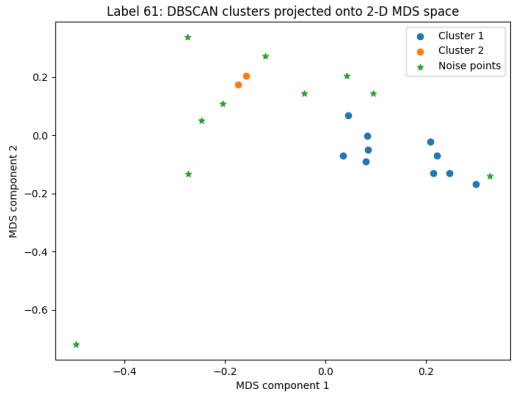

Label 58: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

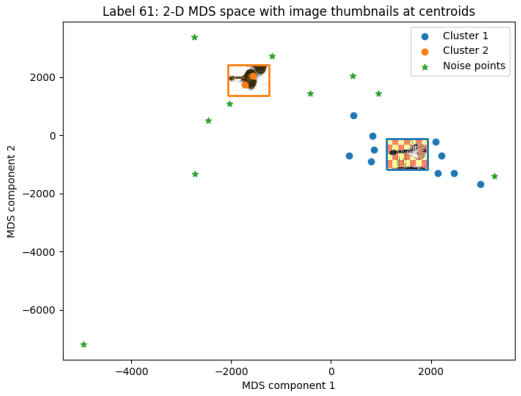

Label 59: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

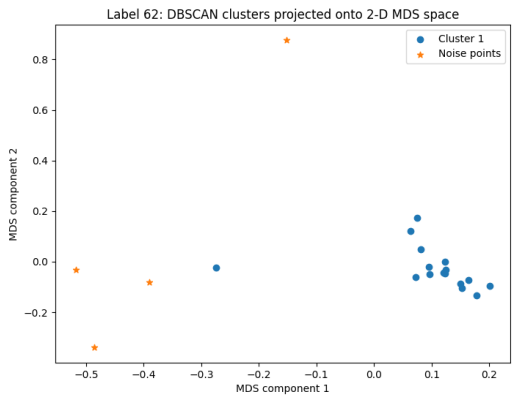

Label 59: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

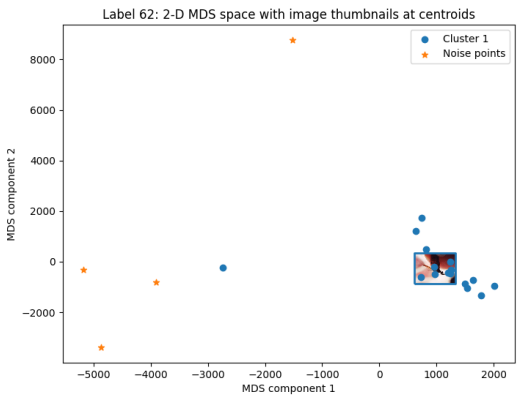

Label 60: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

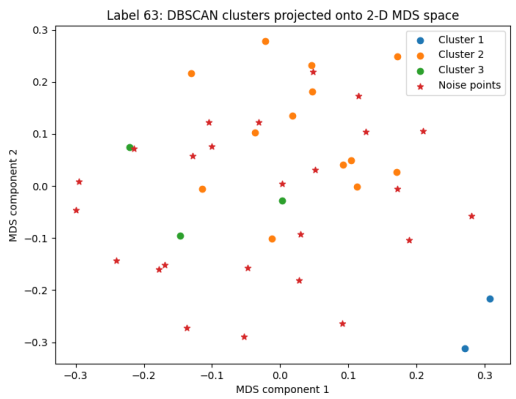

Label 60: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

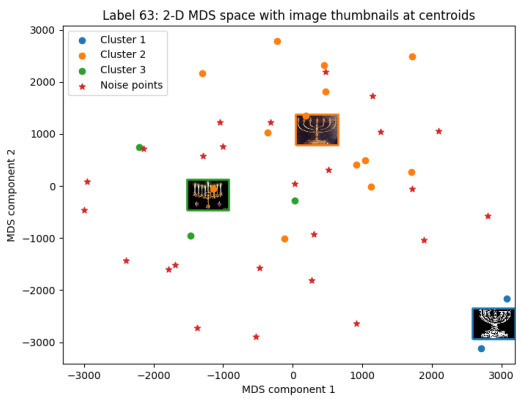

Label 61: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

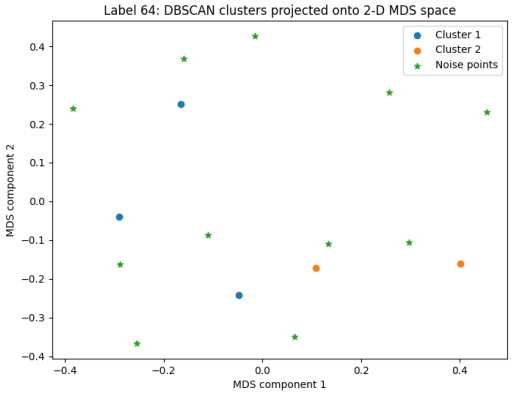

Label 61: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

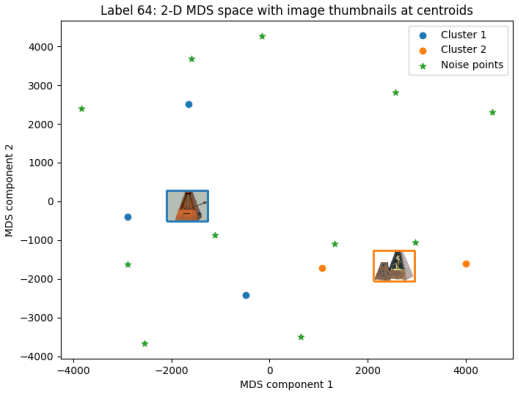

Label 62: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

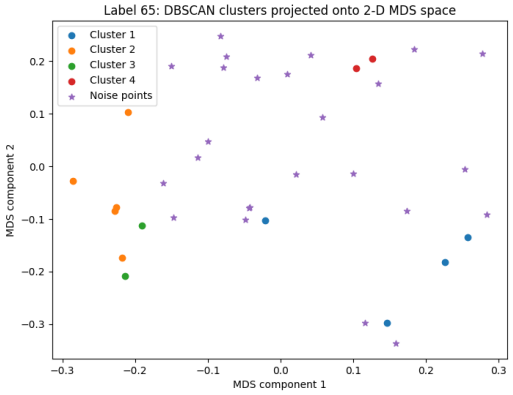

Label 62: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

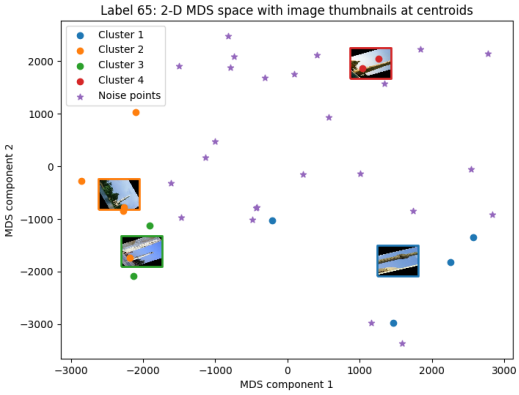

Label 63: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

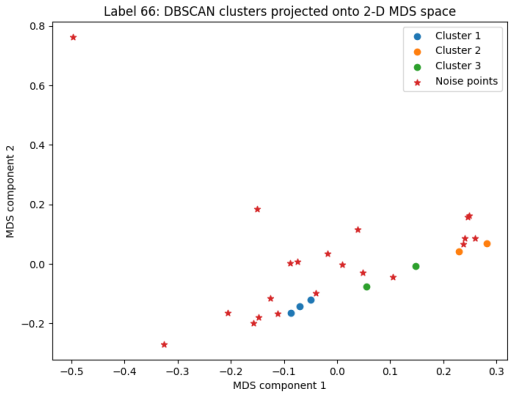

Label 63: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

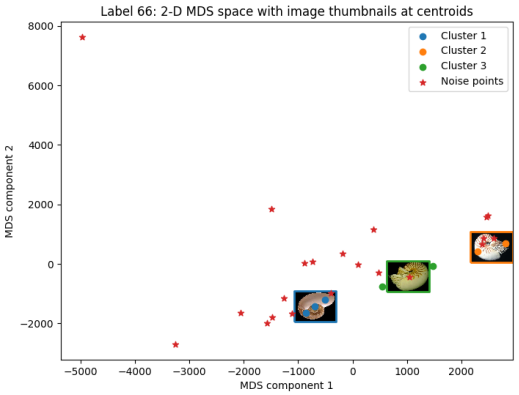

Label 64: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

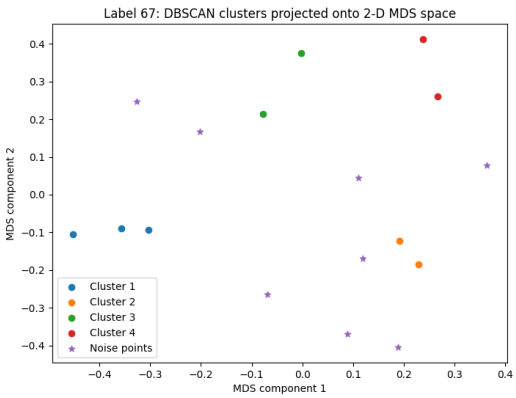

Label 64: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

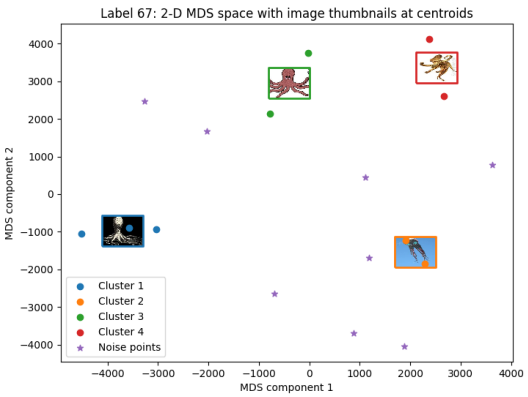

Label 65: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

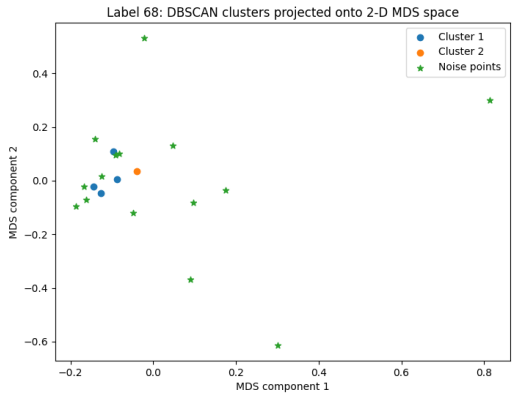

Label 65: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

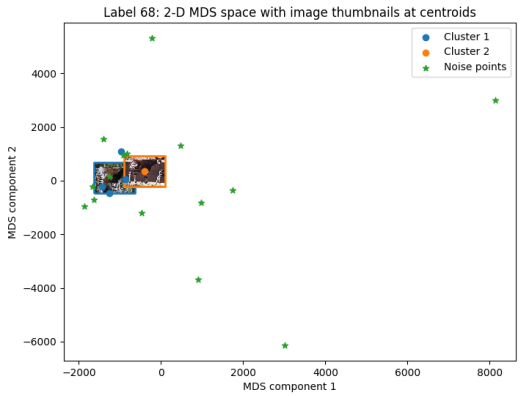

Label 66: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

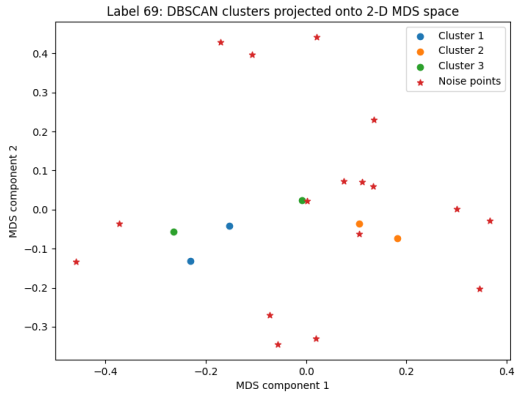

Label 66: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

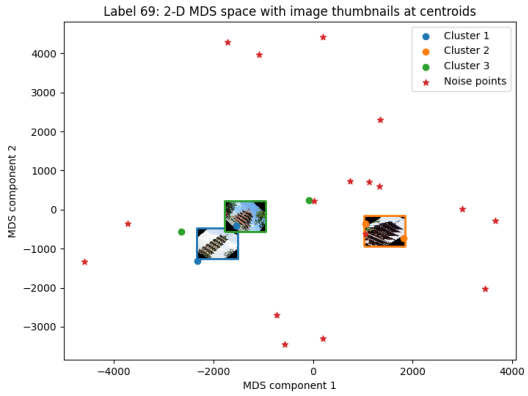

Label 67: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

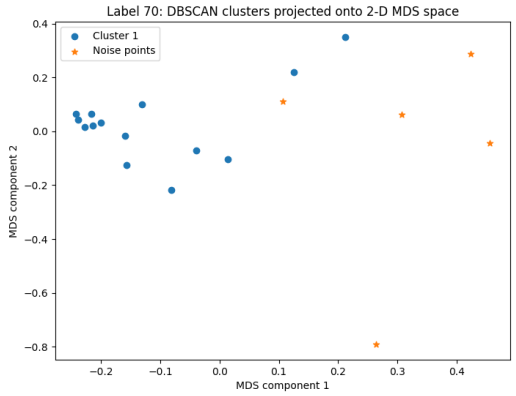

Label 67: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

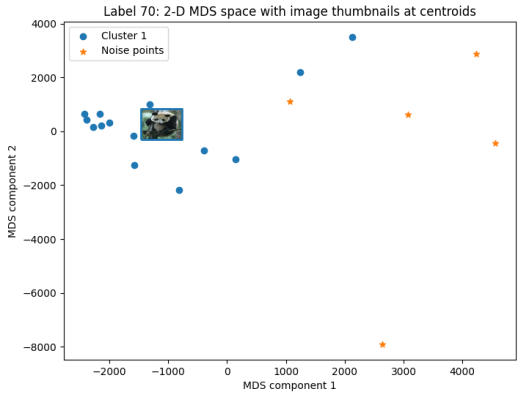

Label 68: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

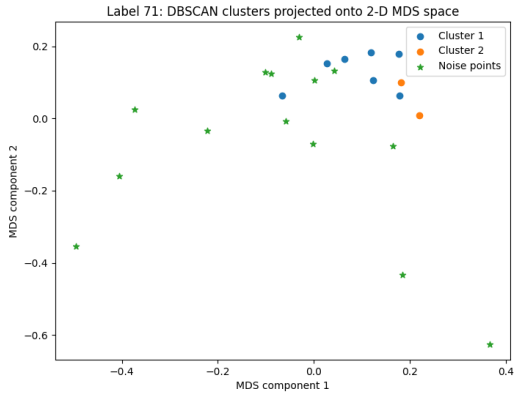

Label 68: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

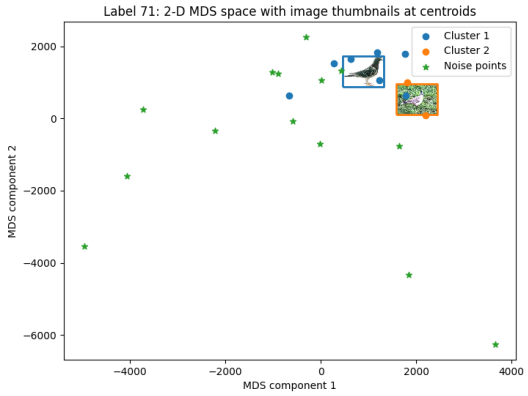

Label 69: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

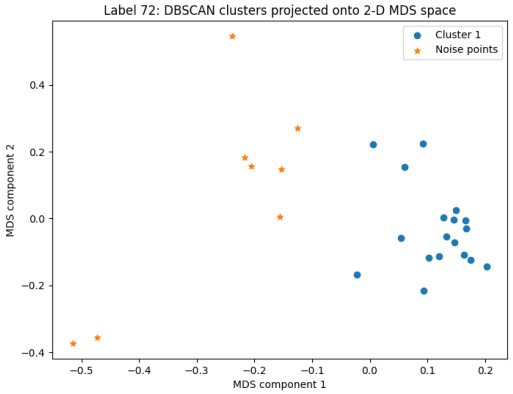

Label 69: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

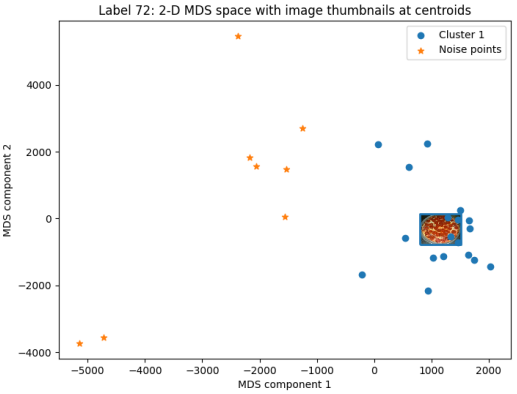

Label 70: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

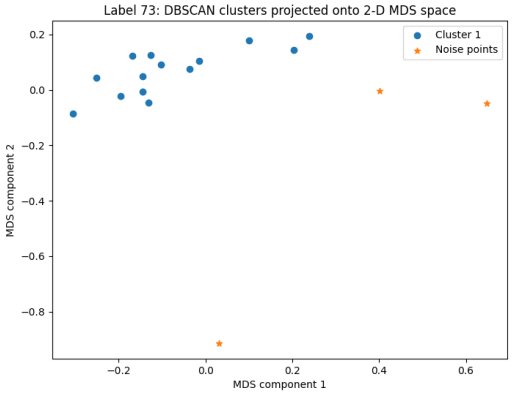

Label 70: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

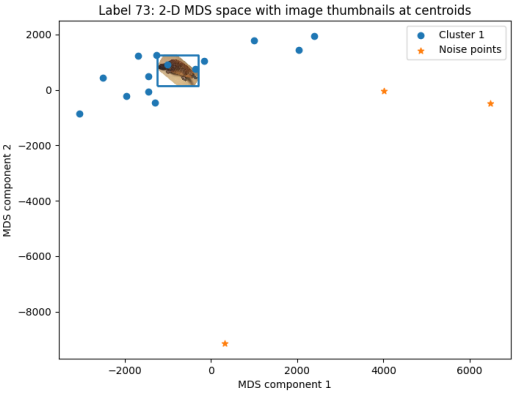

Label 71: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

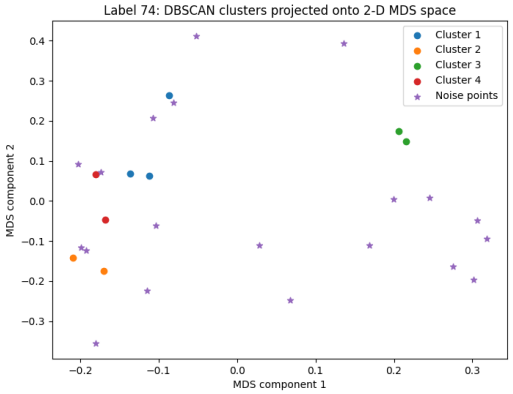

Label 71: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

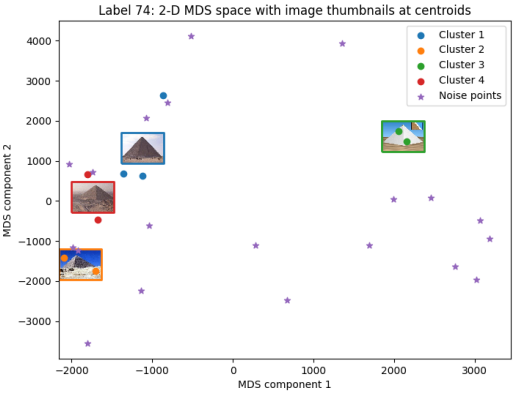

Label 72: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

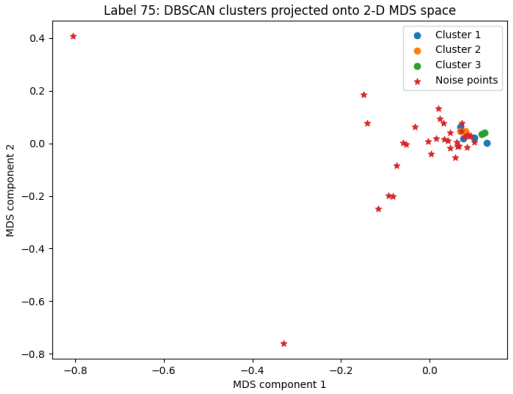

Label 72: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

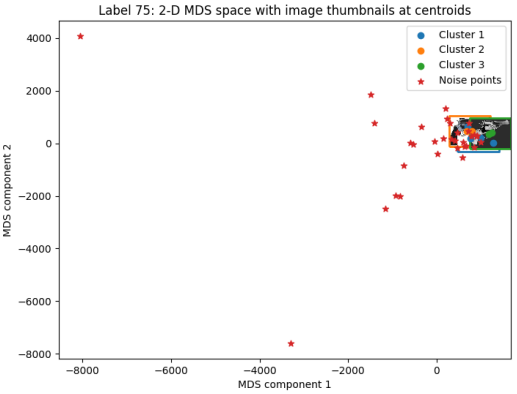

Label 73: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

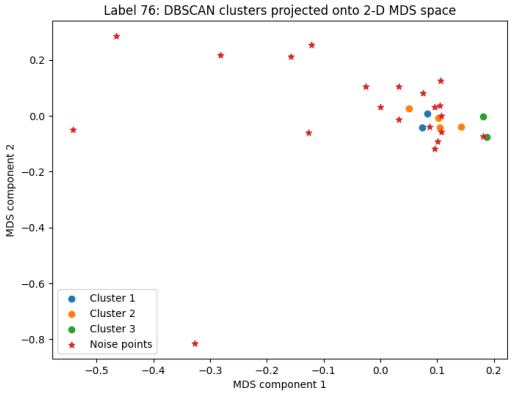

Label 73: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

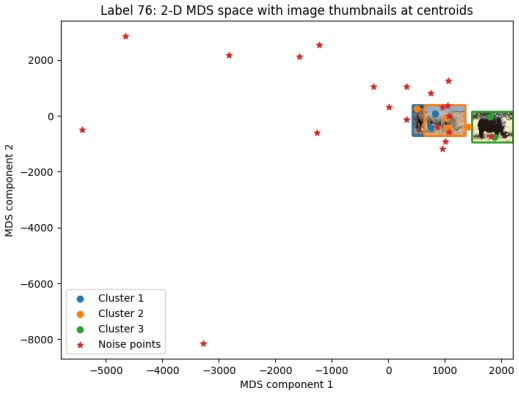

Label 74: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

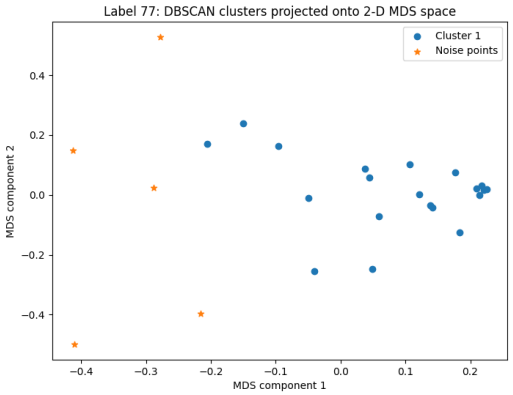

Label 74: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

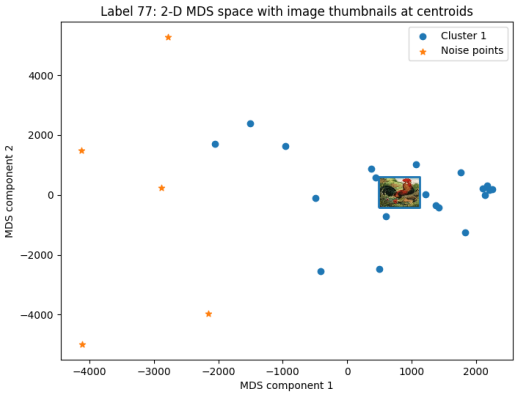

Label 75: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

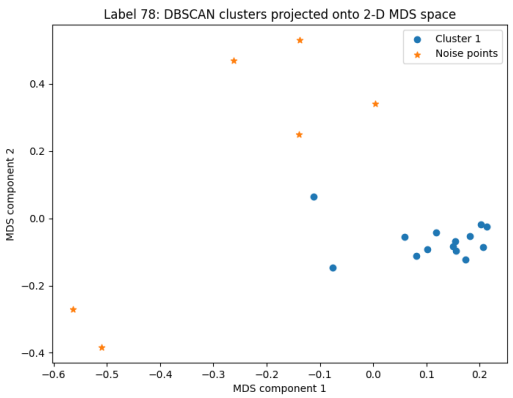

Label 75: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

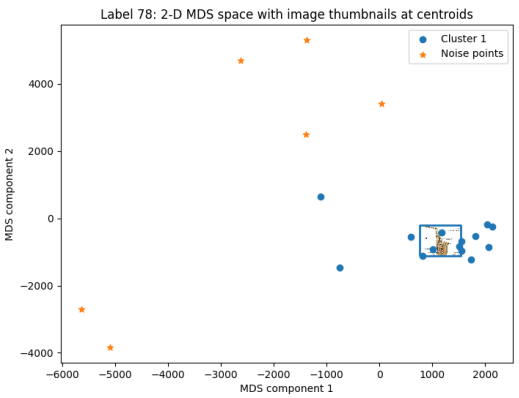

Label 76: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Noise points

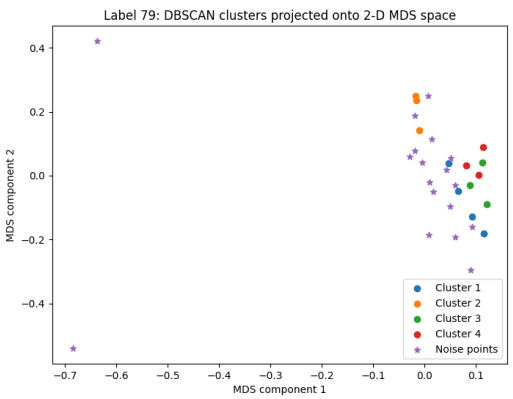

Label 76: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Noise points

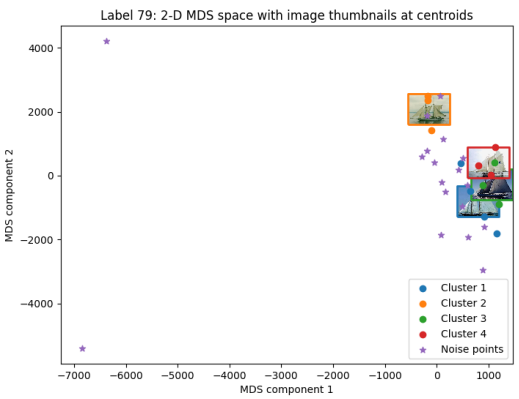

Label 77: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Noise points

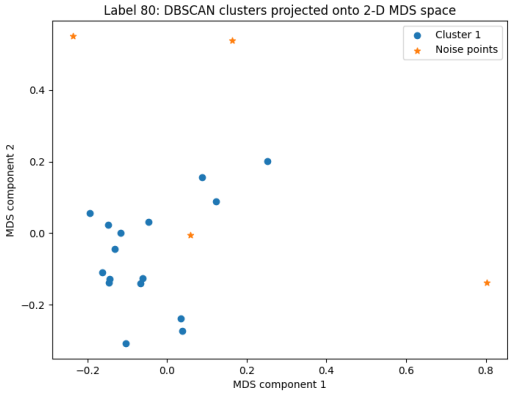

Label 77: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Noise points

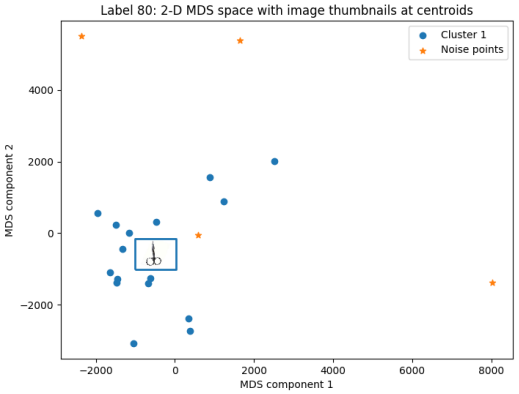

Label 78: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Noise points

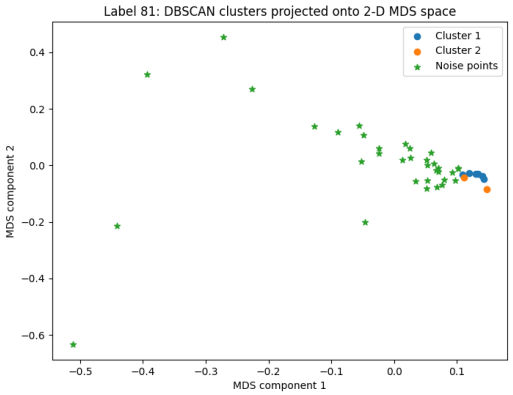

Label 78: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Noise points

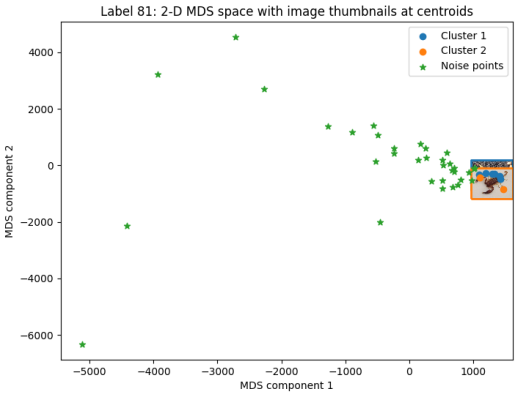

Label 79: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points

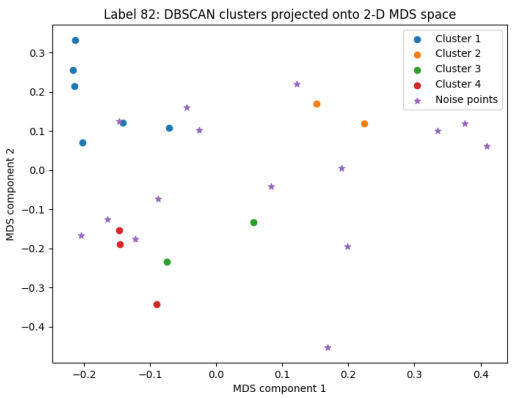

Label 79: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points

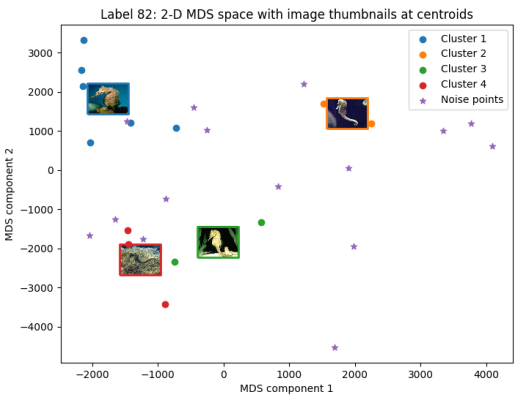

Label 80: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

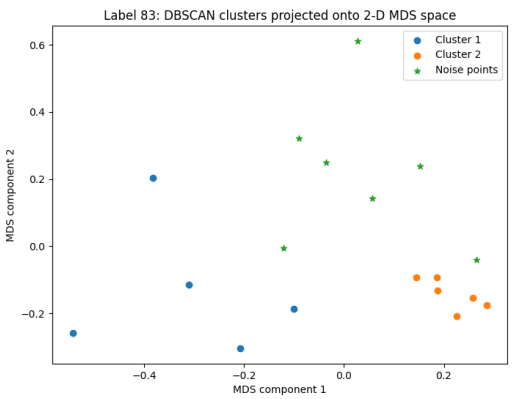

Label 80: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

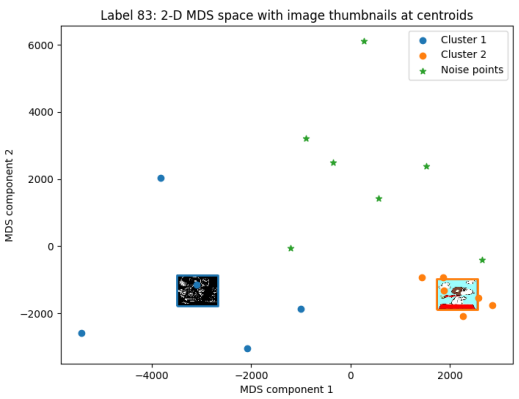

Label 81: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

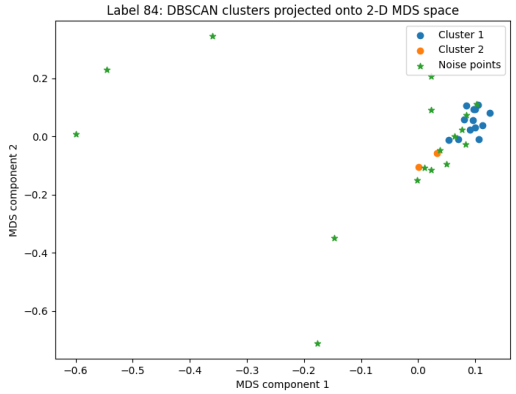

Label 81: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

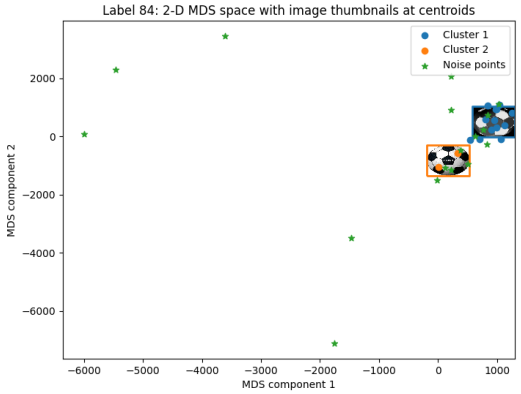

Label 82: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

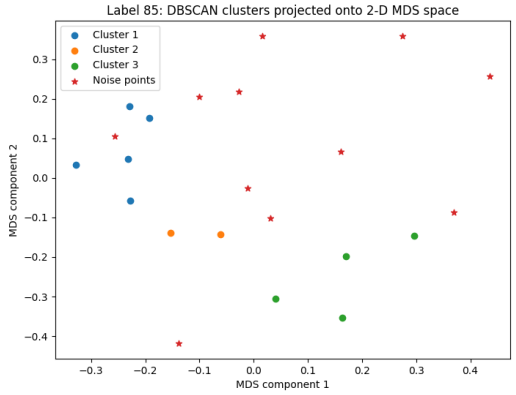

Label 82: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

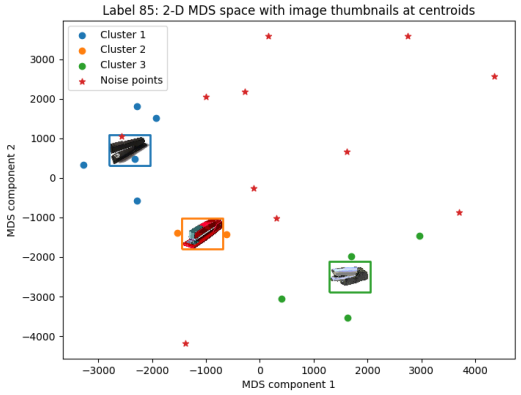

Label 83: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

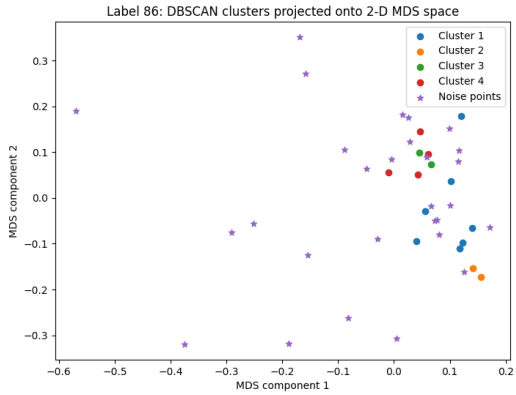

Label 83: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

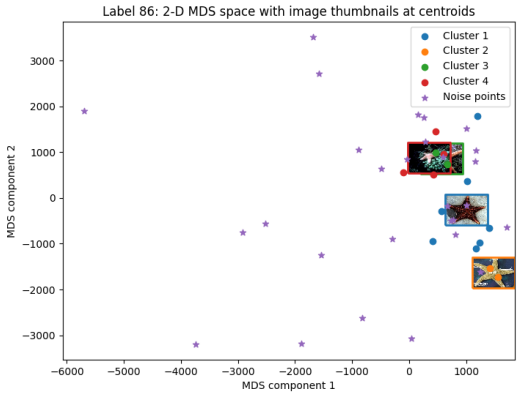

Label 84: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

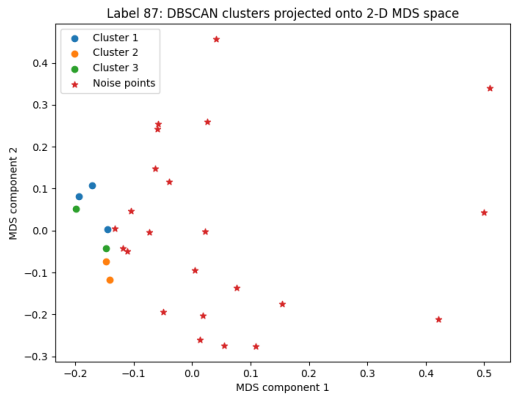

Label 84: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

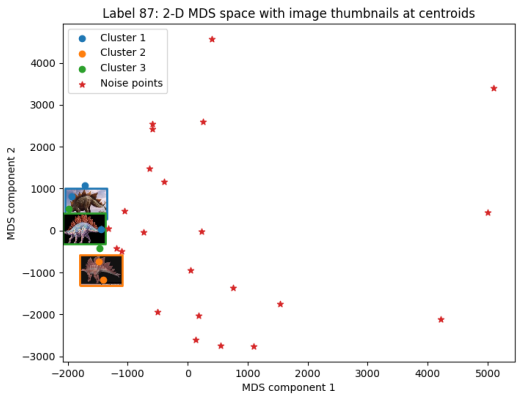

Label 85: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

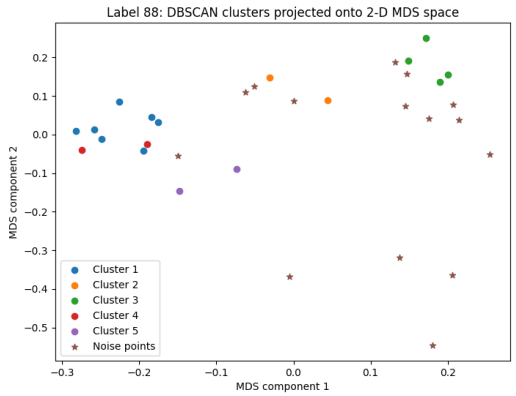

Label 85: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

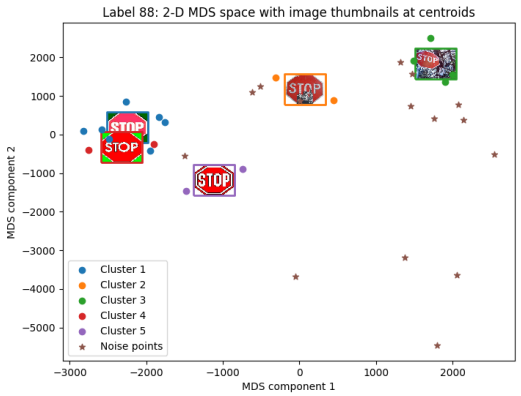

Label 86: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

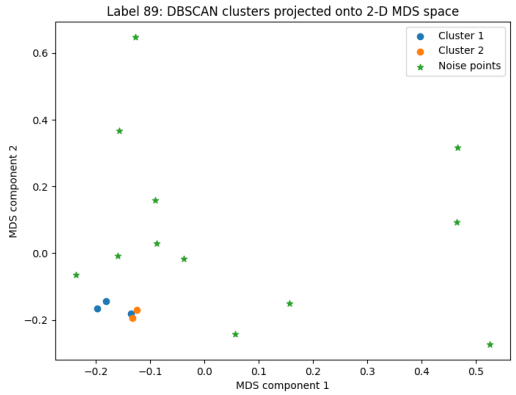

Label 86: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

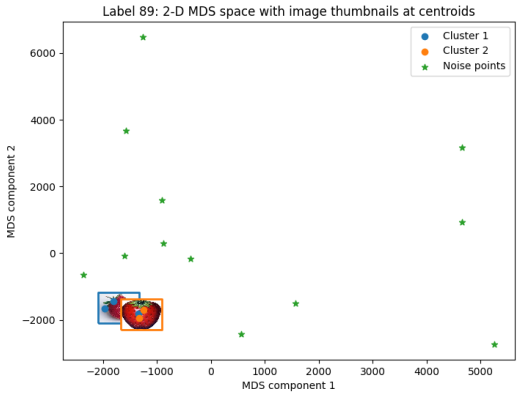

Label 87: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

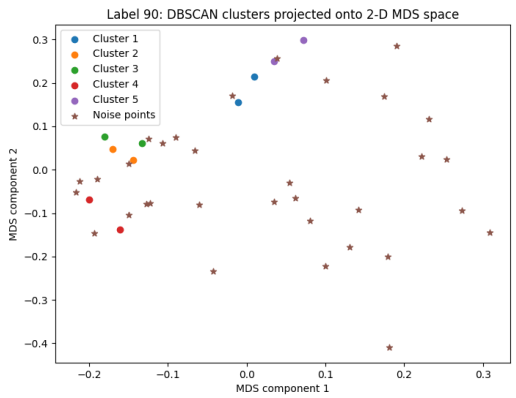

Label 87: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

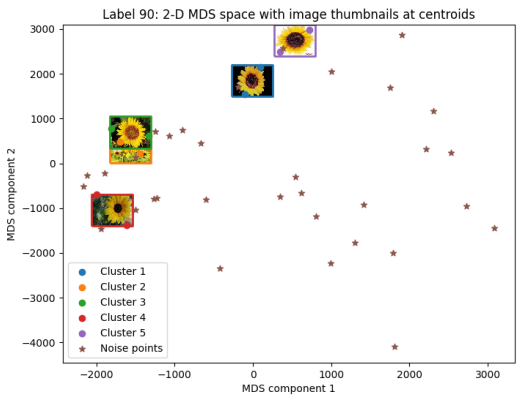

Label 88: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points

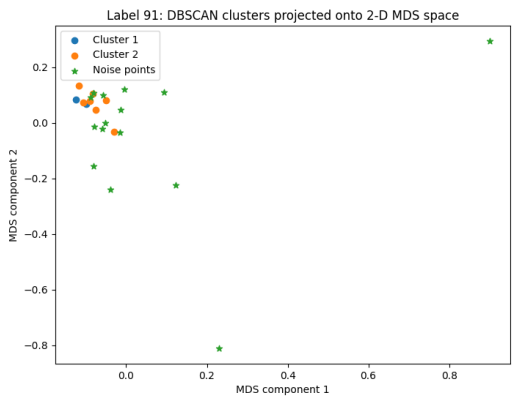

Label 88: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points

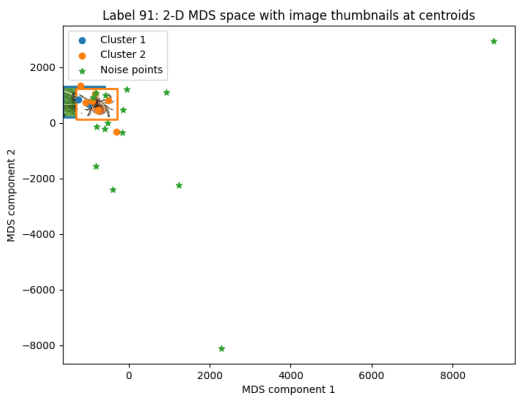

Label 89: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Noise points

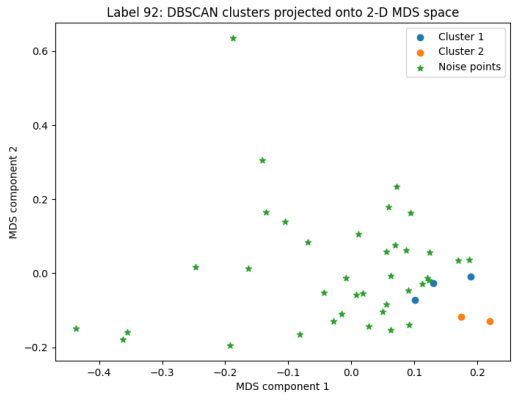

Label 89: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Noise points

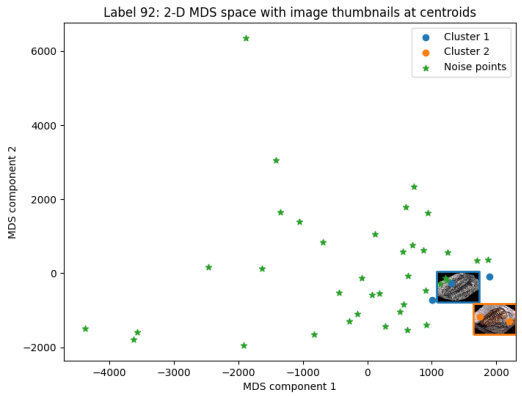

Label 90: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points

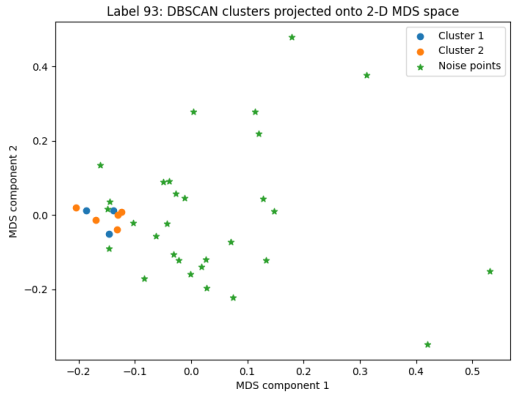

Label 90: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points

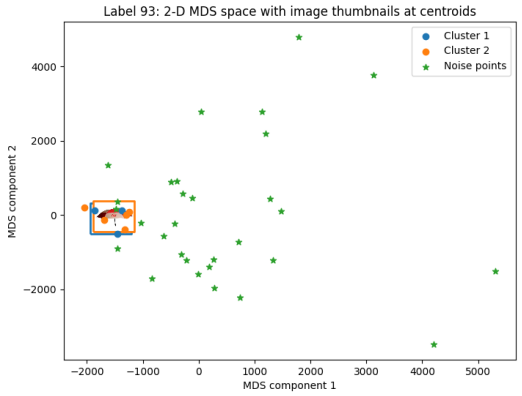

Label 91: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Noise points

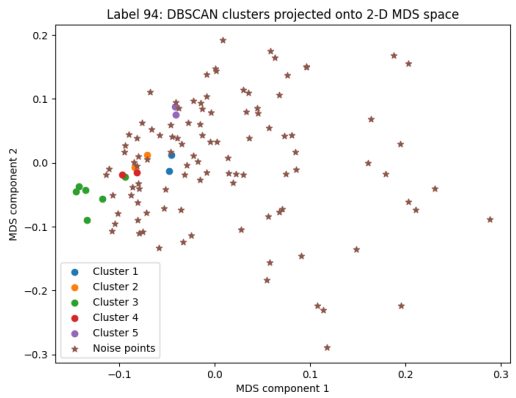

Label 91: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Noise points

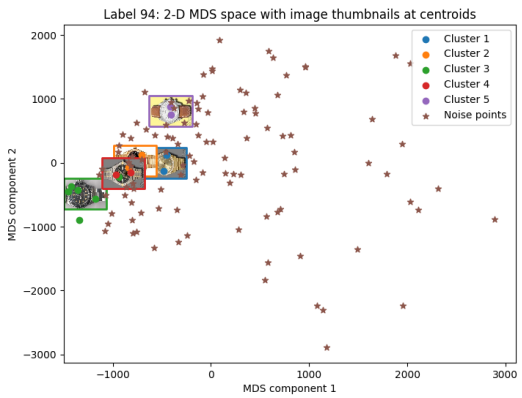

Label 92: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

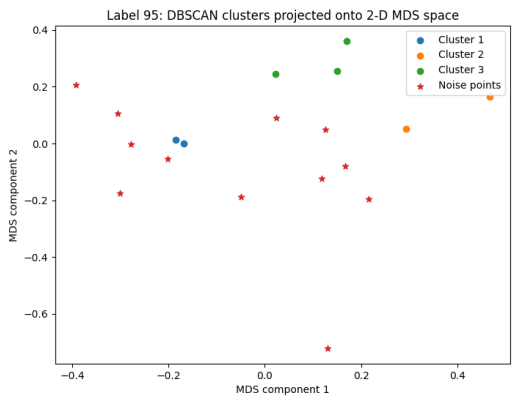

Label 92: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

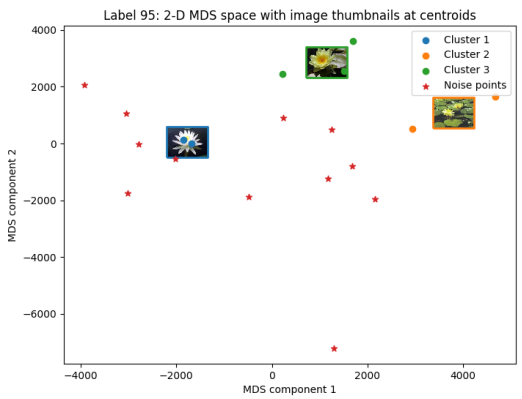

Label 93: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

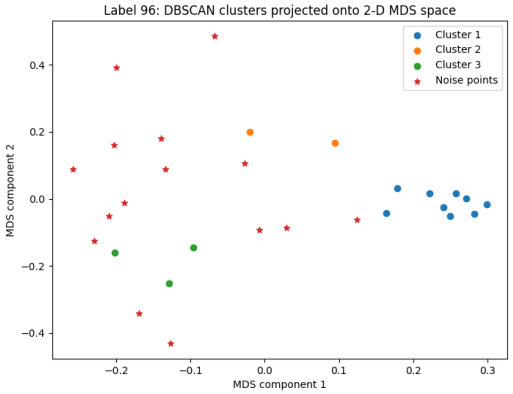

Label 93: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

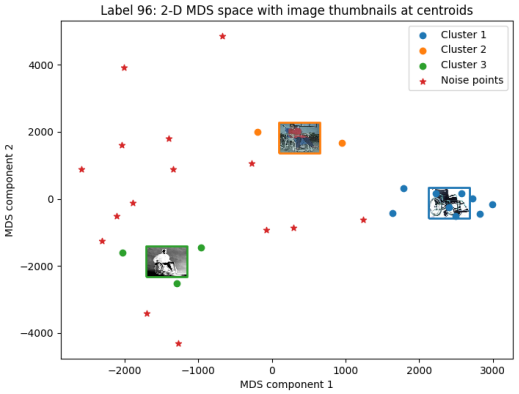

Label 94: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

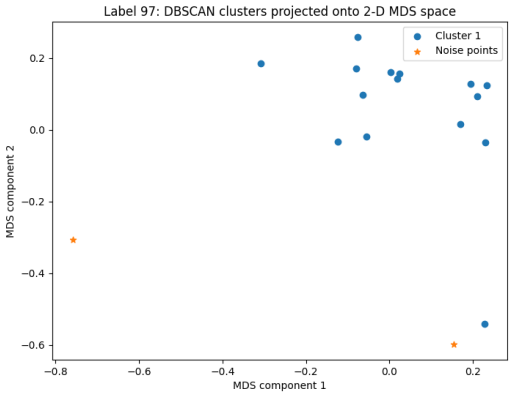

Label 94: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

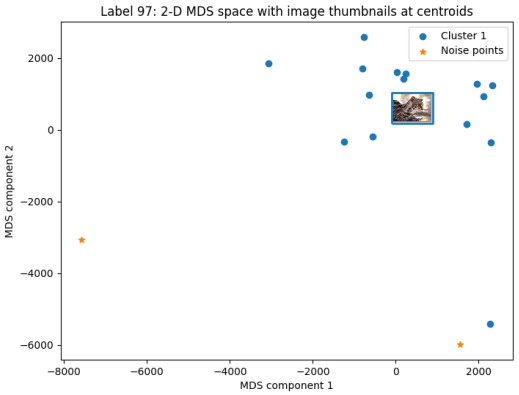

Label 95: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

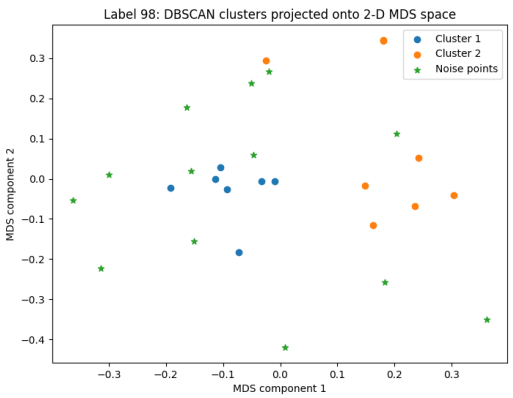

Label 95: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

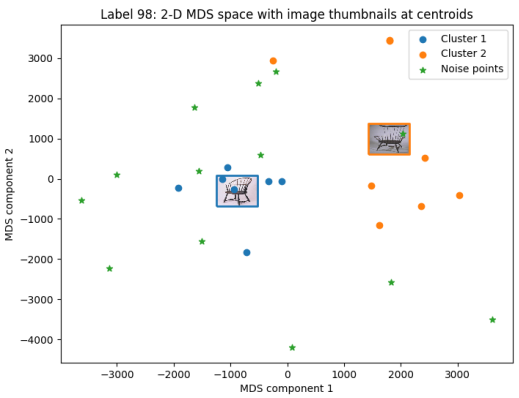

Label 96: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

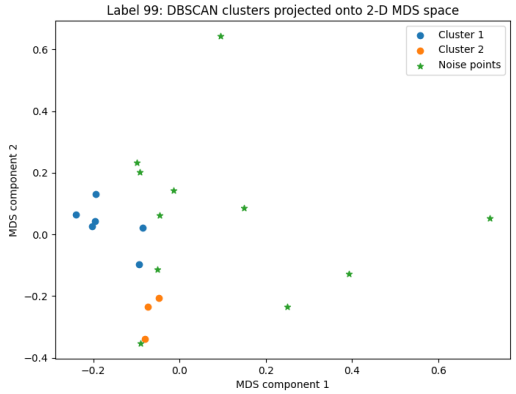

Label 96: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

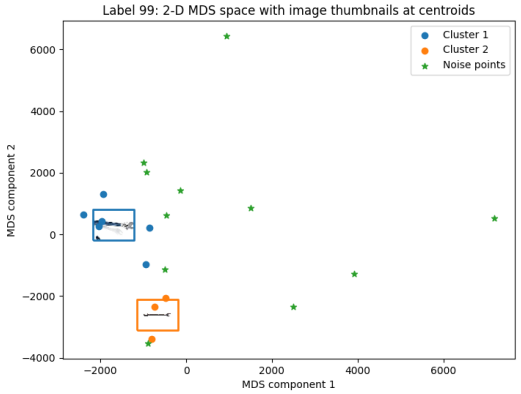

Label 97: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

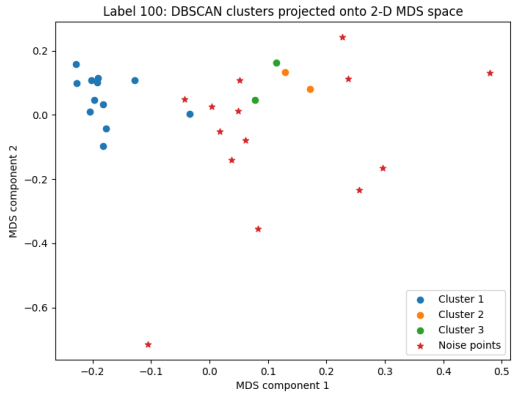

Label 97: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

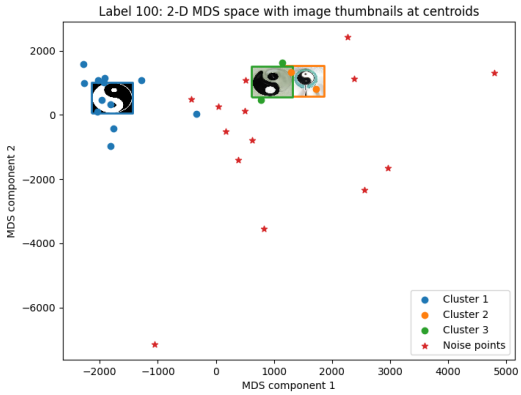

Label 98: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2


Label 98: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

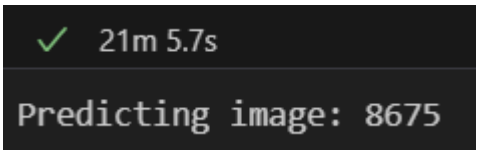








Prediction and metrics:



```
Label 0: Precision=0.95, Recall=0.67, F1-score=0.79
Label 1: Precision=0.75, Recall=0.95, F1-score=0.84
Label 2: Precision=0.94, Recall=1.00, F1-score=0.97
Label 3: Precision=0.99, Recall=1.00, F1-score=0.99
Label 4: Precision=0.90, Recall=1.00, F1-score=0.95
Label 5: Precision=0.97, Recall=0.95, F1-score=0.96
Label 6: Precision=0.92, Recall=0.52, F1-score=0.67
Label 7: Precision=0.90, Recall=0.90, F1-score=0.90
Label 8: Precision=0.88, Recall=0.92, F1-score=0.90
Label 9: Precision=0.88, Recall=0.78, F1-score=0.82
Label 10: Precision=0.85, Recall=0.96, F1-score=0.90
Label 11: Precision=0.94, Recall=1.00, F1-score=0.97
Label 12: Precision=1.00, Recall=0.97, F1-score=0.98
Label 13: Precision=0.88, Recall=0.90, F1-score=0.89
Label 14: Precision=0.62, Recall=0.59, F1-score=0.60
Label 15: Precision=0.95, Recall=0.90, F1-score=0.93
Label 16: Precision=0.88, Recall=0.46, F1-score=0.60
Label 17: Precision=0.89, Recall=0.96, F1-score=0.92
Label 18: Precision=0.95, Recall=0.90, F1-score=0.93
Label 19: Precision=0.97, Recall=1.00, F1-score=0.98
Label 20: Precision=1.00, Recall=0.74, F1-score=0.85
Label 21: Precision=1.00, Recall=0.23, F1-score=0.38
Label 22: Precision=0.83, Recall=0.32, F1-score=0.47
Label 23: Precision=0.89, Recall=0.58, F1-score=0.70
Label 24: Precision=0.71, Recall=0.62, F1-score=0.67
Label 25: Precision=0.76, Recall=0.85, F1-score=0.81
Label 26: Precision=1.00, Recall=0.73, F1-score=0.84
Label 27: Precision=0.60, Recall=0.86, F1-score=0.71
Label 28: Precision=0.70, Recall=0.76, F1-score=0.73
Label 29: Precision=0.90, Recall=0.72, F1-score=0.80
```

```
Label 30: Precision=1.00, Recall=0.86, F1-score=0.93
Label 31: Precision=1.00, Recall=1.00, F1-score=1.00
Label 32: Precision=0.78, Recall=0.96, F1-score=0.86
Label 33: Precision=0.97, Recall=0.91, F1-score=0.94
Label 34: Precision=0.97, Recall=0.88, F1-score=0.92
Label 35: Precision=0.78, Recall=0.78, F1-score=0.78
Label 36: Precision=0.50, Recall=0.06, F1-score=0.11
Label 37: Precision=0.77, Recall=1.00, F1-score=0.87
Label 38: Precision=0.74, Recall=0.91, F1-score=0.82
Label 39: Precision=0.84, Recall=1.00, F1-score=0.91
Label 40: Precision=0.97, Recall=1.00, F1-score=0.99
Label 41: Precision=0.96, Recall=0.73, F1-score=0.83
Label 42: Precision=0.83, Recall=0.83, F1-score=0.83
Label 43: Precision=0.38, Recall=0.18, F1-score=0.24
Label 44: Precision=0.50, Recall=0.94, F1-score=0.65
Label 45: Precision=0.87, Recall=0.80, F1-score=0.83
Label 46: Precision=1.00, Recall=1.00, F1-score=1.00
Label 47: Precision=0.98, Recall=0.98, F1-score=0.98
Label 48: Precision=1.00, Recall=0.43, F1-score=0.60
Label 49: Precision=0.96, Recall=0.89, F1-score=0.92
Label 50: Precision=0.61, Recall=0.86, F1-score=0.72
Label 51: Precision=0.80, Recall=1.00, F1-score=0.89
Label 52: Precision=0.72, Recall=0.87, F1-score=0.79
Label 53: Precision=0.97, Recall=0.88, F1-score=0.92
Label 54: Precision=0.98, Recall=0.98, F1-score=0.98
Label 55: Precision=0.84, Recall=0.75, F1-score=0.80
Label 56: Precision=0.52, Recall=0.97, F1-score=0.67
Label 57: Precision=0.80, Recall=1.00, F1-score=0.89
Label 58: Precision=1.00, Recall=0.87, F1-score=0.93
Label 59: Precision=0.55, Recall=0.52, F1-score=0.54
```

```
Label 60: Precision=0.82, Recall=0.55, F1-score=0.65
Label 61: Precision=0.62, Recall=0.71, F1-score=0.67
Label 62: Precision=0.70, Recall=0.95, F1-score=0.81
Label 63: Precision=0.61, Recall=0.93, F1-score=0.74
Label 64: Precision=0.81, Recall=0.81, F1-score=0.81
Label 65: Precision=0.95, Recall=0.97, F1-score=0.96
Label 66: Precision=0.84, Recall=0.96, F1-score=0.90
Label 67: Precision=0.56, Recall=0.28, F1-score=0.37
Label 68: Precision=0.00, Recall=0.00, F1-score=0.00
Label 69: Precision=1.00, Recall=0.88, F1-score=0.93
Label 70: Precision=1.00, Recall=0.79, F1-score=0.88
Label 71: Precision=0.96, Recall=1.00, F1-score=0.98
Label 72: Precision=1.00, Recall=0.81, F1-score=0.90
Label 73: Precision=1.00, Recall=0.76, F1-score=0.87
Label 74: Precision=1.00, Recall=0.57, F1-score=0.73
Label 75: Precision=0.80, Recall=1.00, F1-score=0.89
Label 76: Precision=0.41, Recall=0.87, F1-score=0.55
Label 77: Precision=1.00, Recall=0.83, F1-score=0.91
Label 78: Precision=0.91, Recall=1.00, F1-score=0.95
Label 79: Precision=0.65, Recall=0.81, F1-score=0.72
Label 80: Precision=0.75, Recall=0.95, F1-score=0.84
Label 81: Precision=0.95, Recall=0.98, F1-score=0.96
Label 82: Precision=0.82, Recall=0.48, F1-score=0.61
Label 83: Precision=0.39, Recall=0.82, F1-score=0.53
Label 84: Precision=1.00, Recall=0.88, F1-score=0.93
Label 85: Precision=0.89, Recall=0.74, F1-score=0.81
Label 86: Precision=0.83, Recall=1.00, F1-score=0.91
Label 87: Precision=0.70, Recall=0.90, F1-score=0.79
Label 88: Precision=0.97, Recall=1.00, F1-score=0.98
Label 89: Precision=0.89, Recall=0.89, F1-score=0.89
```

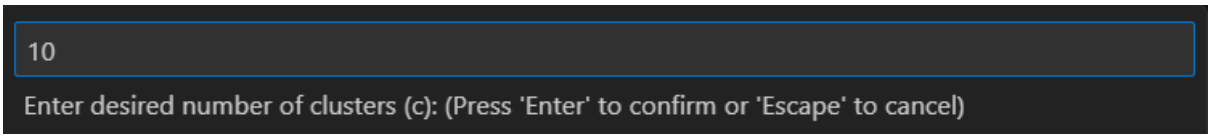
```
Label 90: Precision=0.82, Recall=0.98, F1-score=0.89
Label 91: Precision=1.00, Recall=1.00, F1-score=1.00
Label 92: Precision=0.98, Recall=1.00, F1-score=0.99
Label 93: Precision=0.91, Recall=0.86, F1-score=0.89
Label 94: Precision=0.98, Recall=0.96, F1-score=0.97
Label 95: Precision=0.59, Recall=0.56, F1-score=0.57
Label 96: Precision=0.96, Recall=0.90, F1-score=0.93
Label 97: Precision=0.82, Recall=0.82, F1-score=0.82
Label 98: Precision=0.60, Recall=1.00, F1-score=0.75
Label 99: Precision=0.57, Recall=0.68, F1-score=0.62
Label 100: Precision=0.92, Recall=0.77, F1-score=0.84
Overall Accuracy: 0.86
```

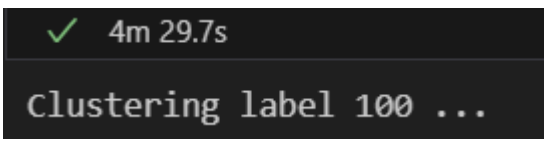
Interpretation of Output:

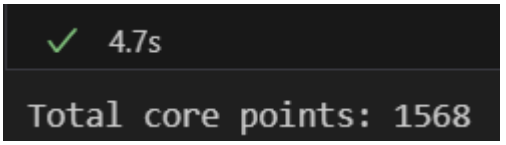
We observe that in many of the labels, we are unable to obtain exactly c clusters and have fewer clusters. Moreover, we see that there are a lot of noise points.

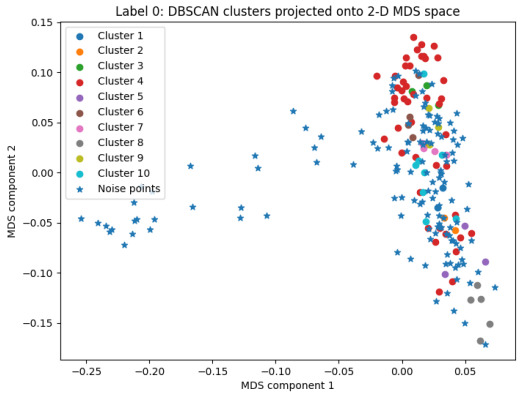
In prediction, we observe that labels with more images usually give better results on both precision and recall, likely because smaller labels don’t have enough inherent clusters to properly classify inputs.

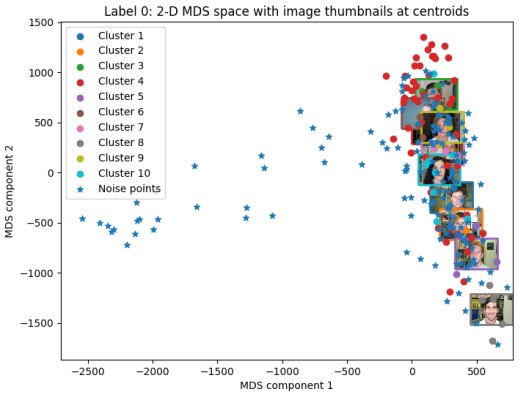
c=10

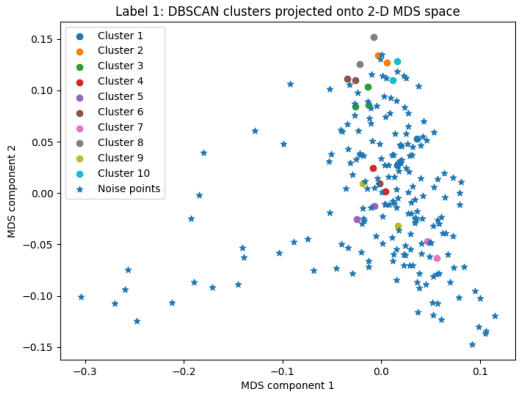


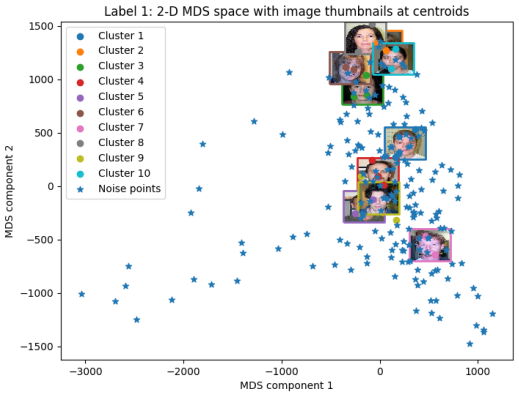


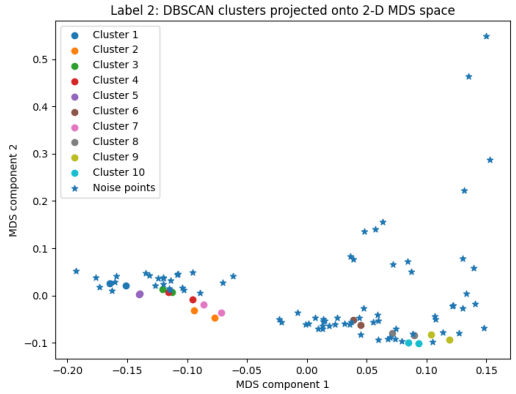


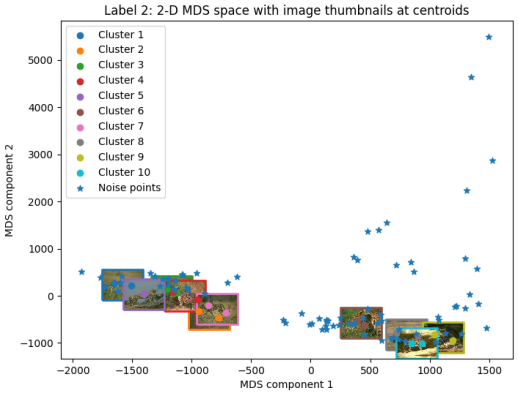


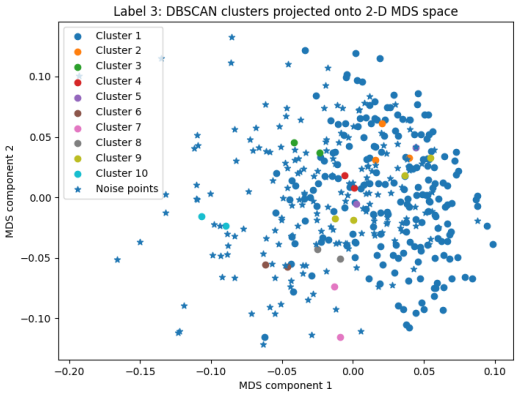


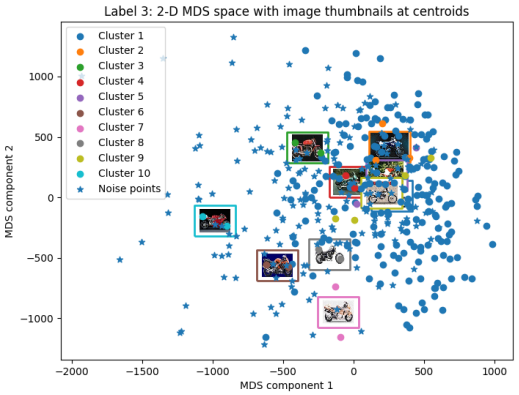


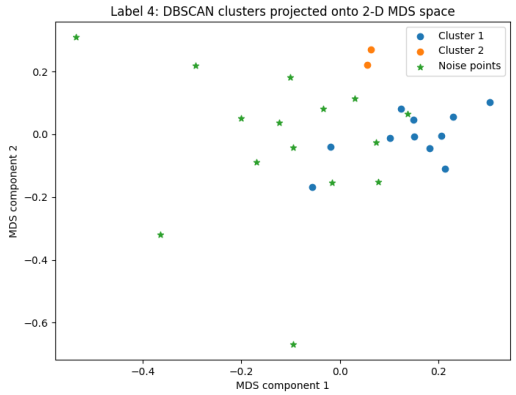


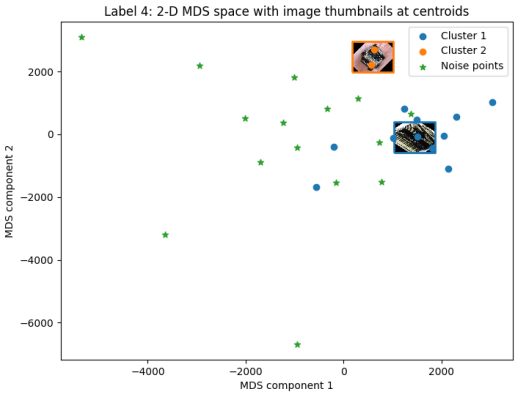

Label 2: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Noise points
MDS component 1
MDS component 2

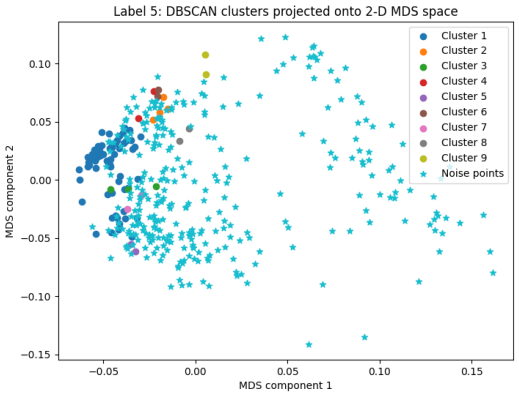

Label 2: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Noise points
MDS component 1
MDS component 2

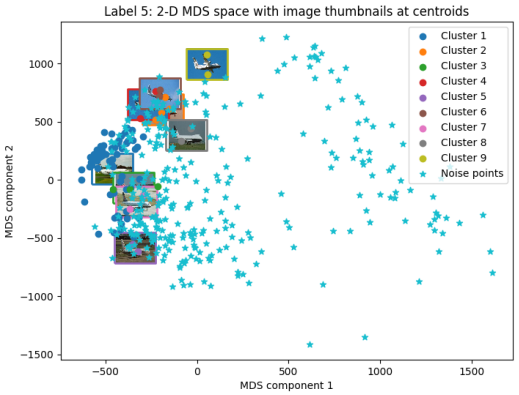

Label 3: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Noise points
MDS component 1
MDS component 2

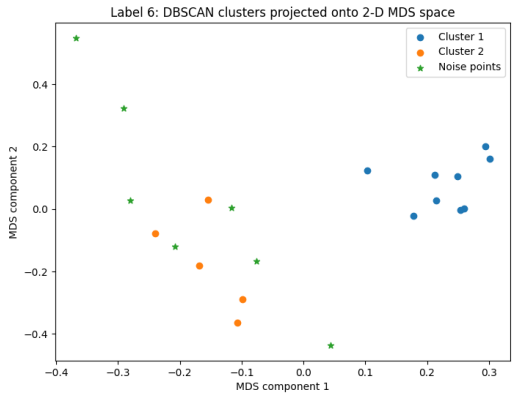

Label 3: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Noise points
MDS component 1
MDS component 2

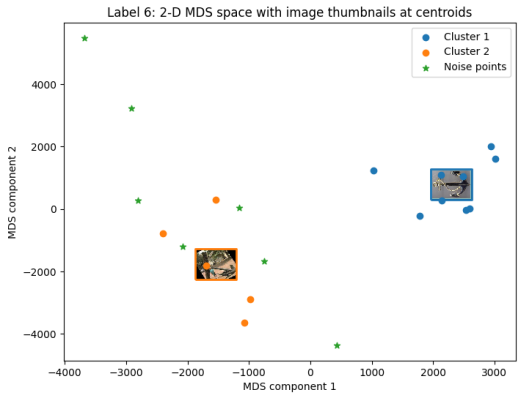

Label 4: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

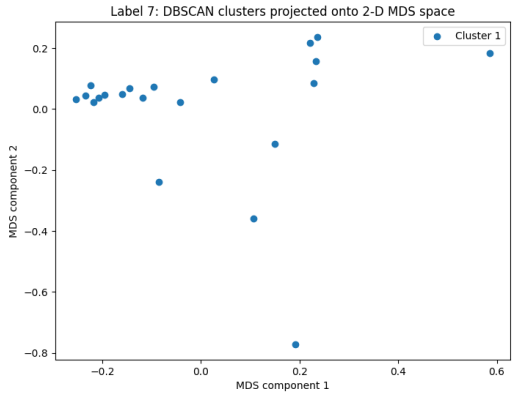

Label 4: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

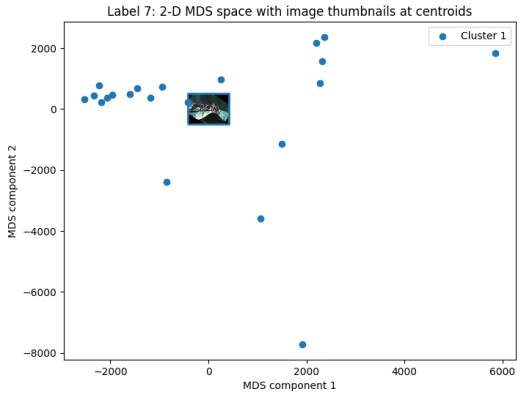

Label 5: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Noise points
MDS component 1
MDS component 2

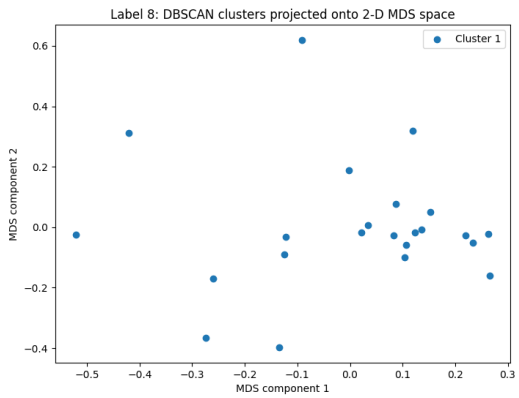

Label 5: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Noise points
MDS component 1
MDS component 2

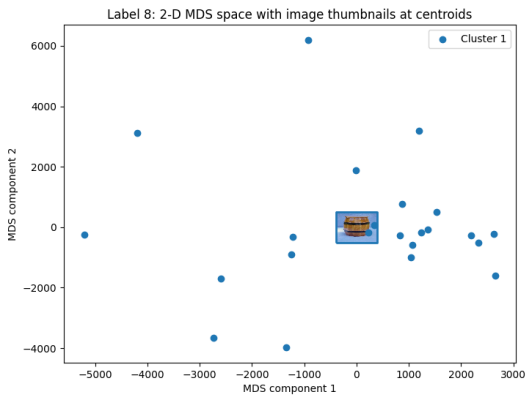

Label 6: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

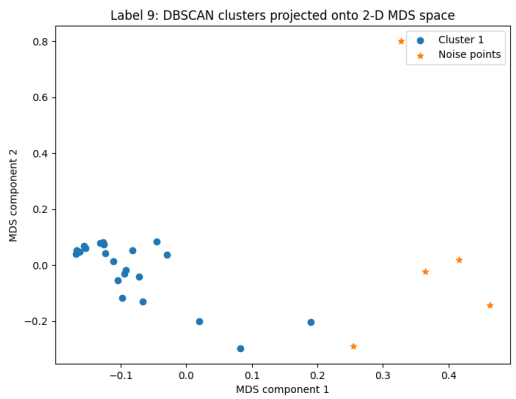

Label 6: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

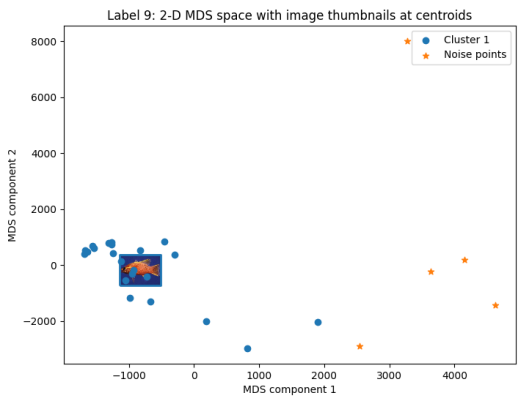

Label 7: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
MDS component 1
MDS component 2

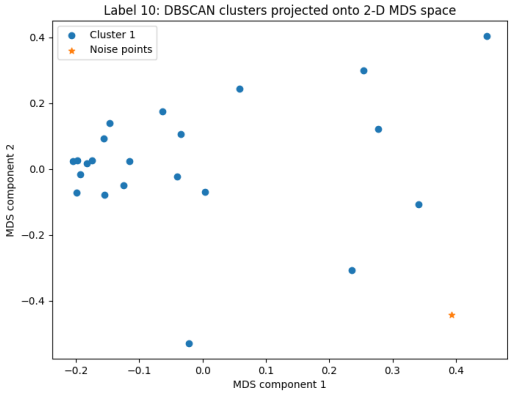

Label 7: 2-D MDS space with image thumbnails at centroids
Cluster 1
MDS component 1
MDS component 2

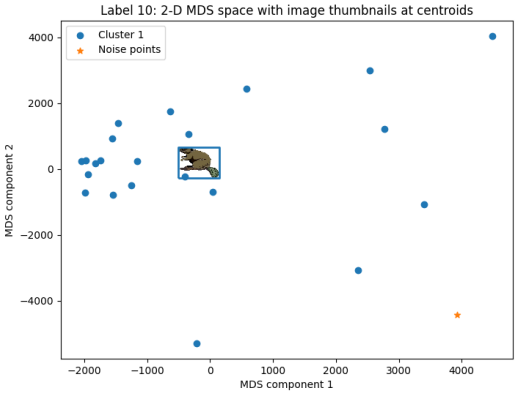

Label 8: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
MDS component 1
MDS component 2

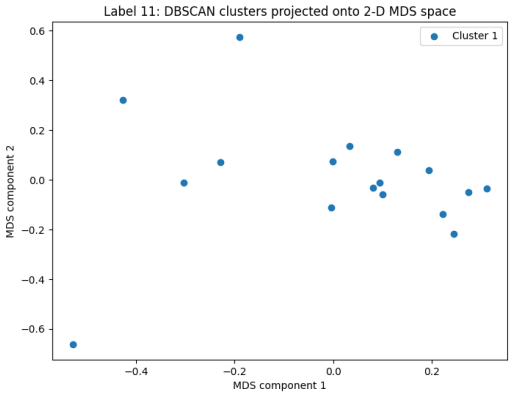

Label 8: 2-D MDS space with image thumbnails at centroids
Cluster 1
MDS component 1
MDS component 2

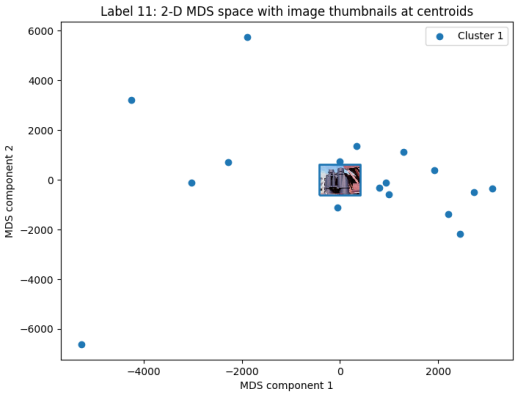

Label 9: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

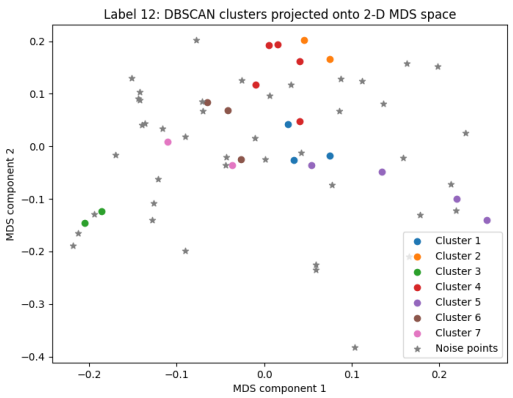

Label 9: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

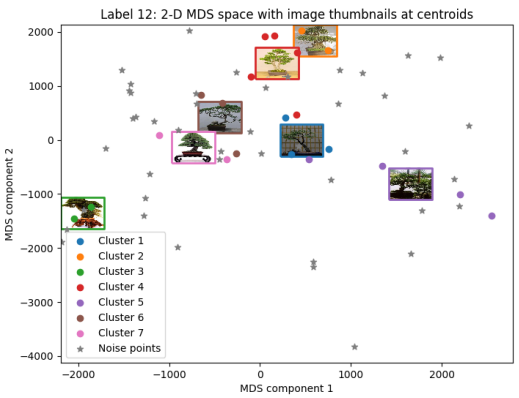

Label 10: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

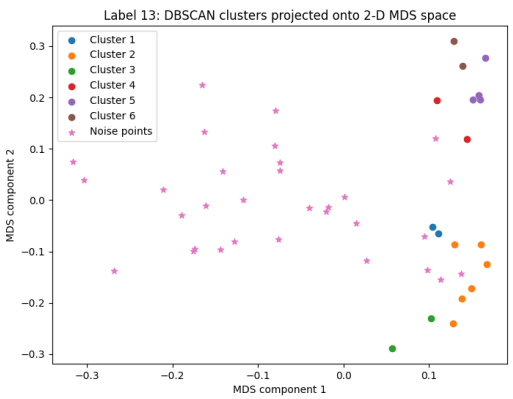

Label 10: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

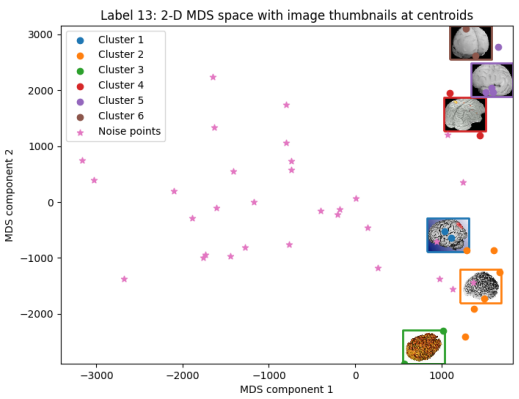

Label 11: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
MDS component 1
MDS component 2

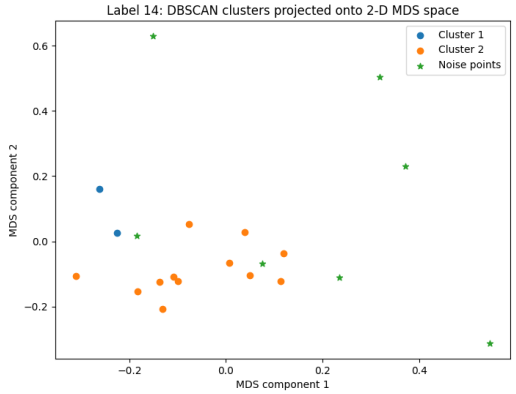

Label 11: 2-D MDS space with image thumbnails at centroids
Cluster 1
MDS component 1
MDS component 2

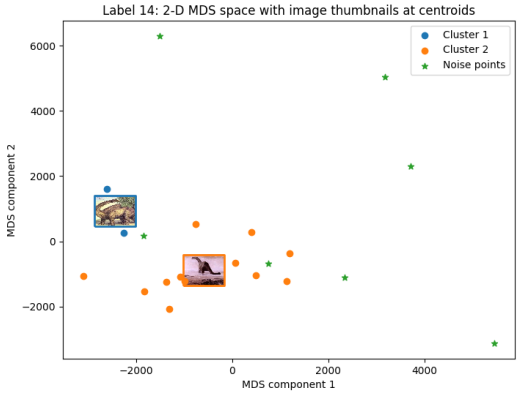

Label 12: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Noise points
MDS component 1
MDS component 2

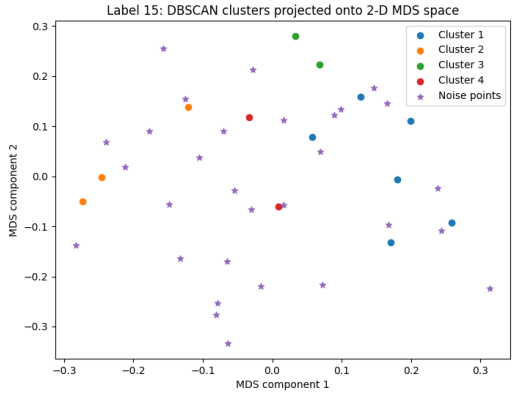

Label 12: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Noise points
MDS component 1
MDS component 2

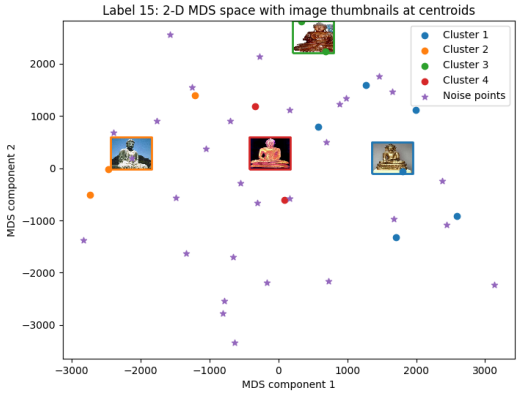

Label 13: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Noise points
MDS component 1
MDS component 2

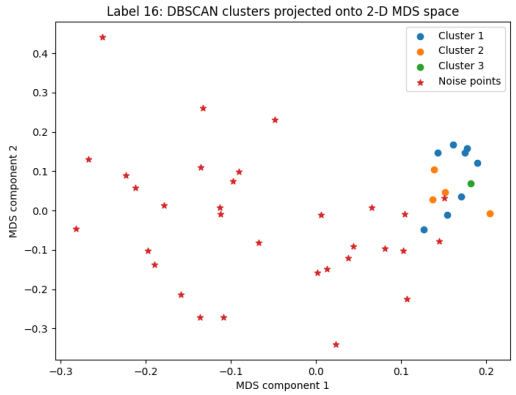

Label 13: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Noise points
MDS component 1
MDS component 2

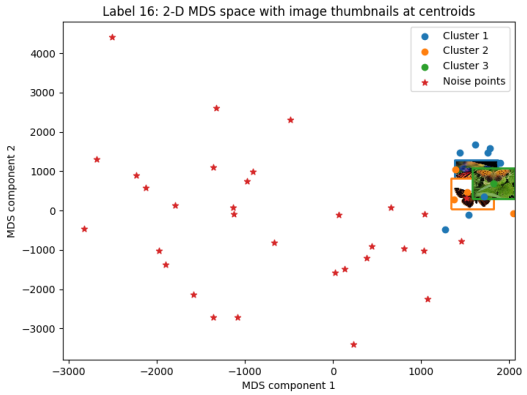

Label 14: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

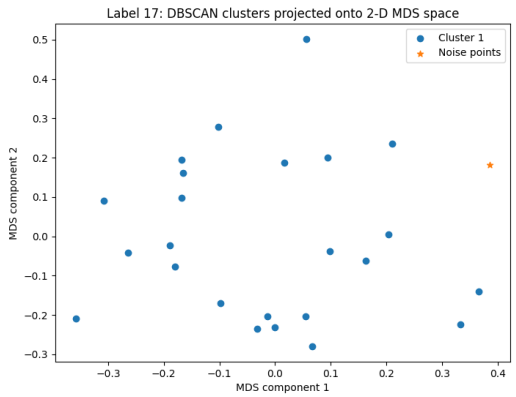

Label 14: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

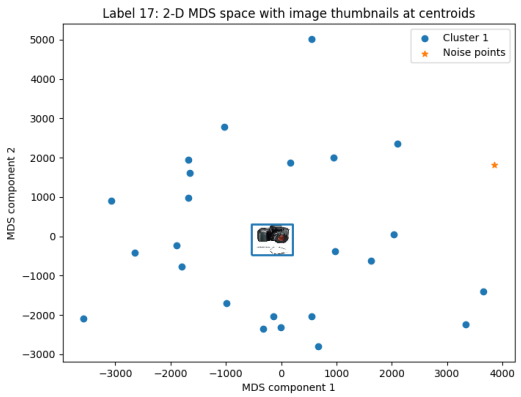

Label 15: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

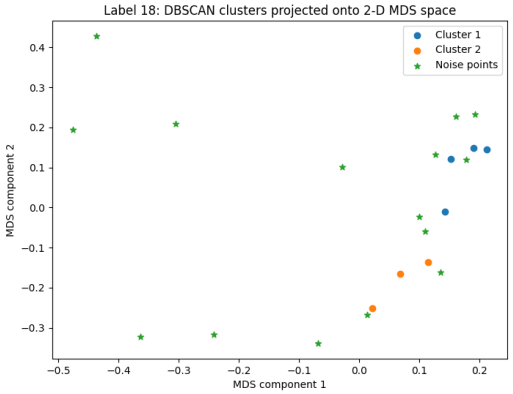

Label 15: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

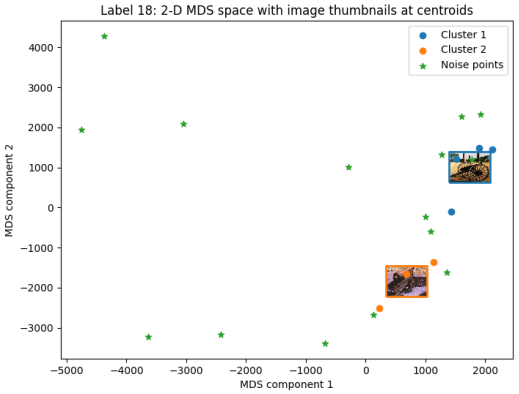

Label 16: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

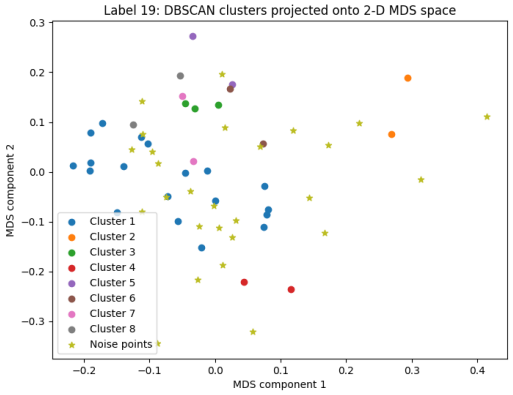

Label 16: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

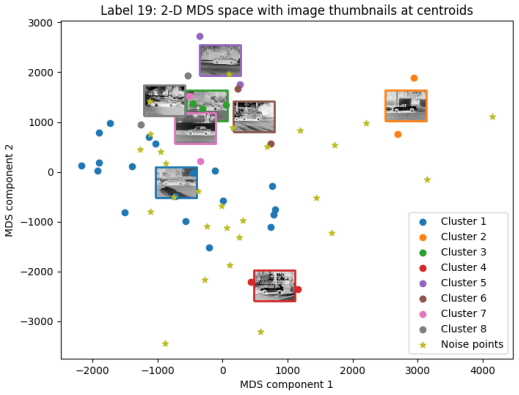

Label 17: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

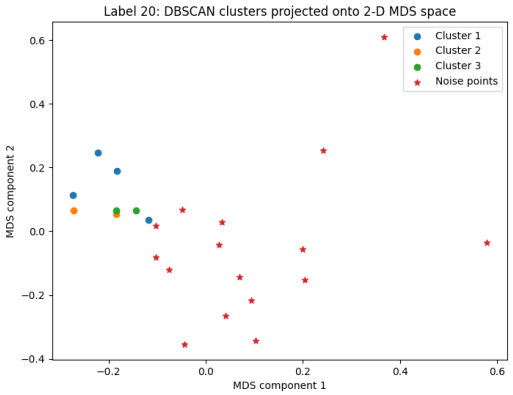

Label 17: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

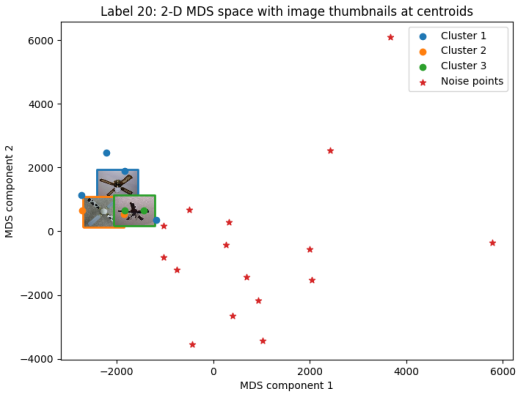

Label 18: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

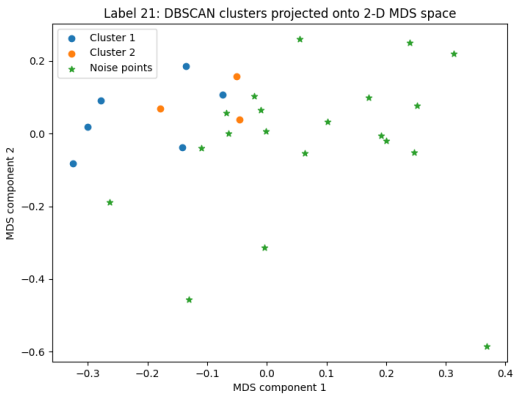

Label 18: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

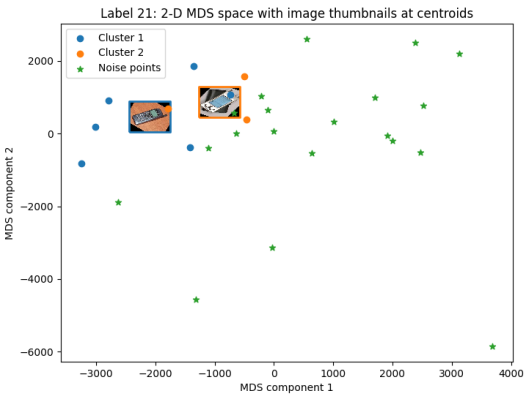

Label 19: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Noise points
MDS component 1
MDS component 2

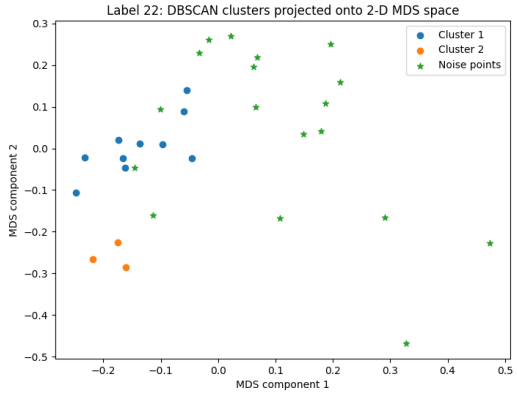

Label 19: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Noise points
MDS component 1
MDS component 2

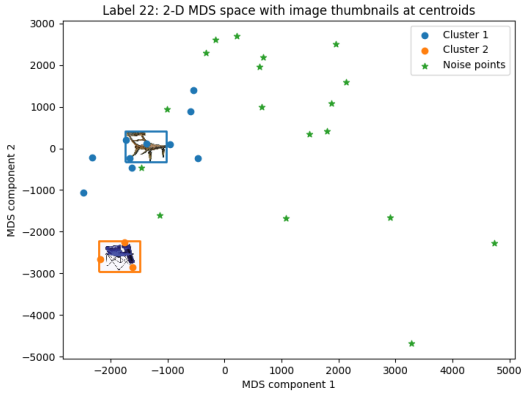

Label 20: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

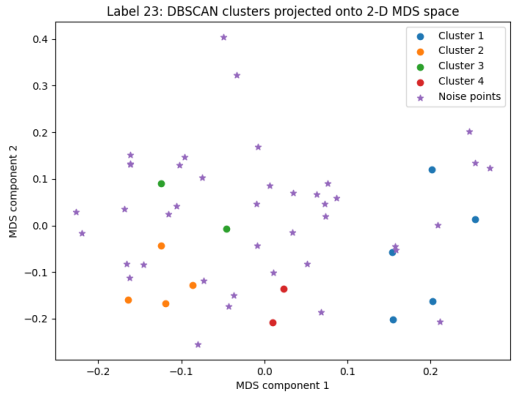

Label 20: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

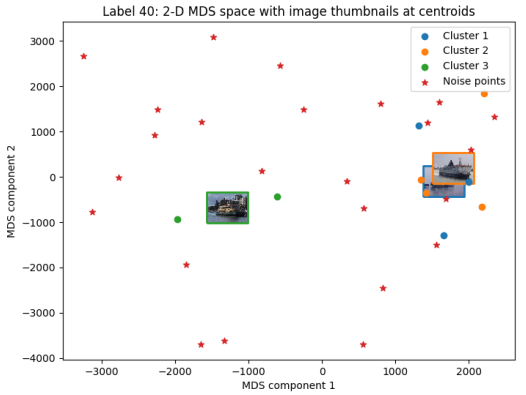

Label 21: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

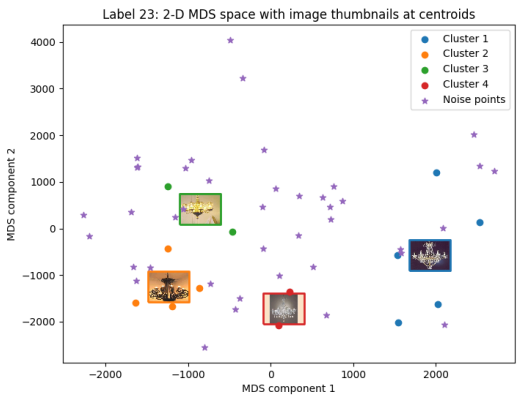

Label 21: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

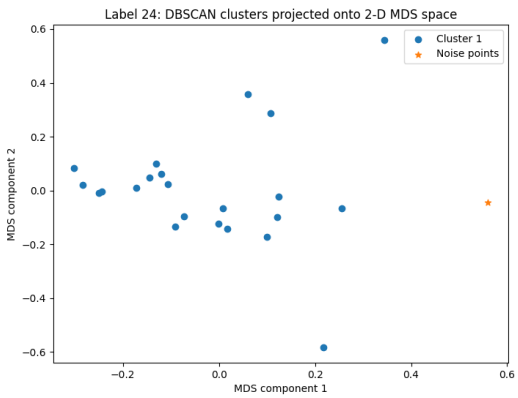

Label 22: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

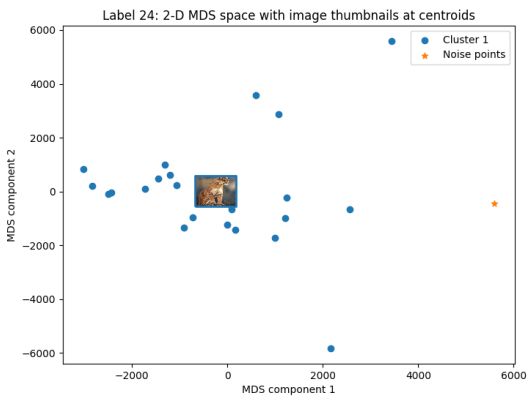

Label 22: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

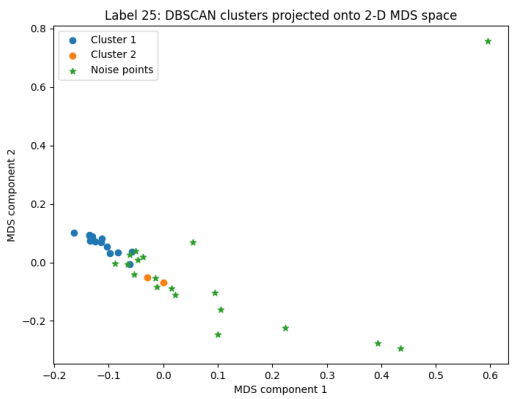

Label 23: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

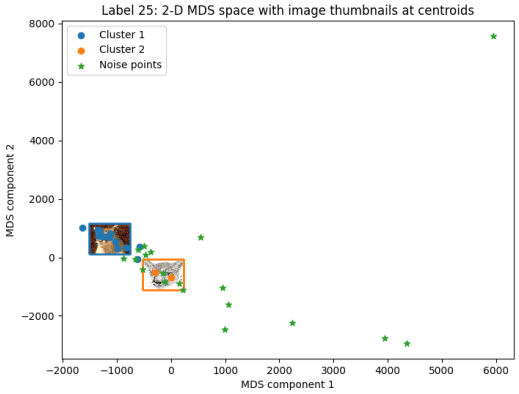

Label 40: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

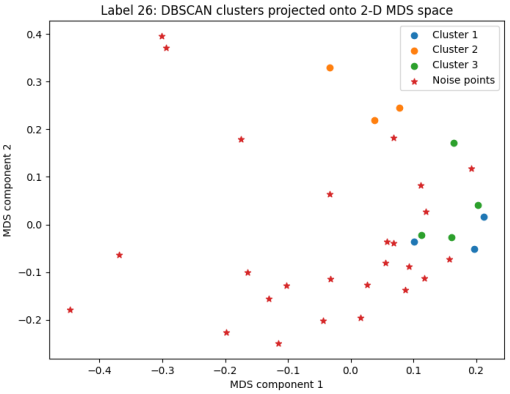

Label 23: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

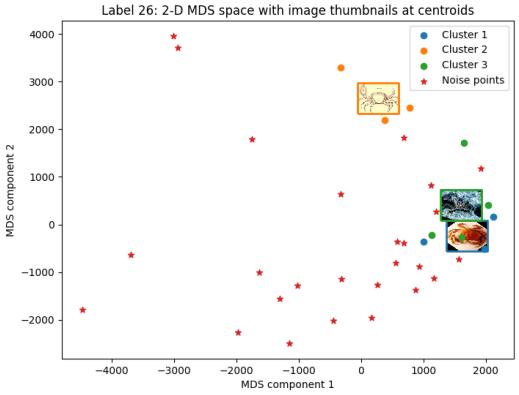

Label 24: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

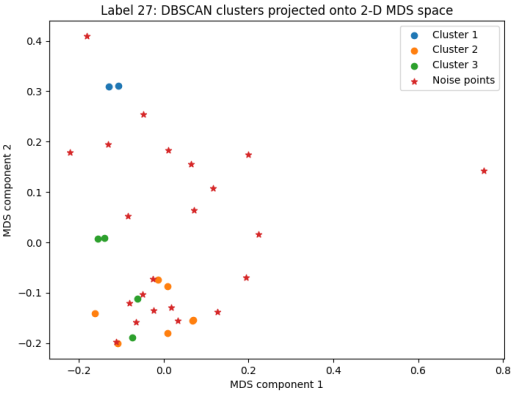

Label 24: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

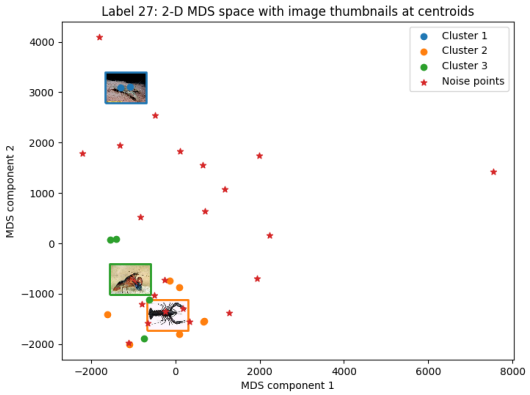

Label 25: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

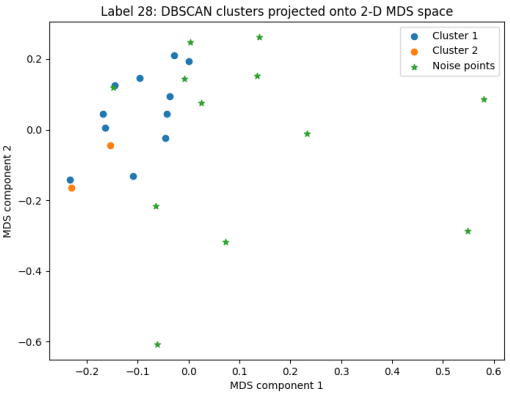

Label 25: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 2
MDS component 1

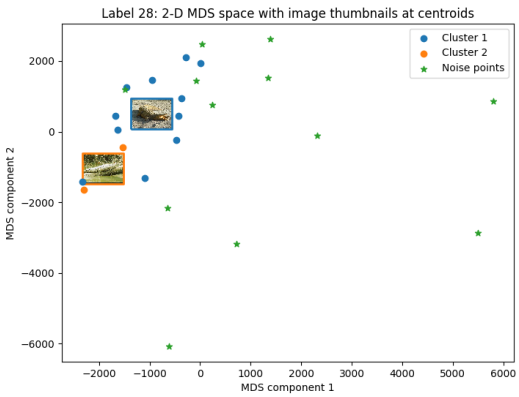

Label 26: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 2
MDS component 1

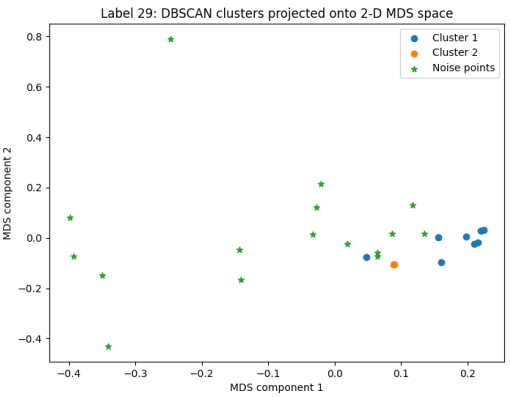

Label 26: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 2
MDS component 1

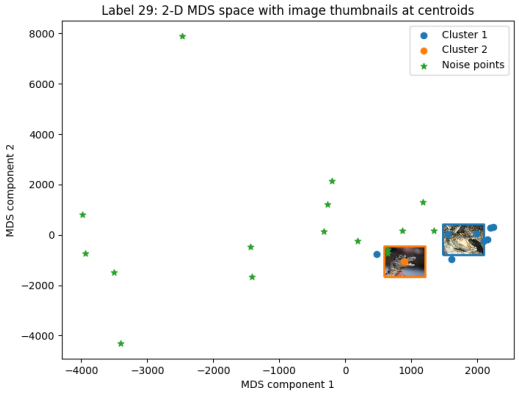

Label 27: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 2
MDS component 1

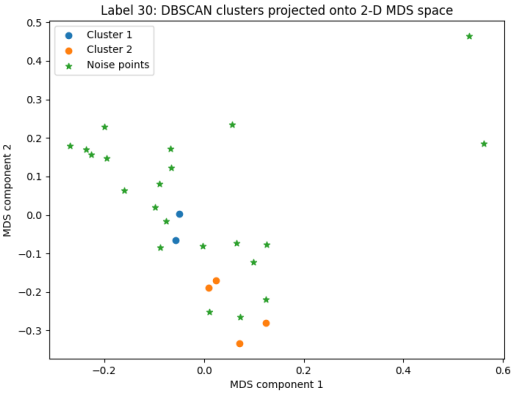

Label 27: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 2
MDS component 1

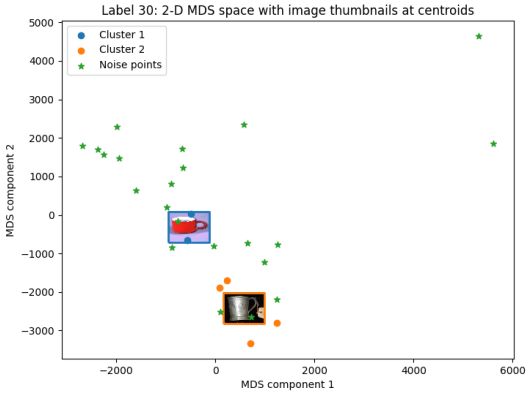

Label 28: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 2
MDS component 1

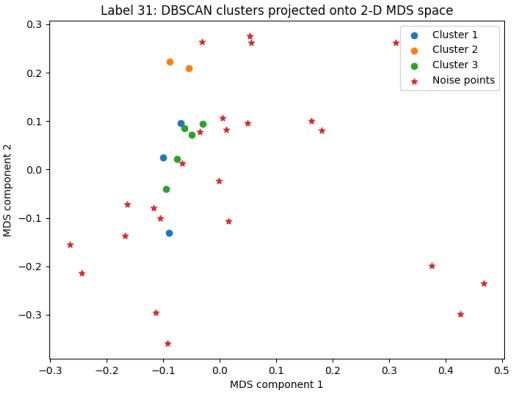

Label 28: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 2
MDS component 1

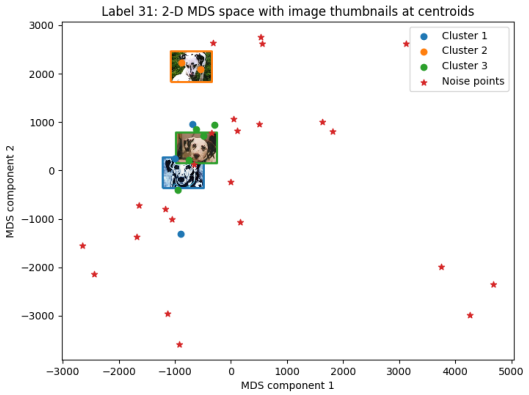

Label 29: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 2
MDS component 1

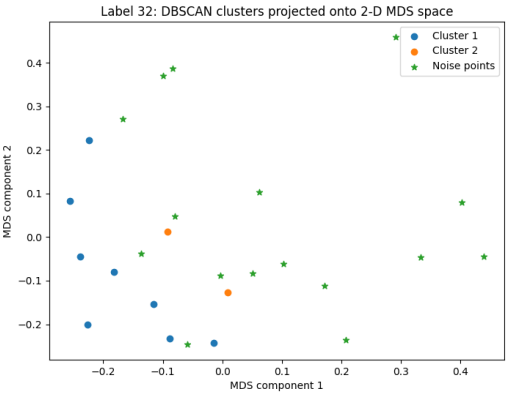

Label 29: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

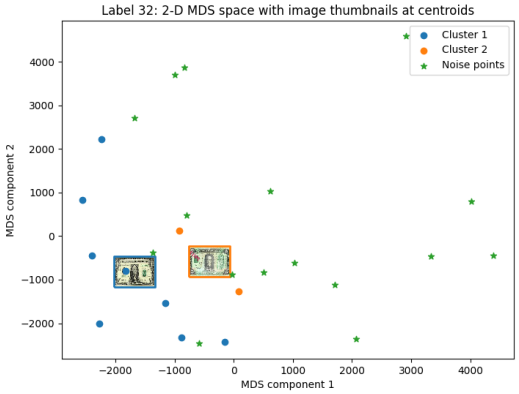

Label 30: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

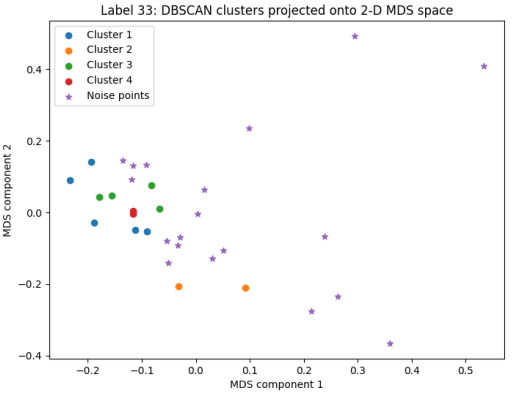

Label 30: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

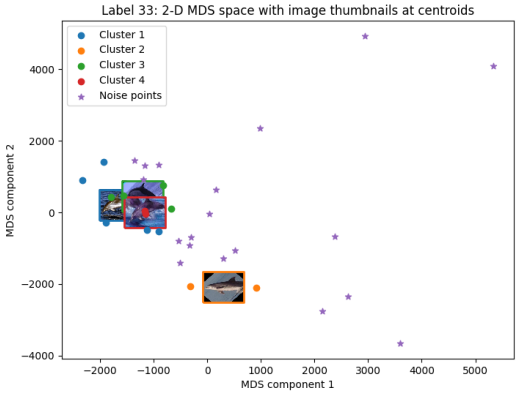

Label 31: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

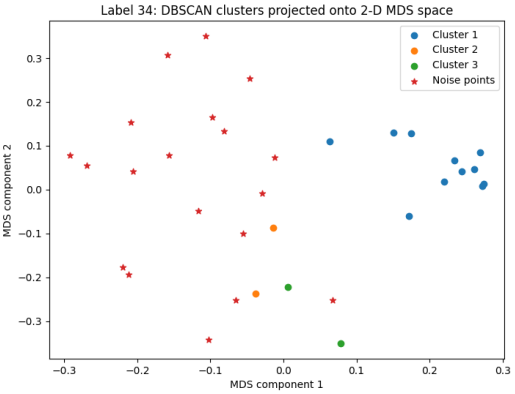

Label 31: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

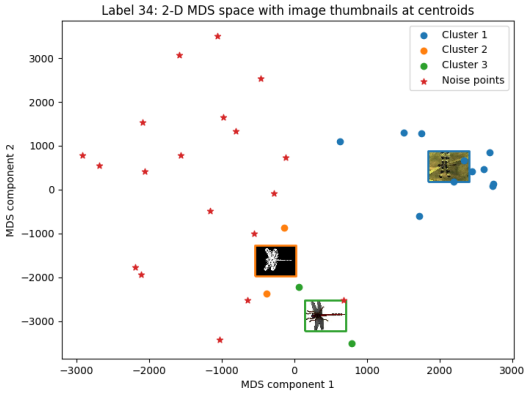

Label 32: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

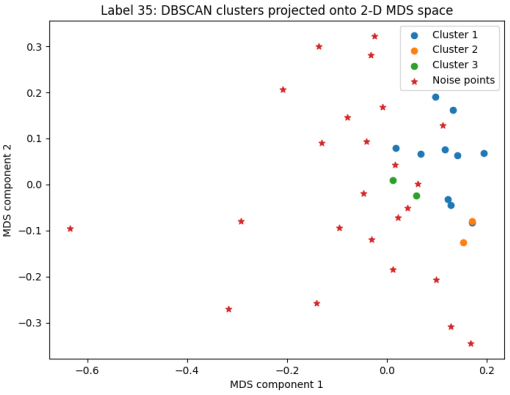

Label 32: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

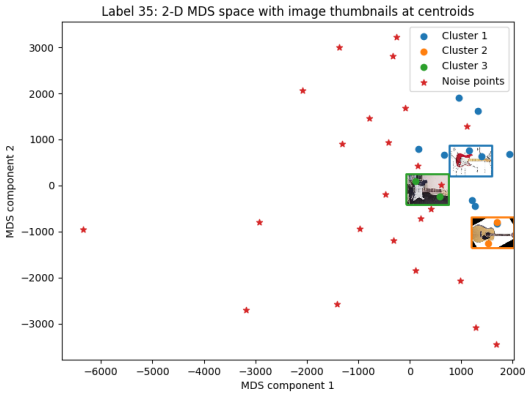

Label 33: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

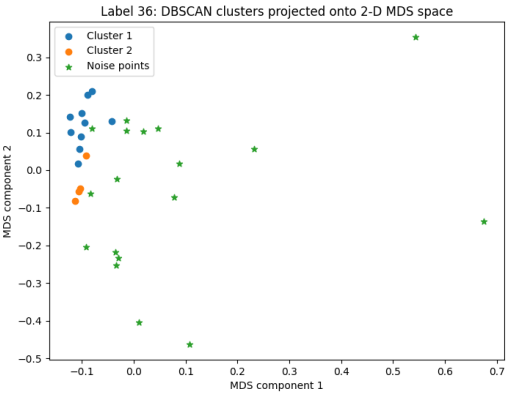

Label 33: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

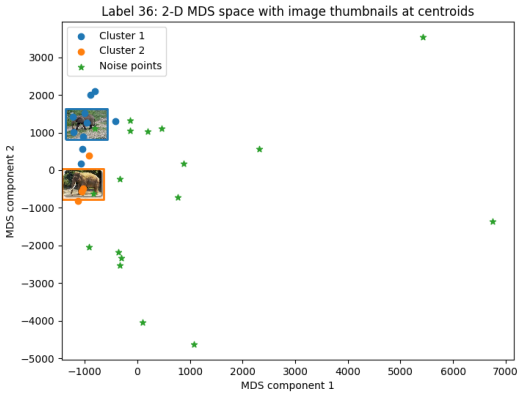

Label 34: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

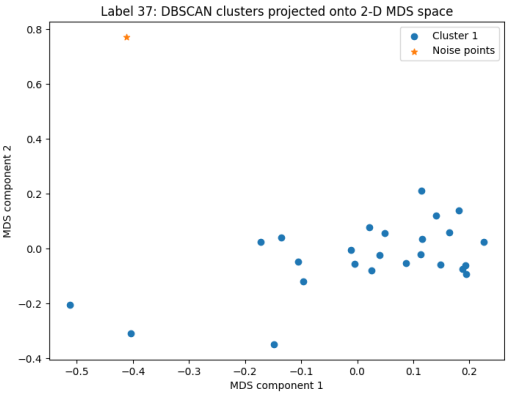

Label 34: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

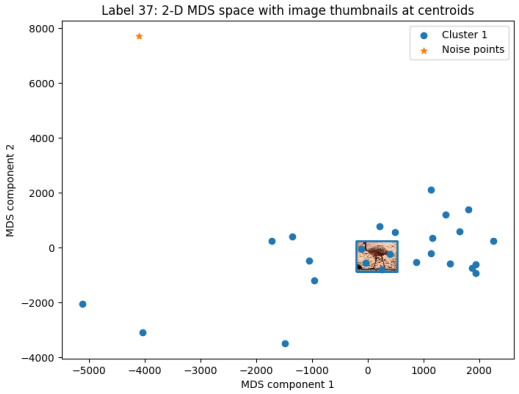

Label 35: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

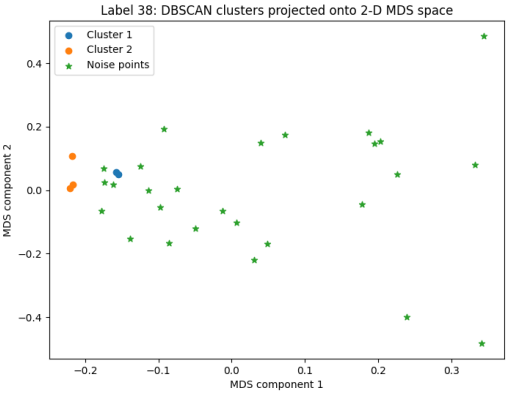

Label 35: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

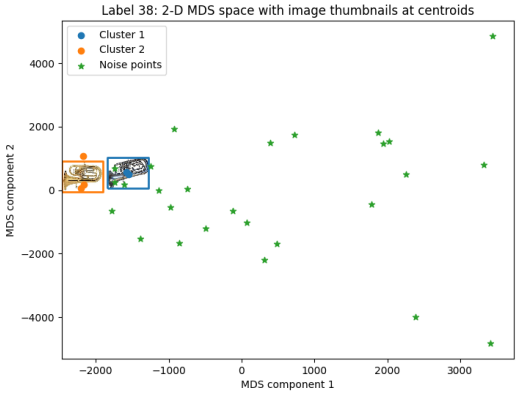

Label 36: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

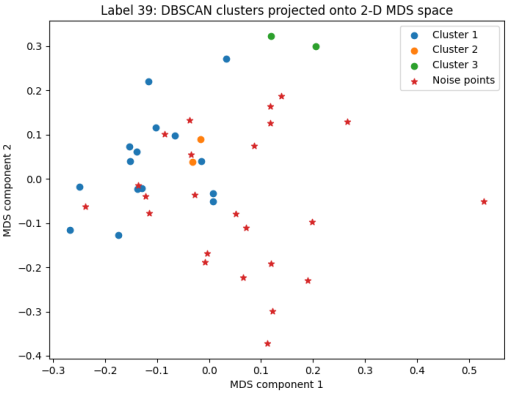

Label 36: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

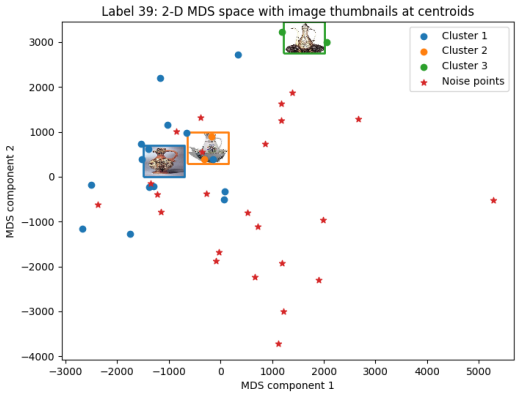

Label 37: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

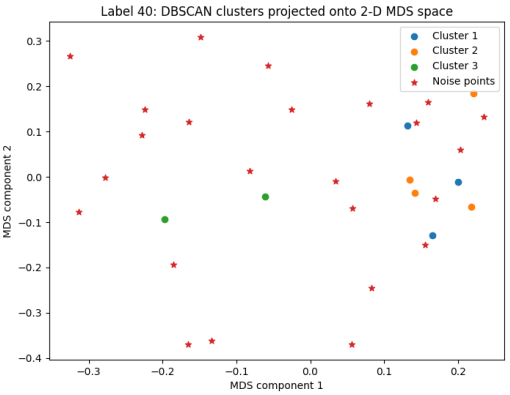

Label 37: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

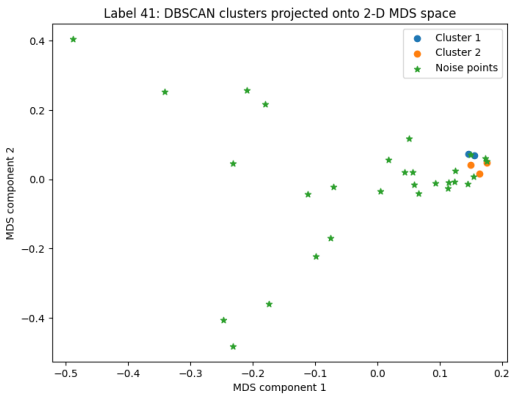

Label 38: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

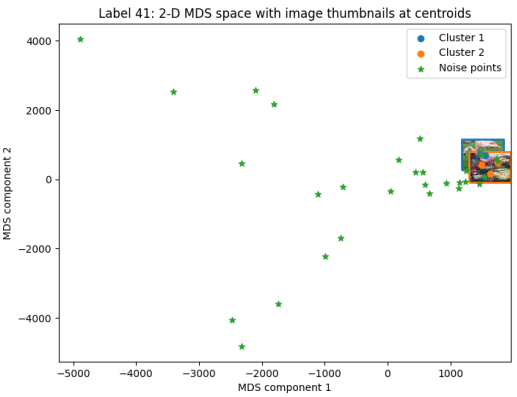

Label 38: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

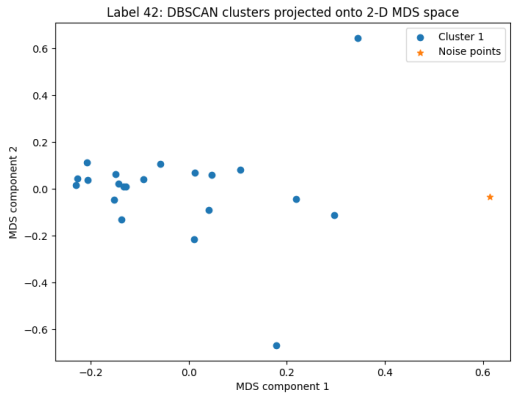

Label 39: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

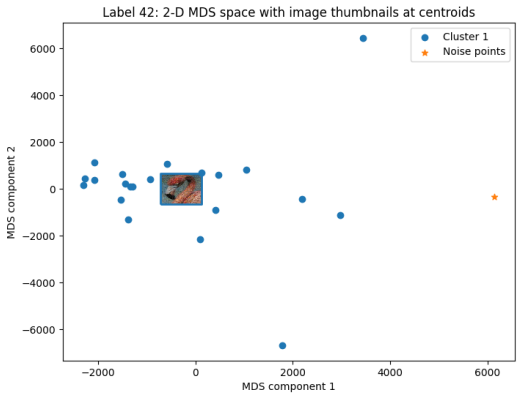

Label 39: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

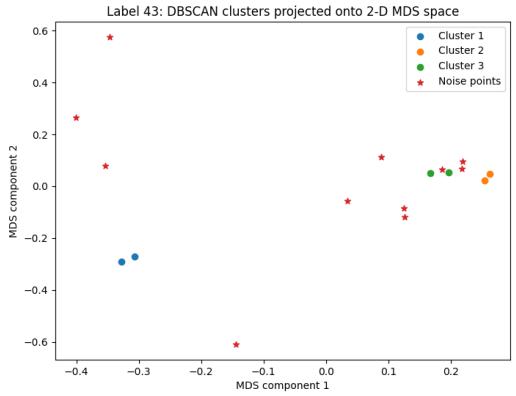

Label 40: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

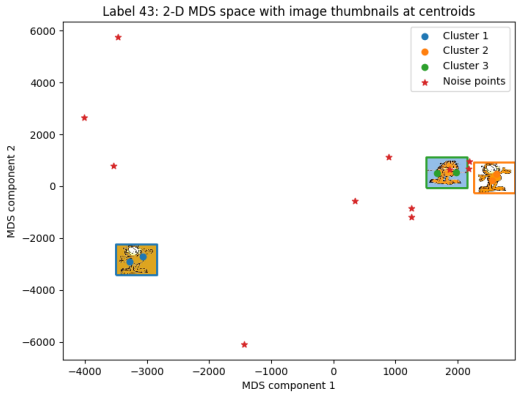

Label 41: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

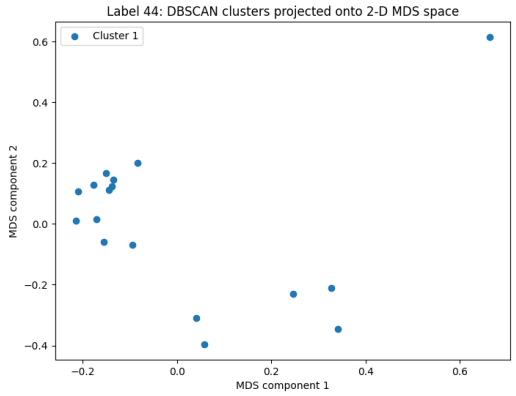

Label 41: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

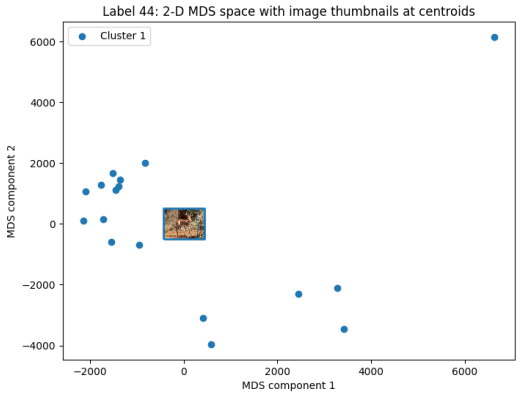

Label 42: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

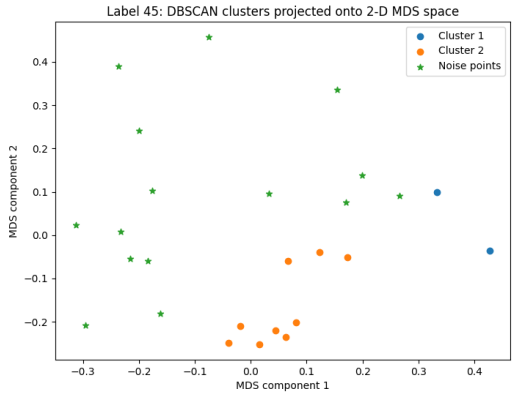

Label 42: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

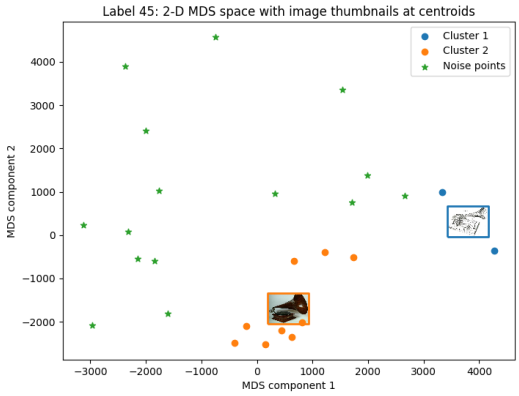

Label 43: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

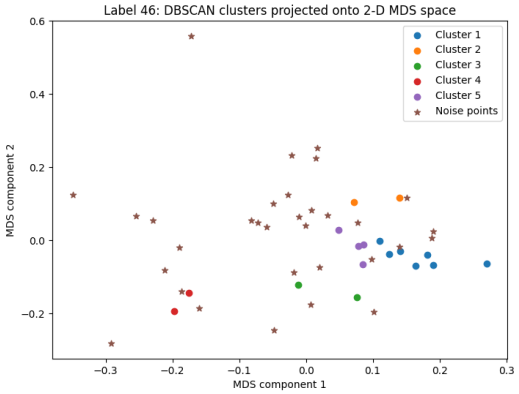

Label 43: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

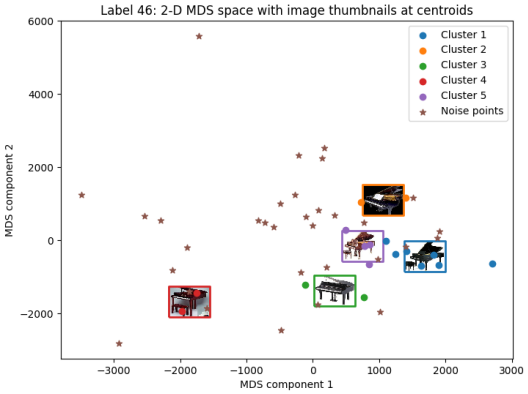

Label 44: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
MDS component 1
MDS component 2

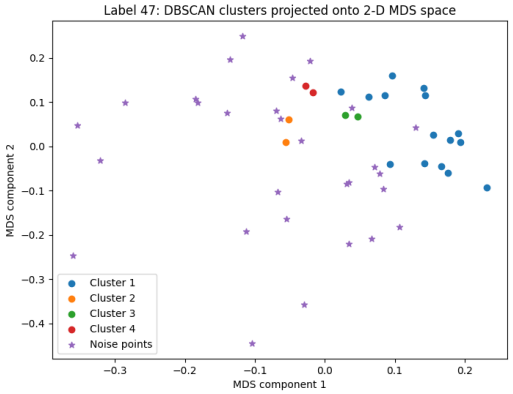

Label 44: 2-D MDS space with image thumbnails at centroids
Cluster 1
MDS component 1
MDS component 2

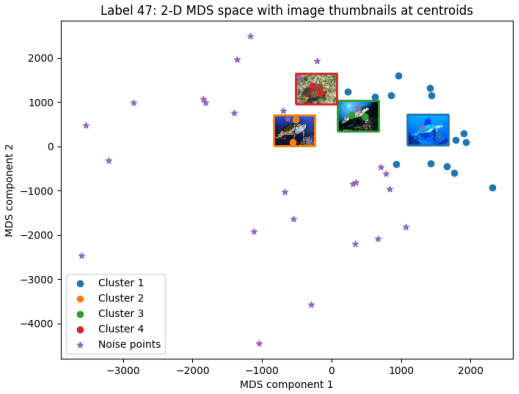

Label 45: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

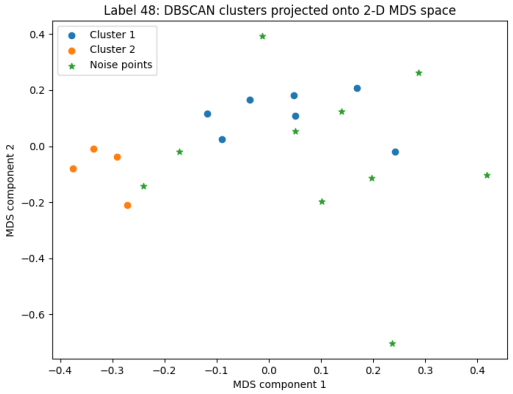

Label 45: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

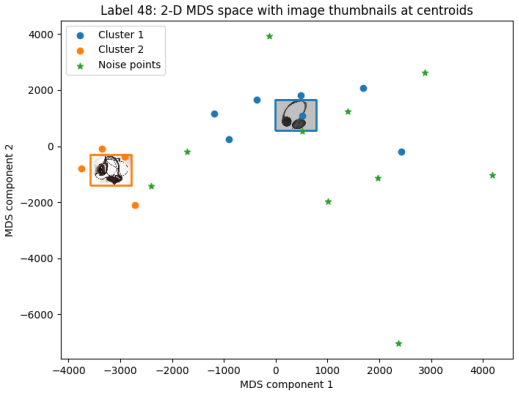

Label 46: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

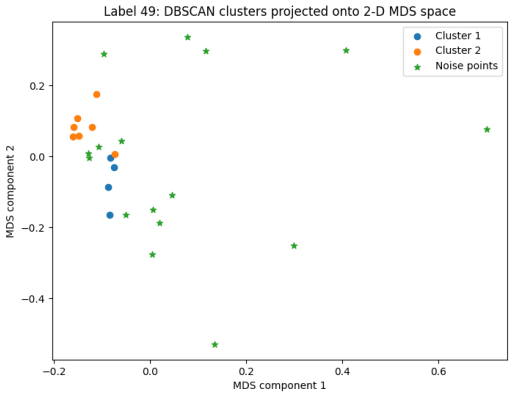

Label 46: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

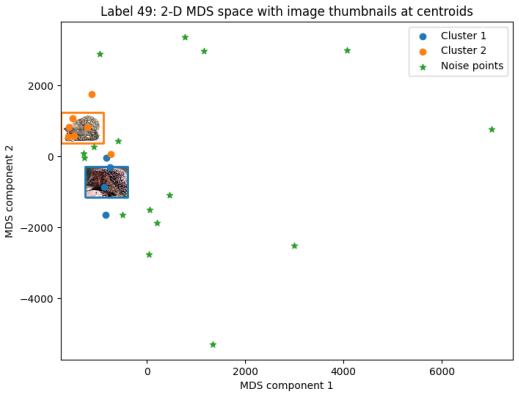

Label 47: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

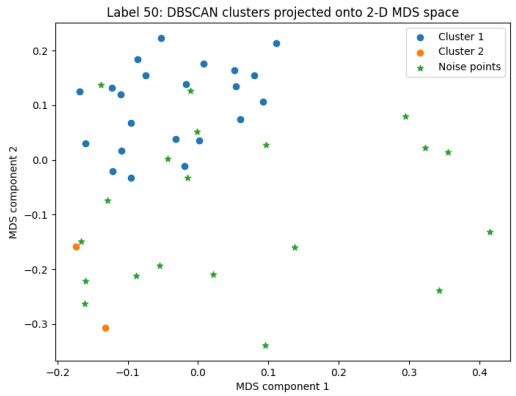

Label 47: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

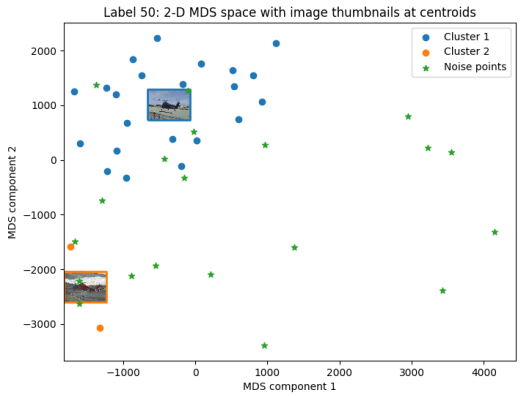

Label 48: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

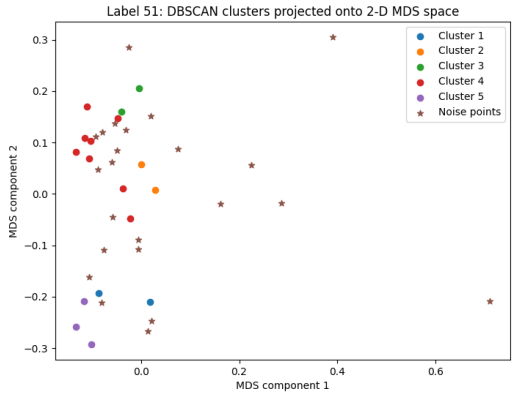

Label 48: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

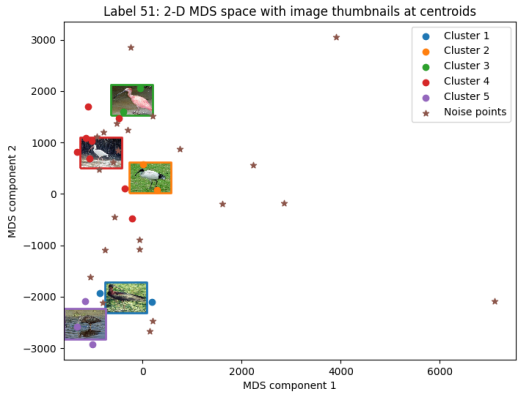

Label 49: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

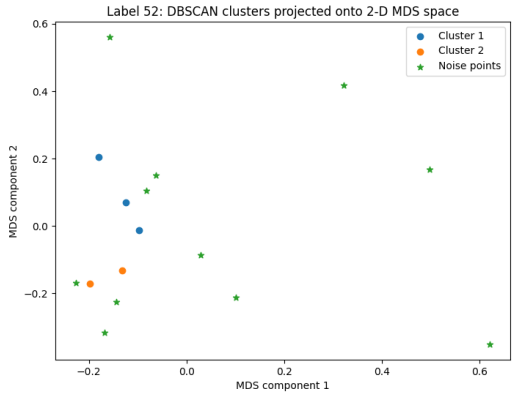

Label 49: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

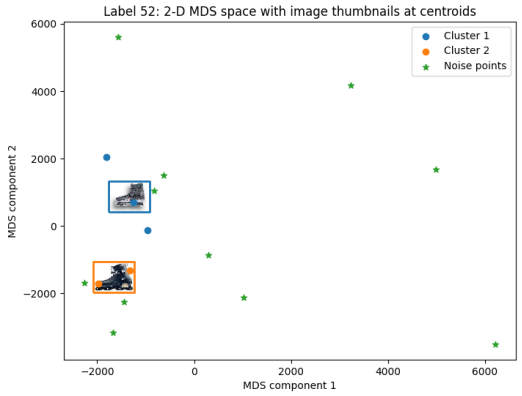

Label 50: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

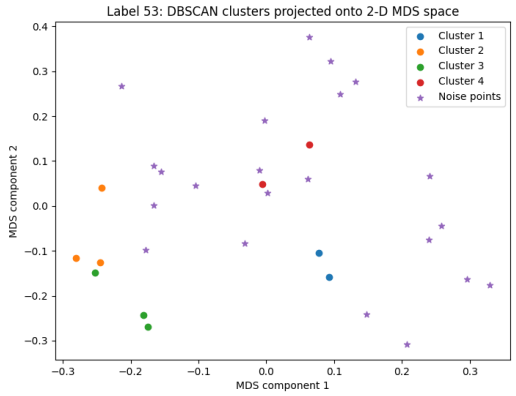

Label 50: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

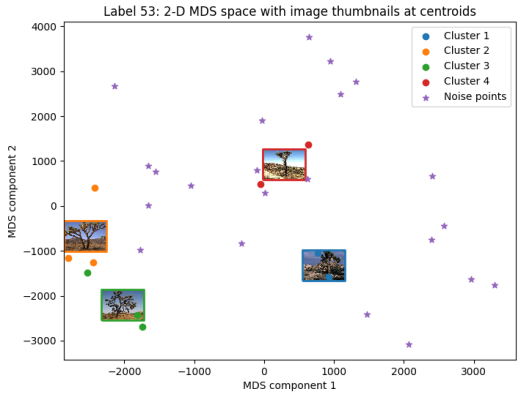

Label 51: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

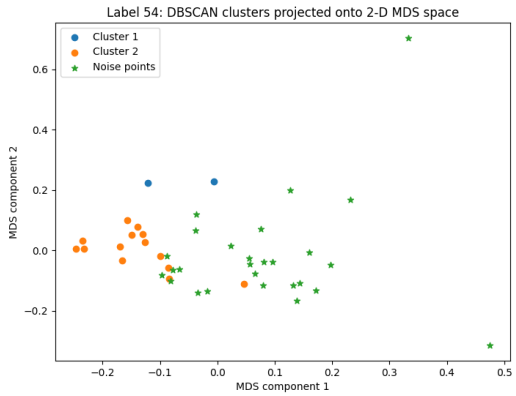

Label 51: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

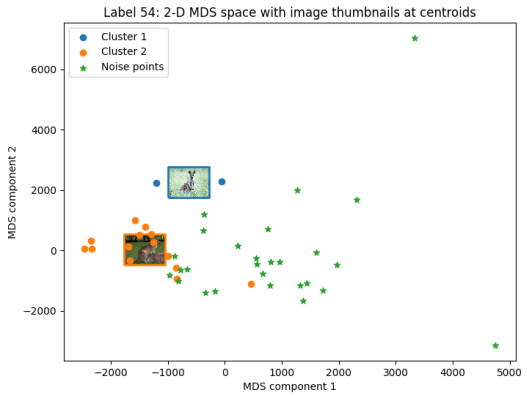

Label 52: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

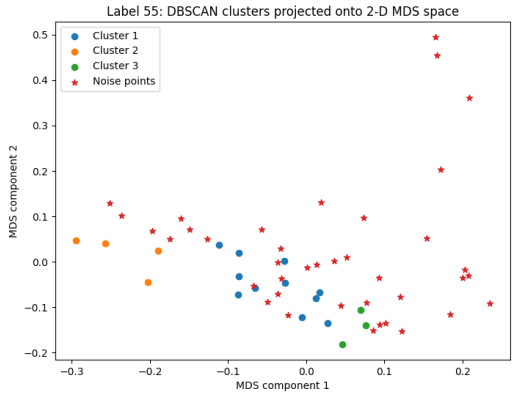

Label 52: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

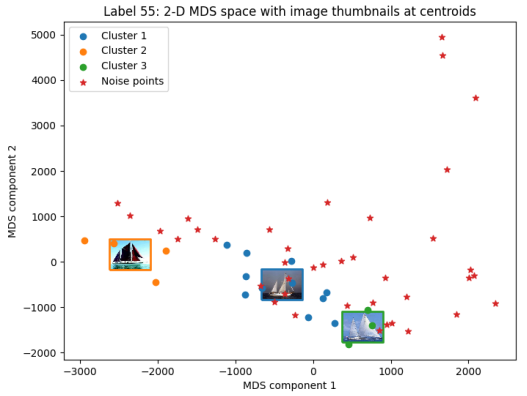

Label 53: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

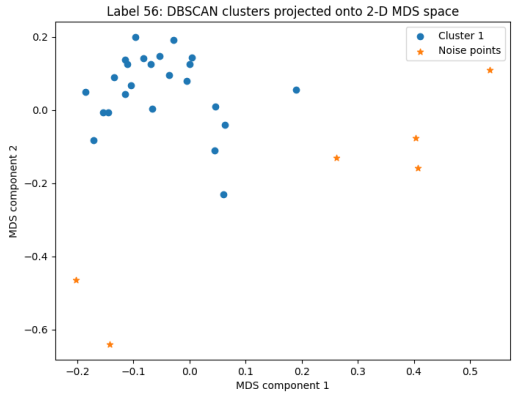

Label 53: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

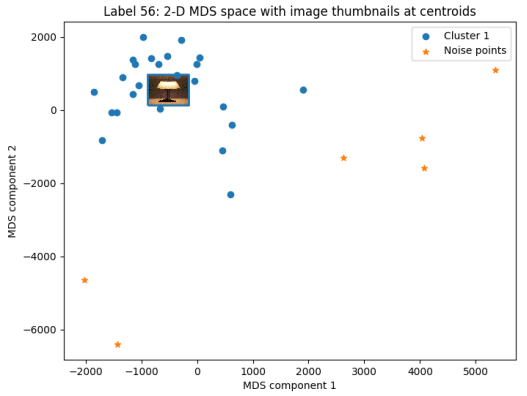

Label 54: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

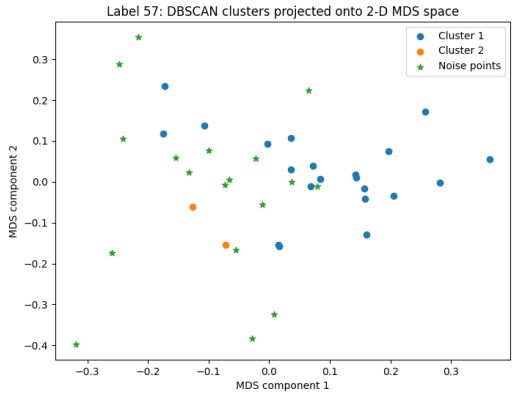

Label 54: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

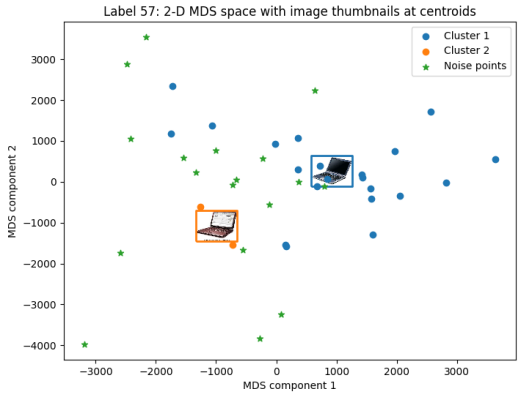

Label 55: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

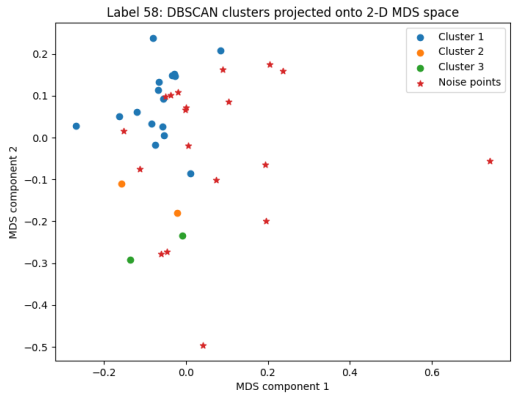

Label 55: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

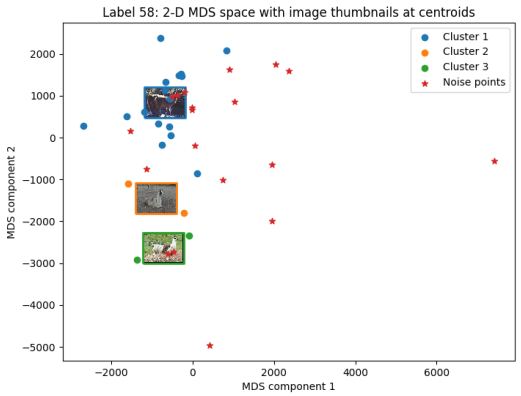

Label 56: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

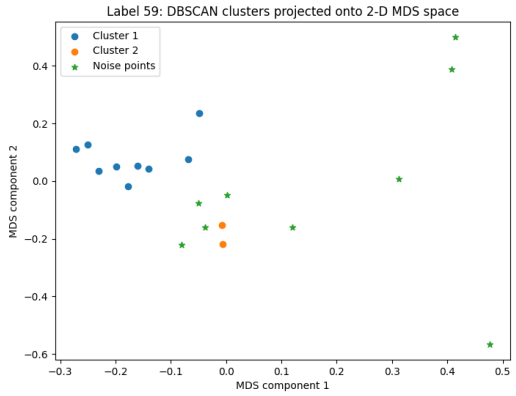

Label 56: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

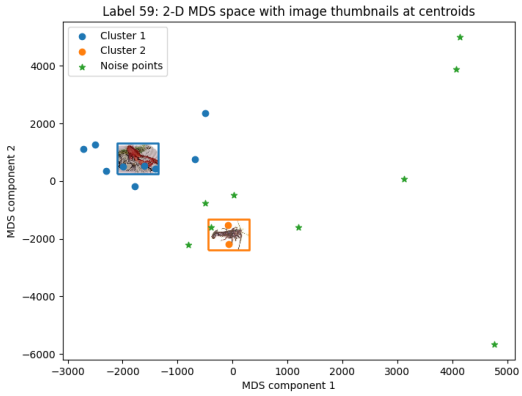

Label 57: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

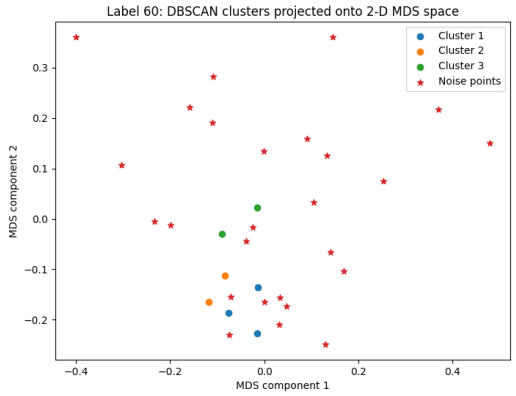

Label 57: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

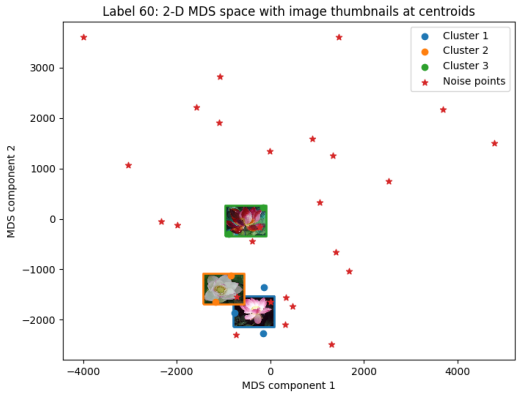

Label 58: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

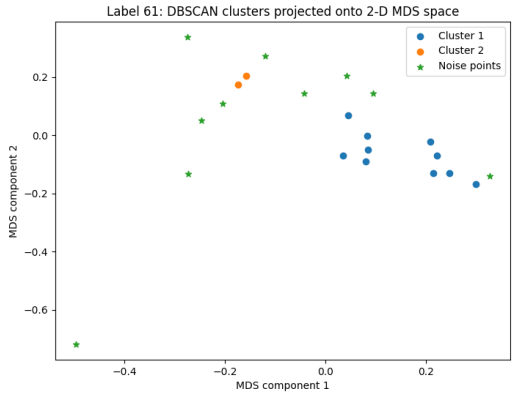

Label 58: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

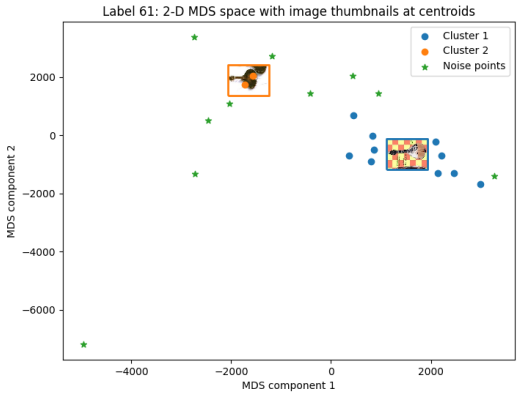

Label 59: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

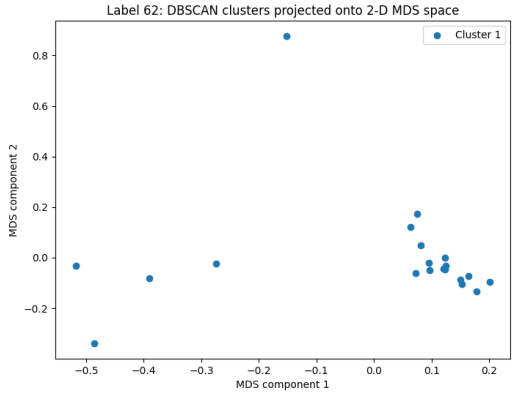

Label 59: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

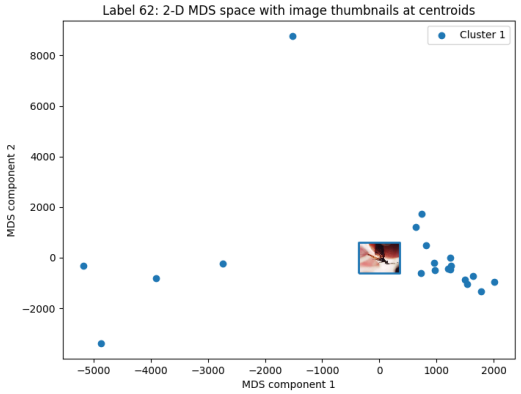

Label 60: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

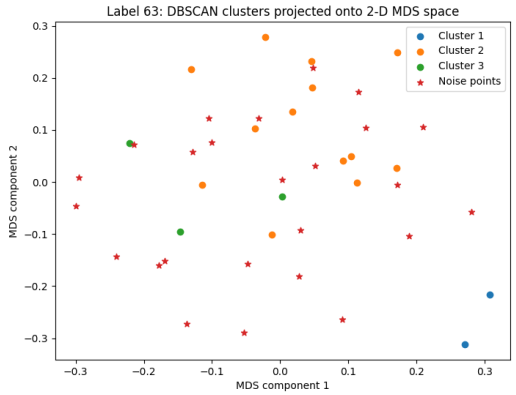

Label 60: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

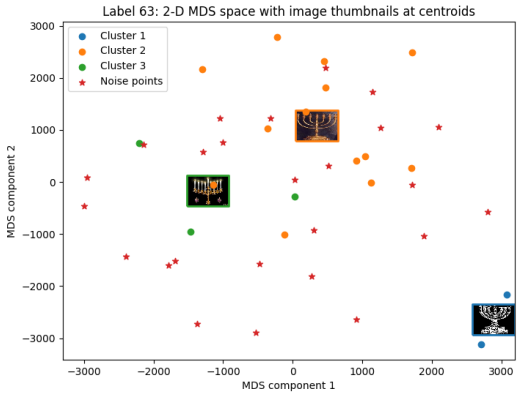

Label 61: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

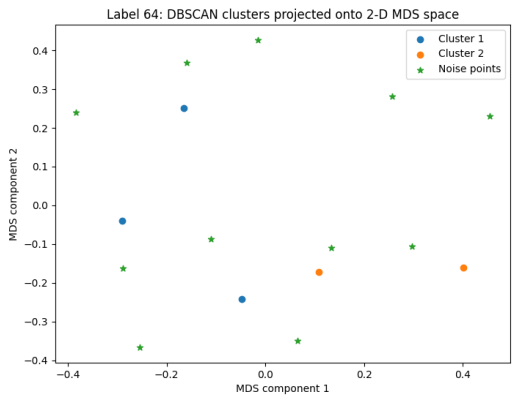

Label 61: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

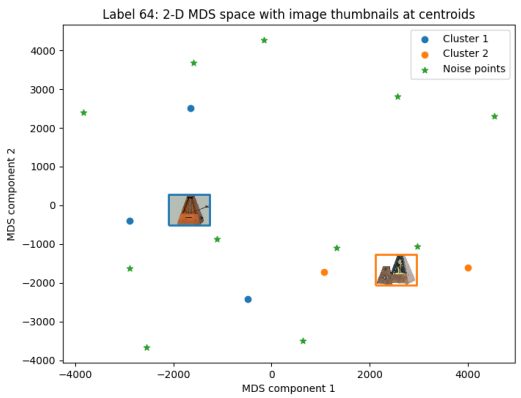

Label 62: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
MDS component 1
MDS component 2

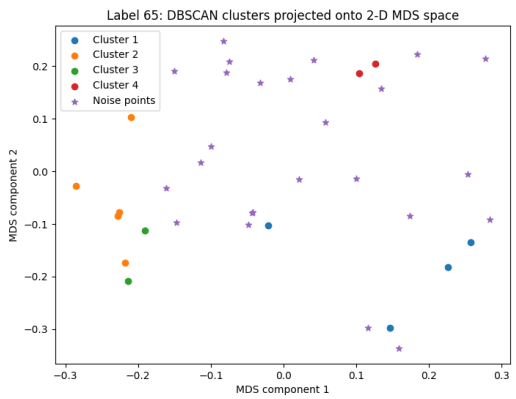

Label 62: 2-D MDS space with image thumbnails at centroids
Cluster 1
MDS component 1
MDS component 2

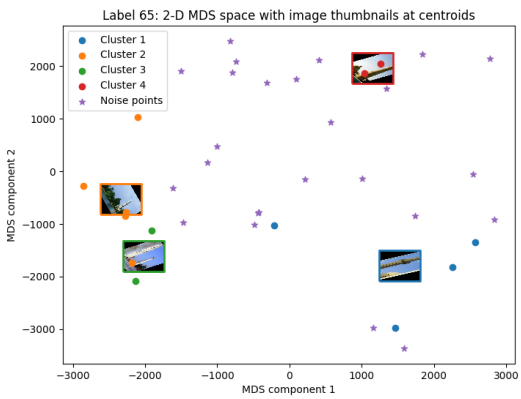

Label 63: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

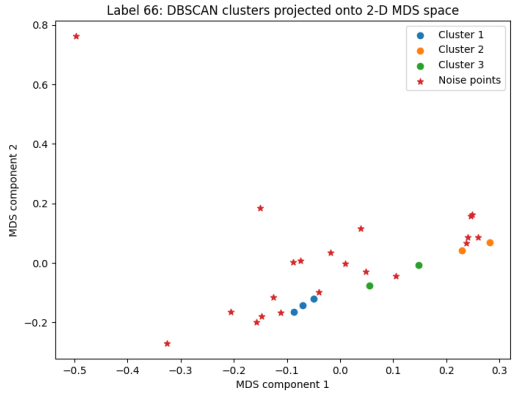

Label 63: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

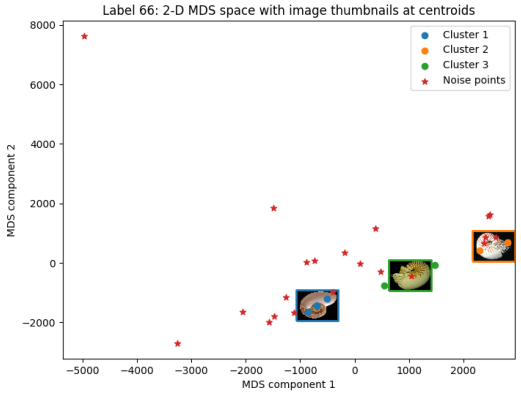

Label 64: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

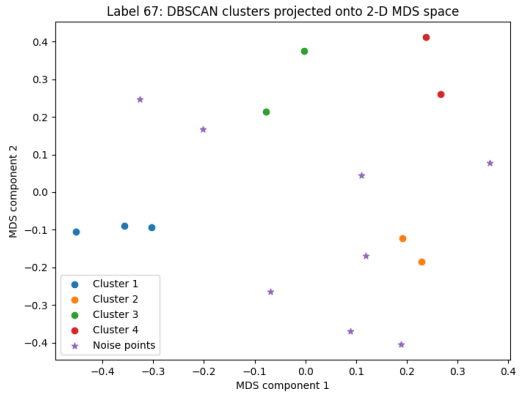

Label 64: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

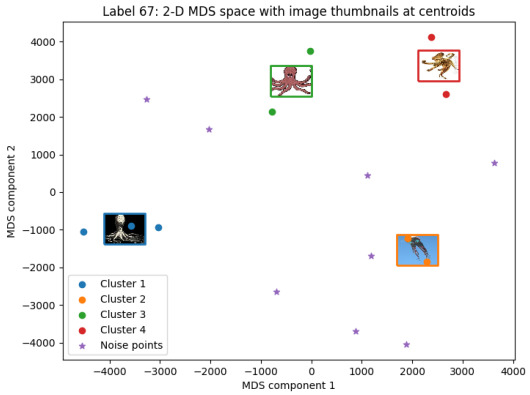

Label 65: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

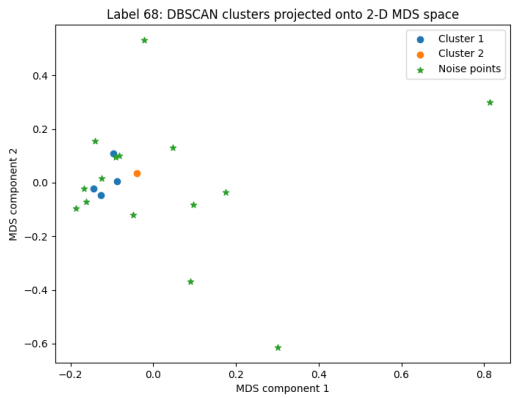

Label 65: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

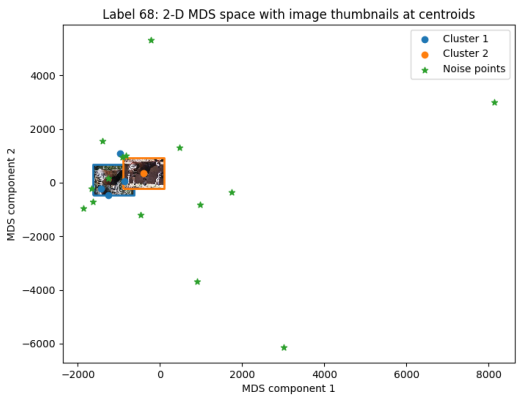

Label 66: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

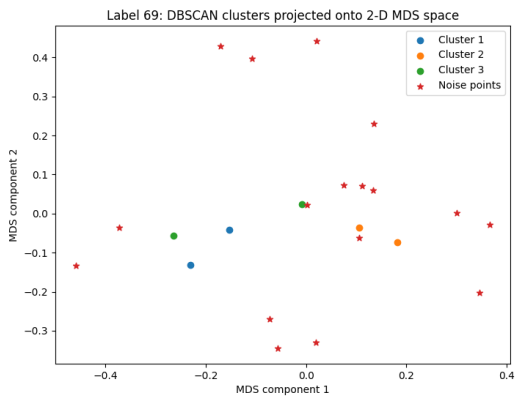

Label 66: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

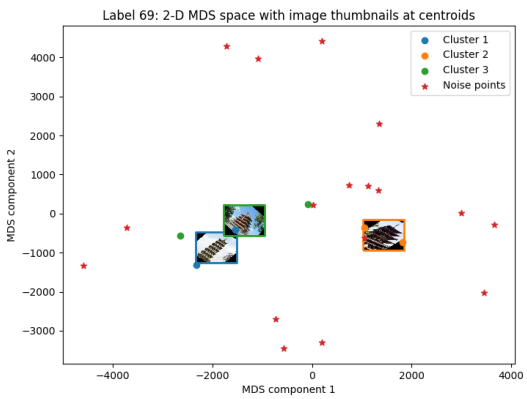

Label 67: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

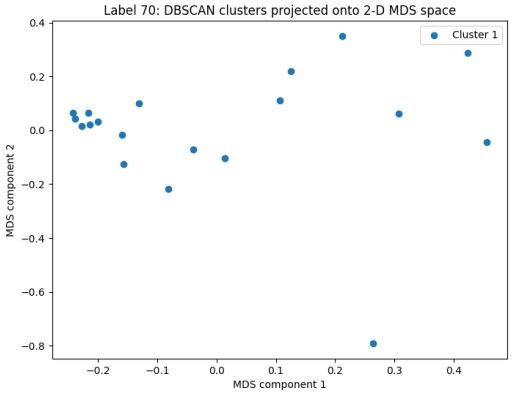

Label 67: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

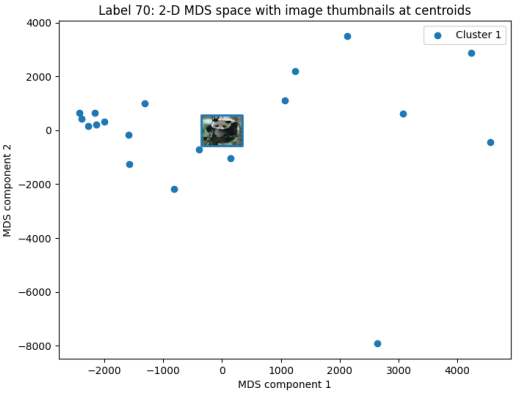

Label 68: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

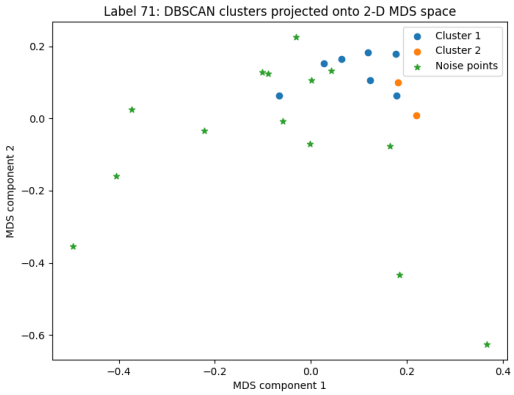

Label 68: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

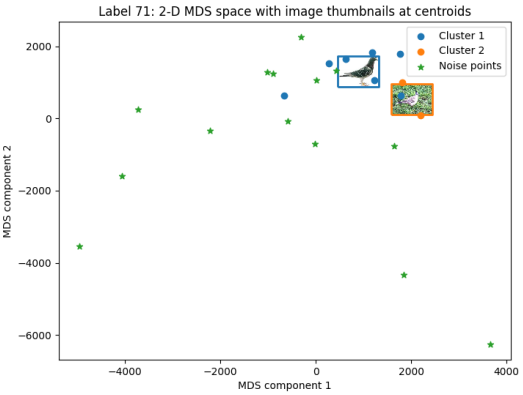

Label 69: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

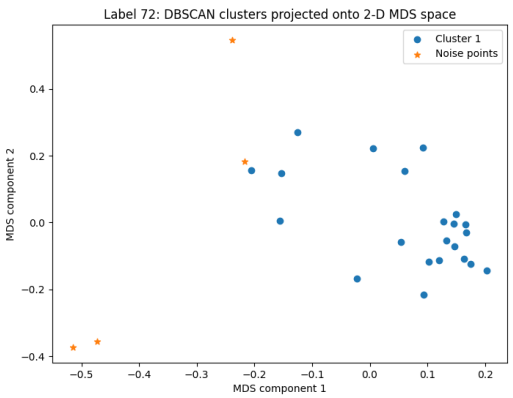

Label 69: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

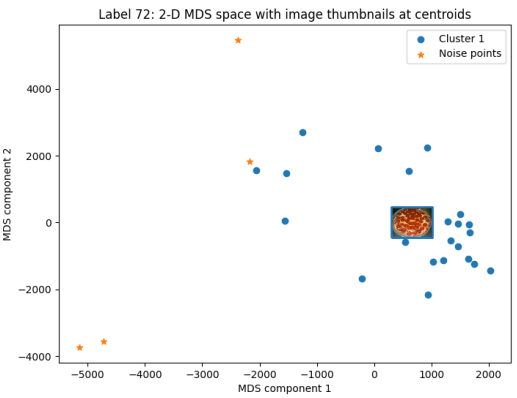

Label 70: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
MDS component 1
MDS component 2

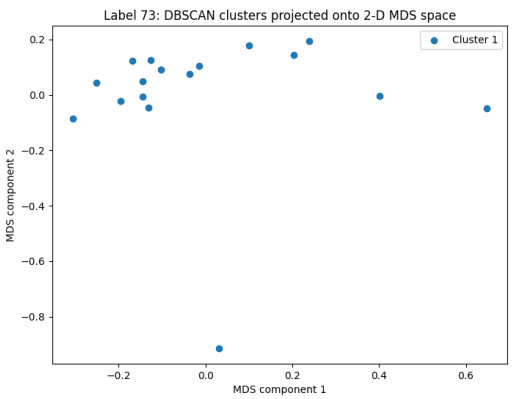

Label 70: 2-D MDS space with image thumbnails at centroids
Cluster 1
MDS component 1
MDS component 2

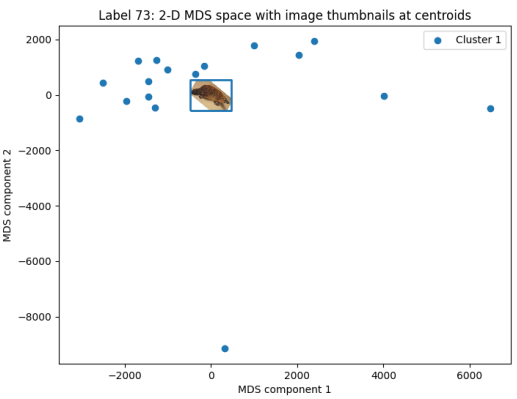

Label 71: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

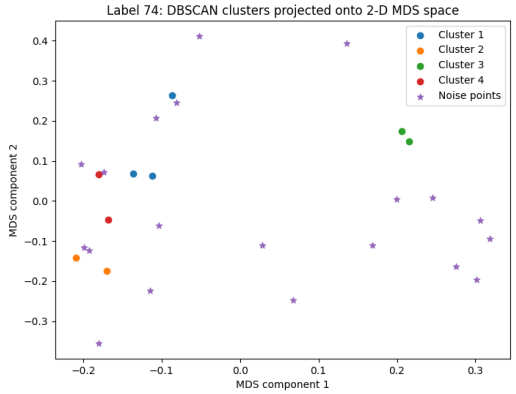

Label 71: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

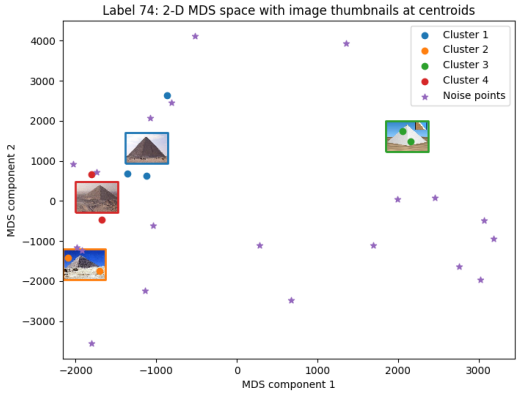

Label 72: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

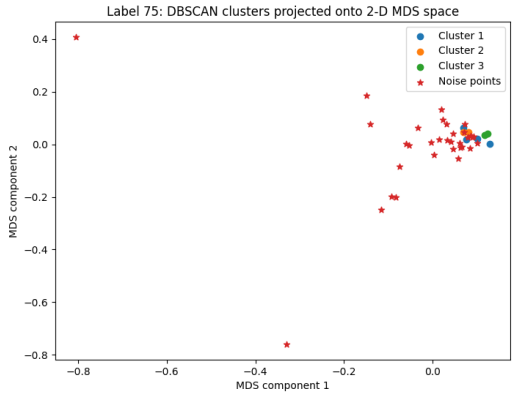

Label 72: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

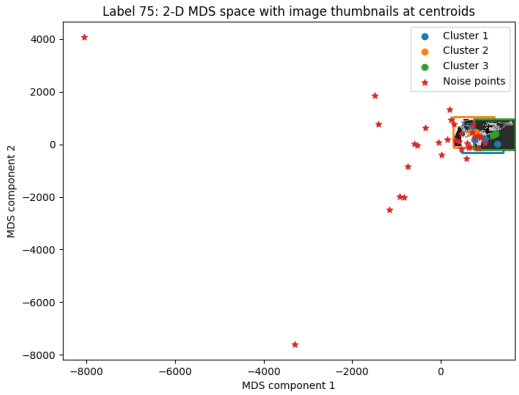

Label 73: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
MDS component 1
MDS component 2

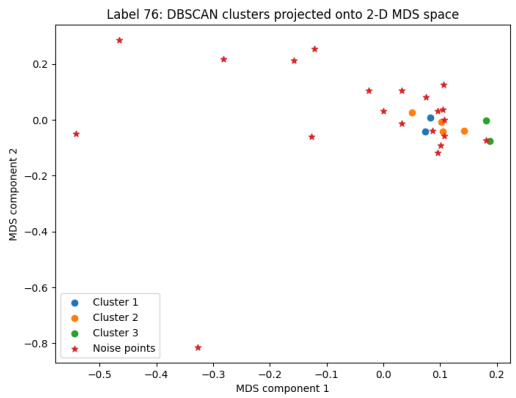

Label 73: 2-D MDS space with image thumbnails at centroids
Cluster 1
MDS component 1
MDS component 2

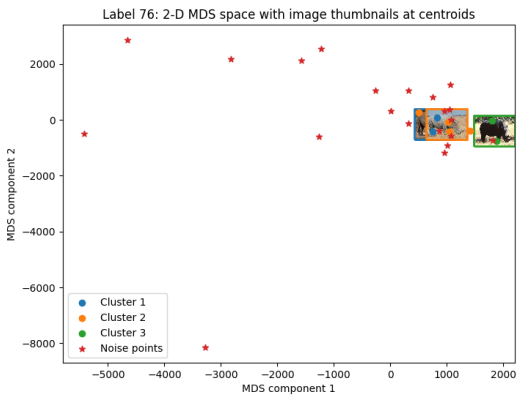

Label 74: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

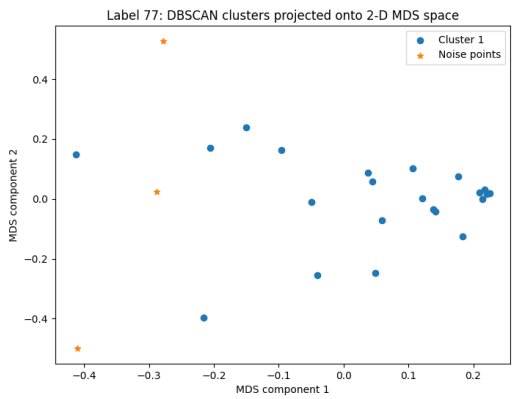

Label 74: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

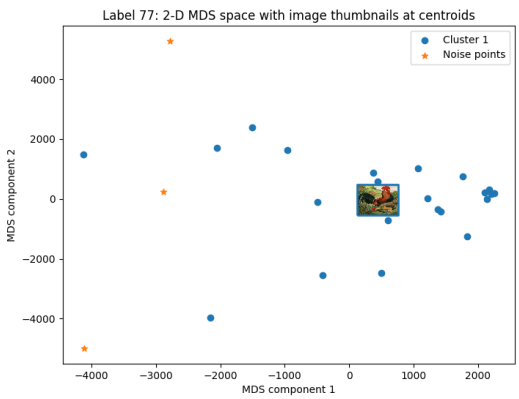

Label 75: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

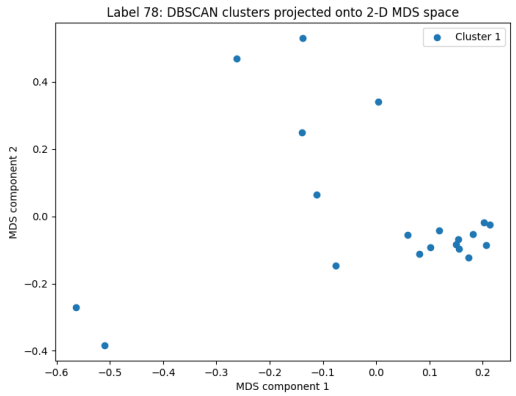

Label 75: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

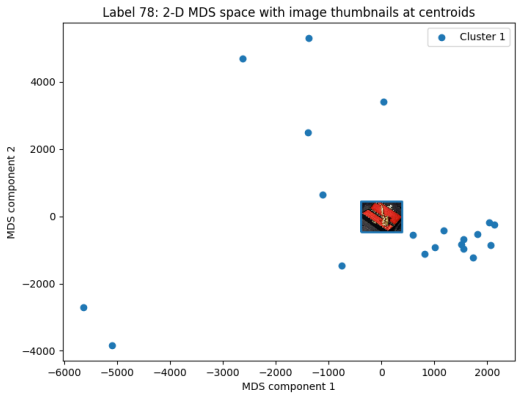

Label 76: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

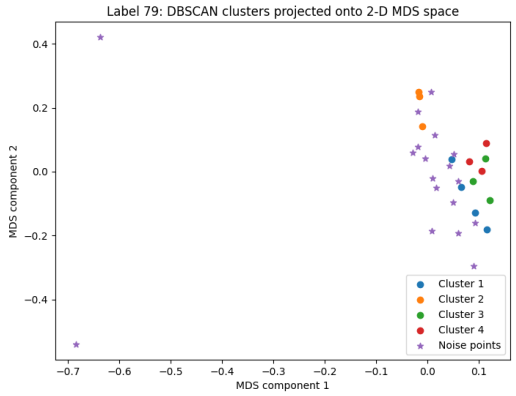

Label 76: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

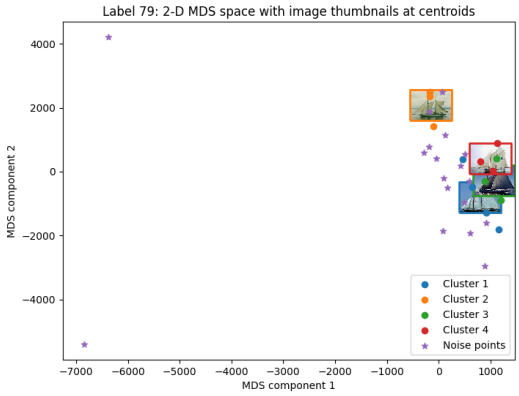

Label 77: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Noise points
MDS component 1
MDS component 2

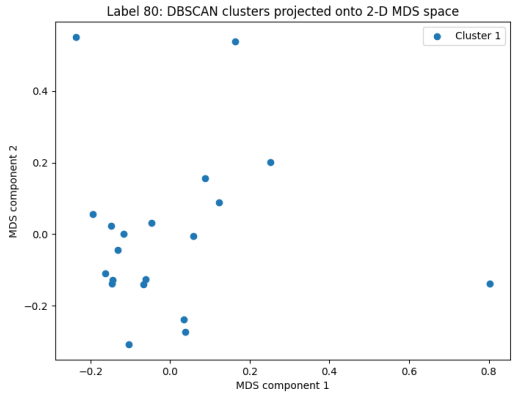

Label 77: 2-D MDS space with image thumbnails at centroids
Cluster 1
Noise points
MDS component 1
MDS component 2

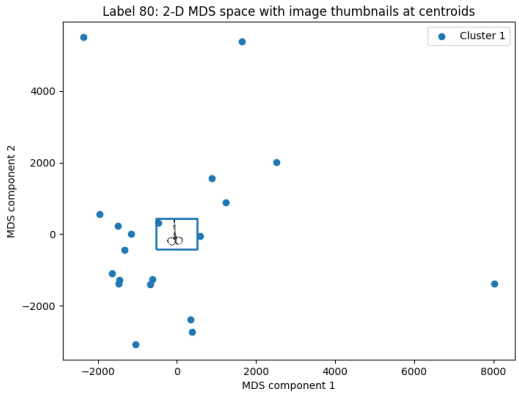

Label 78: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
MDS component 1
MDS component 2

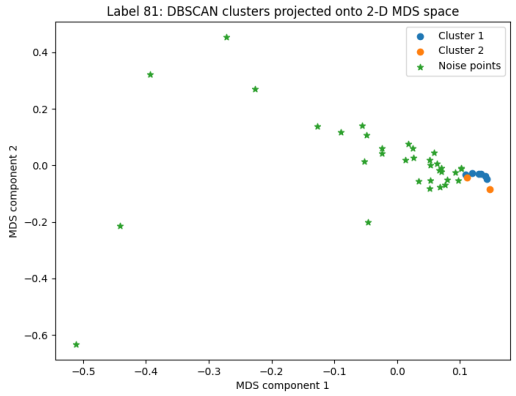

Label 78: 2-D MDS space with image thumbnails at centroids
Cluster 1
MDS component 1
MDS component 2

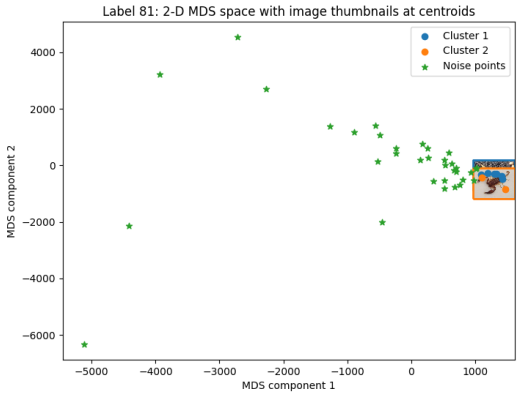

Label 79: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

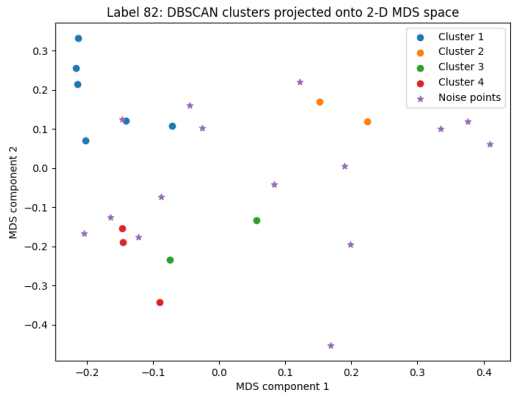

Label 79: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

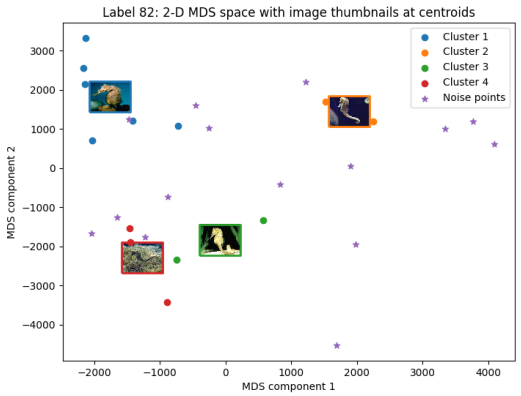

Label 80: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
MDS component 1
MDS component 2

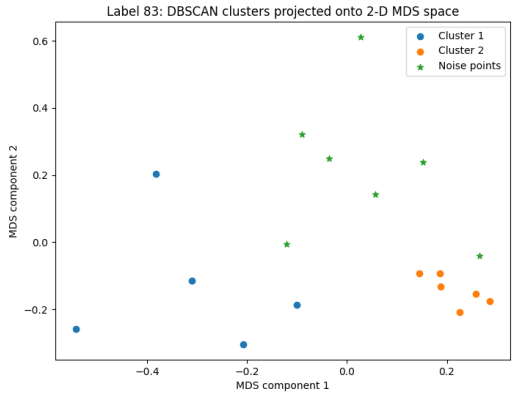

Label 80: 2-D MDS space with image thumbnails at centroids
Cluster 1
MDS component 1
MDS component 2

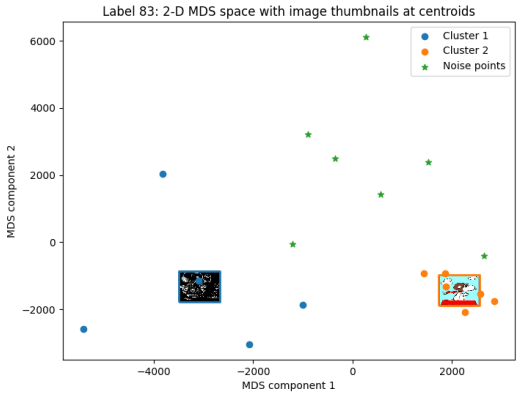

Label 81: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

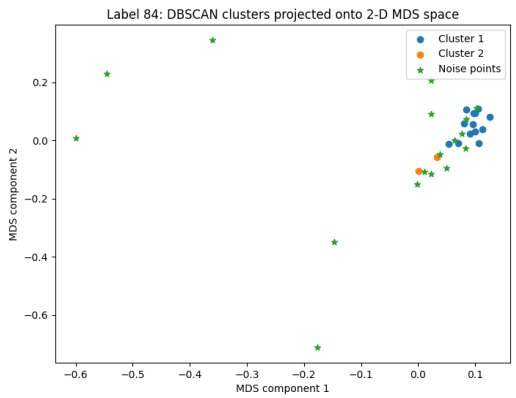

Label 81: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

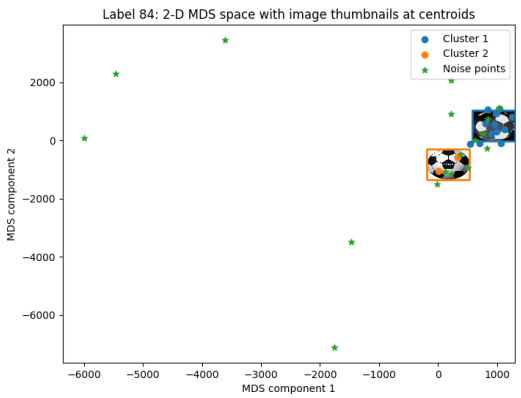

Label 82: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

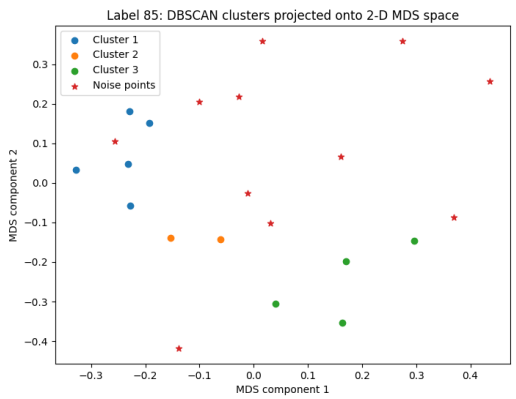

Label 82: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

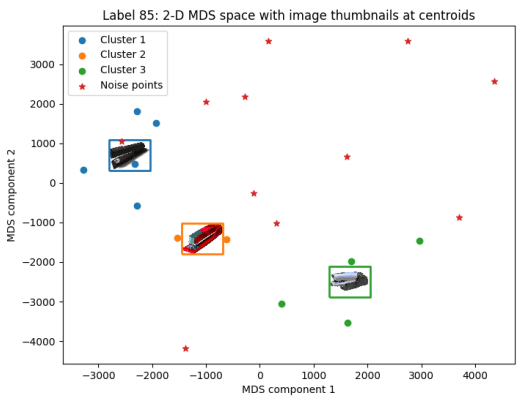

Label 83: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

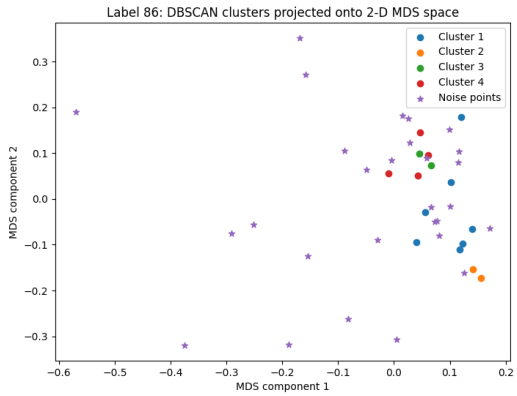

Label 83: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

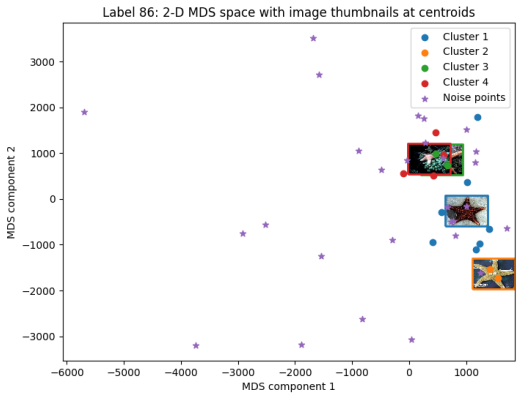

Label 84: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

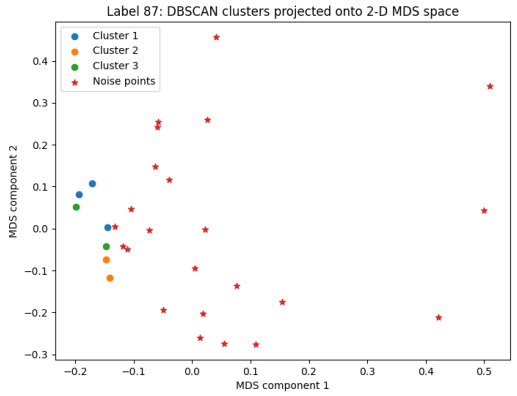

Label 84: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

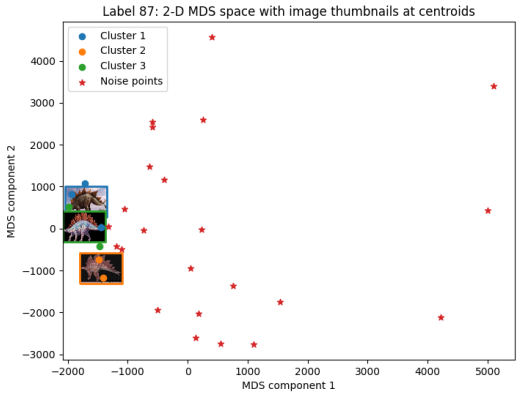

Label 85: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

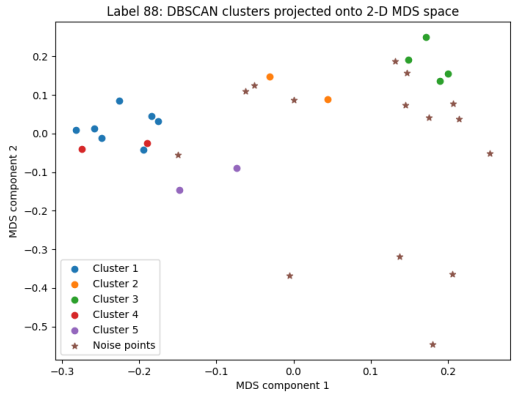

Label 85: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

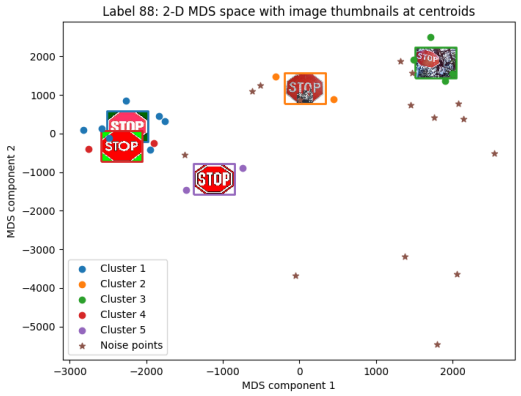

Label 86: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

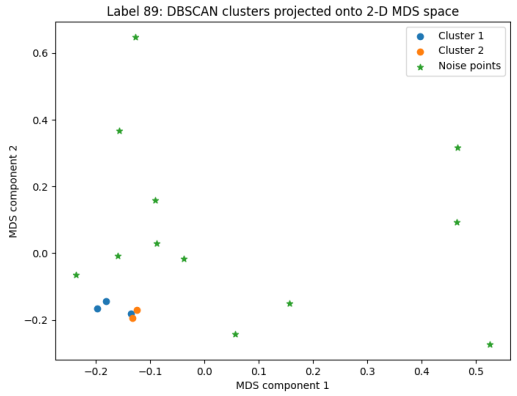

Label 86: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Noise points
MDS component 1
MDS component 2

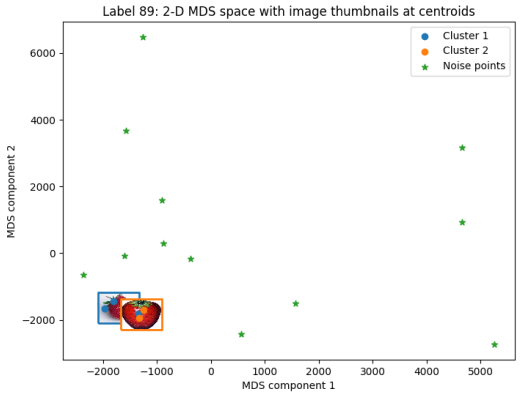

Label 87: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

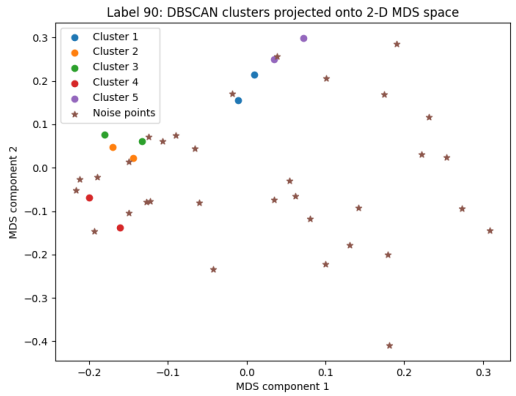

Label 87: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2

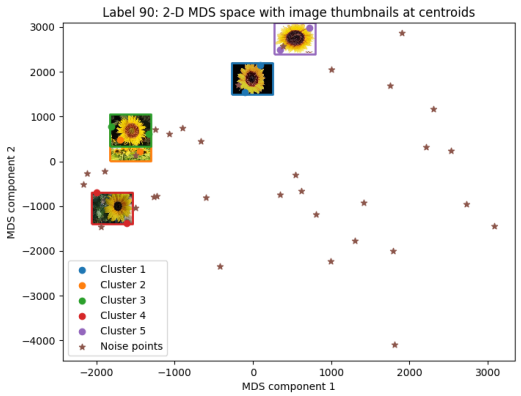

Label 88: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

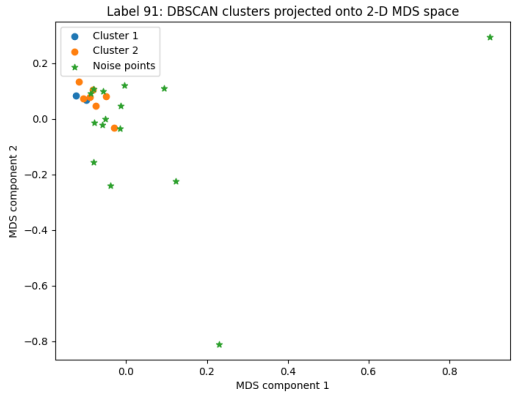

Label 88: 2-D MDS space with image thumbnails at centroids
STOP
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

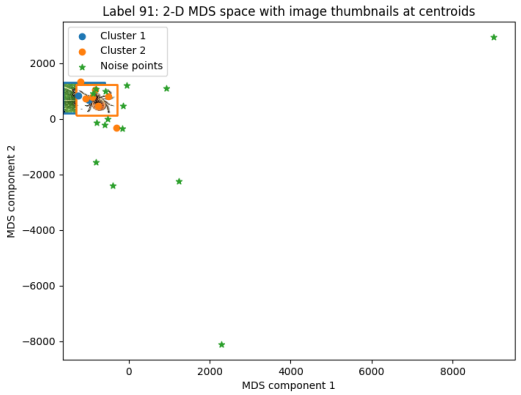

Label 89: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

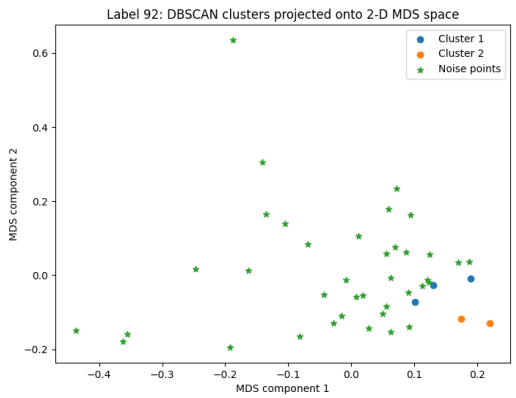

Label 89: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

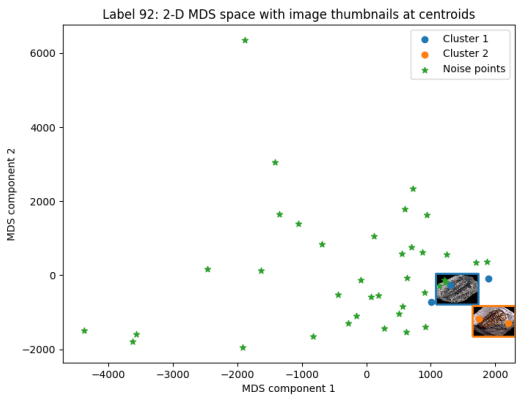

Label 90: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

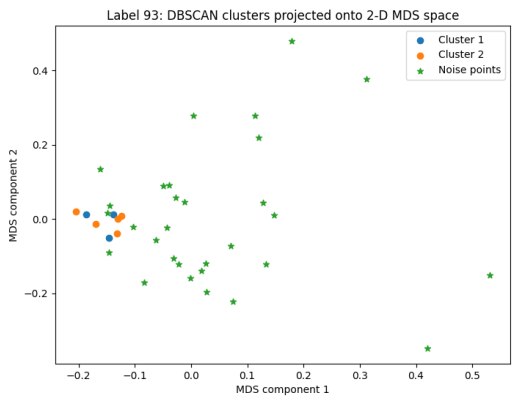

Label 90: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Noise points
MDS component 1
MDS component 2

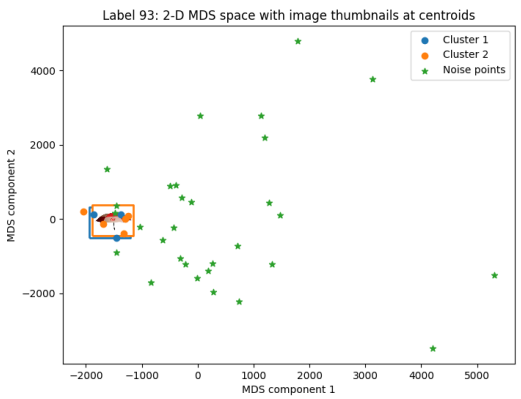

Label 91: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

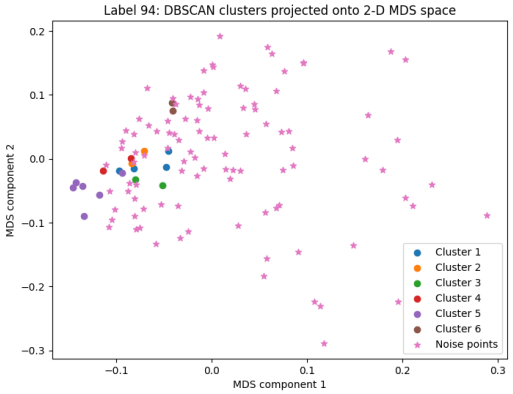

Label 91: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

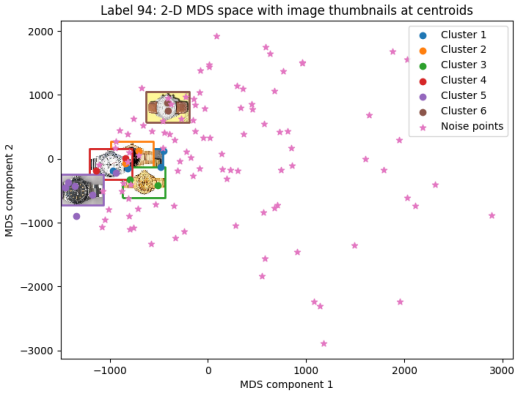

Label 92: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

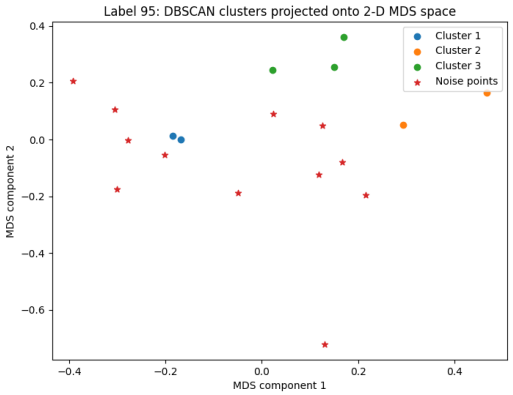

Label 92: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

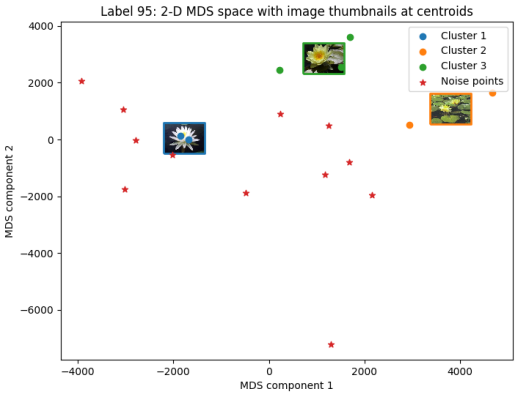

Label 93: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

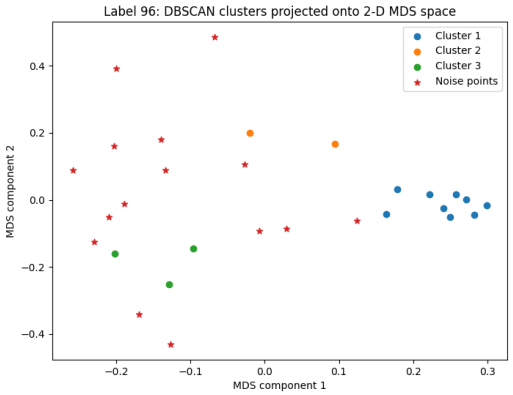

Label 93: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

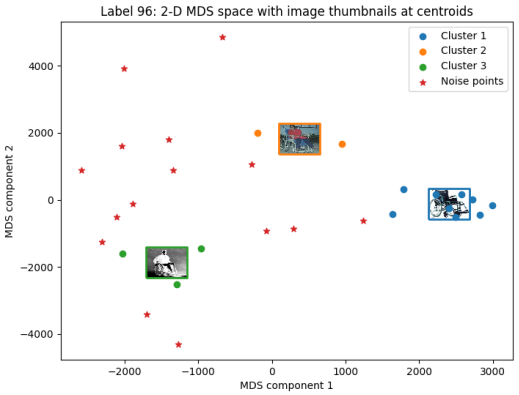

Label 94: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Noise points

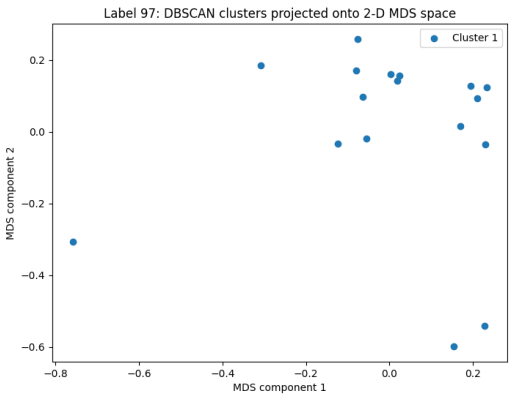

Label 94: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Noise points

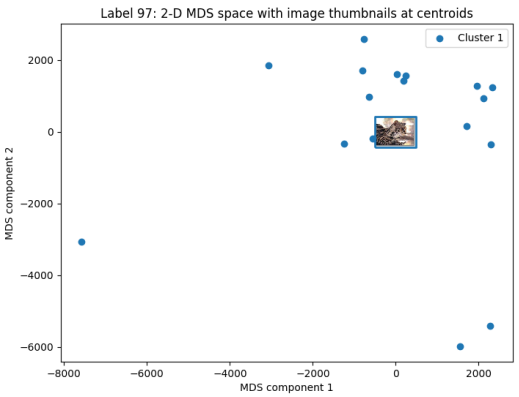

Label 95: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Noise points

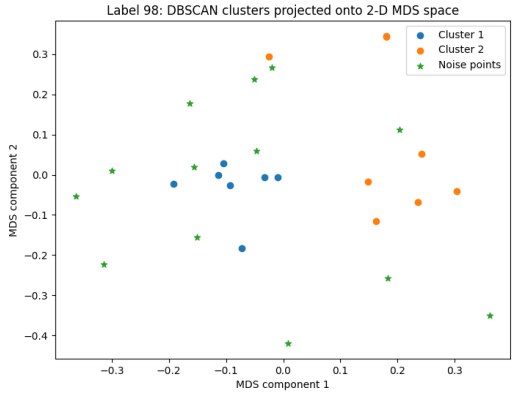

Label 95: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Noise points

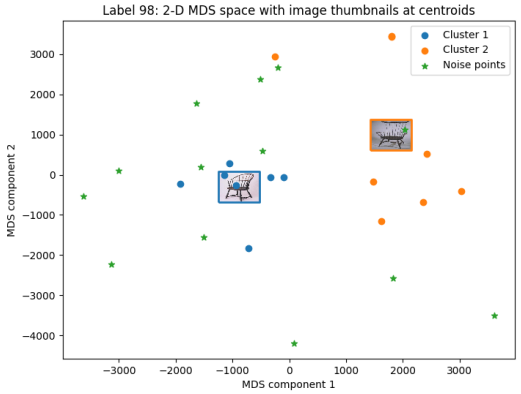

Label 96: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Noise points

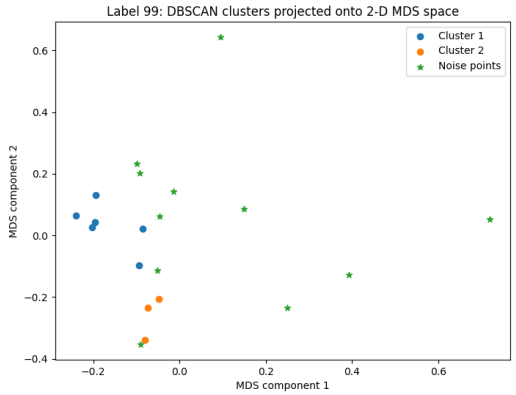

Label 96: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1
Cluster 2
Cluster 3
Noise points

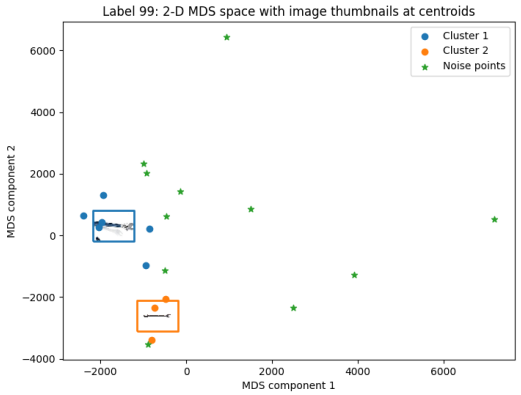

Label 97: DBSCAN clusters projected onto 2-D MDS space
MDS component 1
MDS component 2
Cluster 1

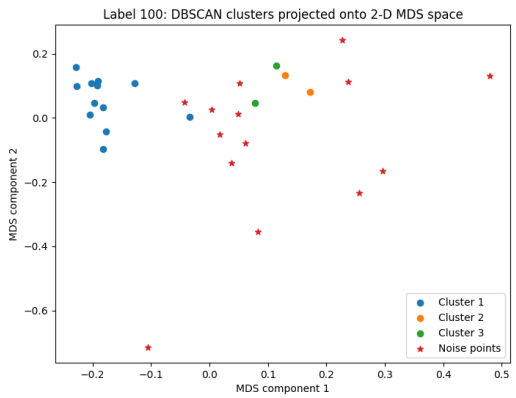

Label 97: 2-D MDS space with image thumbnails at centroids
MDS component 1
MDS component 2
Cluster 1

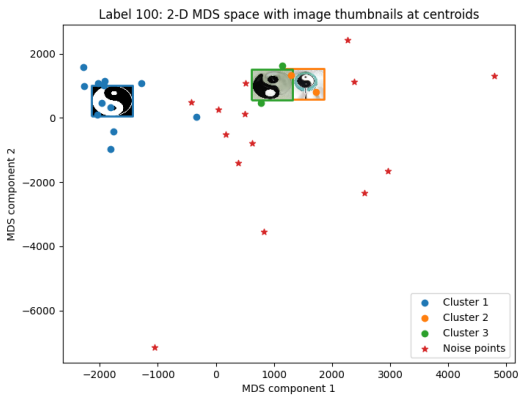

Label 98: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2

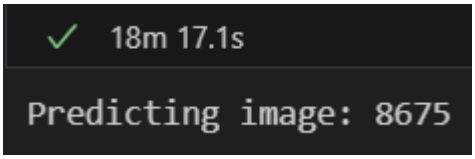

Label 98: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2


Label 99: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2


Label 99: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Noise points
MDS component 1
MDS component 2


Label 100: DBSCAN clusters projected onto 2-D MDS space
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2


Label 100: 2-D MDS space with image thumbnails at centroids
Cluster 1
Cluster 2
Cluster 3
Noise points
MDS component 1
MDS component 2


18m 17.1s
Predicting image: 8675

```
Label 0: Precision=0.59, Recall=0.98, F1-score=0.74
Label 1: Precision=0.96, Recall=0.31, F1-score=0.47
Label 2: Precision=0.90, Recall=1.00, F1-score=0.95
Label 3: Precision=0.95, Recall=1.00, F1-score=0.98
Label 4: Precision=0.82, Recall=1.00, F1-score=0.90
Label 5: Precision=0.90, Recall=0.97, F1-score=0.93
Label 6: Precision=0.90, Recall=0.43, F1-score=0.58
Label 7: Precision=0.88, Recall=0.71, F1-score=0.79
Label 8: Precision=0.88, Recall=0.92, F1-score=0.90
Label 9: Precision=0.72, Recall=0.78, F1-score=0.75
Label 10: Precision=0.73, Recall=0.96, F1-score=0.83
Label 11: Precision=0.80, Recall=1.00, F1-score=0.89
Label 12: Precision=1.00, Recall=0.92, F1-score=0.96
Label 13: Precision=0.88, Recall=0.94, F1-score=0.91
Label 14: Precision=0.81, Recall=0.59, F1-score=0.68
Label 15: Precision=0.97, Recall=0.90, F1-score=0.94
Label 16: Precision=0.95, Recall=0.43, F1-score=0.60
Label 17: Precision=0.71, Recall=1.00, F1-score=0.83
Label 18: Precision=0.94, Recall=0.76, F1-score=0.84
Label 19: Precision=0.93, Recall=1.00, F1-score=0.96
Label 20: Precision=1.00, Recall=0.57, F1-score=0.72
Label 21: Precision=1.00, Recall=0.03, F1-score=0.06
Label 22: Precision=1.00, Recall=0.03, F1-score=0.06
Label 23: Precision=0.96, Recall=0.43, F1-score=0.60
Label 24: Precision=0.54, Recall=0.29, F1-score=0.38
Label 25: Precision=0.65, Recall=0.91, F1-score=0.76
Label 26: Precision=0.96, Recall=0.70, F1-score=0.81
Label 27: Precision=0.71, Recall=0.83, F1-score=0.76
Label 28: Precision=0.75, Recall=0.60, F1-score=0.67
Label 29: Precision=0.78, Recall=0.84, F1-score=0.81
```

```
Label 30: Precision=1.00, Recall=0.28, F1-score=0.43
Label 31: Precision=1.00, Recall=1.00, F1-score=1.00
Label 32: Precision=0.78, Recall=0.96, F1-score=0.86
Label 33: Precision=0.97, Recall=0.85, F1-score=0.90
Label 34: Precision=0.77, Recall=0.88, F1-score=0.82
Label 35: Precision=0.79, Recall=0.84, F1-score=0.82
Label 36: Precision=0.00, Recall=0.00, F1-score=0.00
Label 37: Precision=0.66, Recall=1.00, F1-score=0.79
Label 38: Precision=0.46, Recall=0.19, F1-score=0.27
Label 39: Precision=0.60, Recall=1.00, F1-score=0.75
Label 40: Precision=0.97, Recall=0.94, F1-score=0.96
Label 41: Precision=0.00, Recall=0.00, F1-score=0.00
Label 42: Precision=0.45, Recall=0.96, F1-score=0.61
Label 43: Precision=0.00, Recall=0.00, F1-score=0.00
Label 44: Precision=0.70, Recall=0.94, F1-score=0.80
Label 45: Precision=0.90, Recall=0.76, F1-score=0.83
Label 46: Precision=0.98, Recall=1.00, F1-score=0.99
Label 47: Precision=0.98, Recall=0.98, F1-score=0.98
Label 48: Precision=1.00, Recall=0.05, F1-score=0.09
Label 49: Precision=1.00, Recall=0.48, F1-score=0.65
Label 50: Precision=0.68, Recall=0.77, F1-score=0.72
Label 51: Precision=0.70, Recall=1.00, F1-score=0.82
Label 52: Precision=0.40, Recall=0.13, F1-score=0.20
Label 53: Precision=1.00, Recall=0.78, F1-score=0.88
Label 54: Precision=0.93, Recall=0.98, F1-score=0.95
Label 55: Precision=0.77, Recall=0.81, F1-score=0.79
Label 56: Precision=0.36, Recall=0.97, F1-score=0.53
Label 57: Precision=0.83, Recall=1.00, F1-score=0.91
Label 58: Precision=0.90, Recall=0.95, F1-score=0.92
Label 59: Precision=0.58, Recall=0.67, F1-score=0.62
```

```
Label 60: Precision=0.87, Recall=0.39, F1-score=0.54
Label 61: Precision=0.73, Recall=0.76, F1-score=0.74
Label 62: Precision=0.58, Recall=0.70, F1-score=0.64
Label 63: Precision=0.61, Recall=0.91, F1-score=0.73
Label 64: Precision=1.00, Recall=0.44, F1-score=0.61
Label 65: Precision=0.69, Recall=0.92, F1-score=0.79
Label 66: Precision=0.54, Recall=0.52, F1-score=0.53
Label 67: Precision=0.70, Recall=0.39, F1-score=0.50
Label 68: Precision=0.00, Recall=0.00, F1-score=0.00
Label 69: Precision=1.00, Recall=0.29, F1-score=0.45
Label 70: Precision=1.00, Recall=0.58, F1-score=0.73
Label 71: Precision=0.94, Recall=0.68, F1-score=0.79
Label 72: Precision=0.96, Recall=0.89, F1-score=0.92
Label 73: Precision=1.00, Recall=0.71, F1-score=0.83
Label 74: Precision=1.00, Recall=0.32, F1-score=0.49
Label 75: Precision=0.60, Recall=1.00, F1-score=0.75
Label 76: Precision=0.37, Recall=0.87, F1-score=0.52
Label 77: Precision=0.88, Recall=0.88, F1-score=0.88
Label 78: Precision=0.49, Recall=1.00, F1-score=0.66
Label 79: Precision=0.71, Recall=0.84, F1-score=0.77
Label 80: Precision=0.52, Recall=0.89, F1-score=0.65
Label 81: Precision=0.88, Recall=0.86, F1-score=0.87
Label 82: Precision=0.83, Recall=0.34, F1-score=0.49
Label 83: Precision=0.39, Recall=0.82, F1-score=0.53
Label 84: Precision=1.00, Recall=0.88, F1-score=0.93
Label 85: Precision=0.94, Recall=0.65, F1-score=0.77
Label 86: Precision=0.78, Recall=1.00, F1-score=0.88
Label 87: Precision=0.68, Recall=0.86, F1-score=0.76
Label 88: Precision=1.00, Recall=0.81, F1-score=0.90
Label 89: Precision=1.00, Recall=0.89, F1-score=0.94
```

```
Label 90: Precision=0.69, Recall=0.98, F1-score=0.81
Label 91: Precision=0.74, Recall=1.00, F1-score=0.85
Label 92: Precision=1.00, Recall=0.79, F1-score=0.88
Label 93: Precision=0.82, Recall=0.73, F1-score=0.77
Label 94: Precision=0.97, Recall=0.97, F1-score=0.97
Label 95: Precision=0.64, Recall=0.50, F1-score=0.56
Label 96: Precision=0.81, Recall=0.70, F1-score=0.75
Label 97: Precision=0.90, Recall=0.53, F1-score=0.67
Label 98: Precision=0.47, Recall=1.00, F1-score=0.64
Label 99: Precision=0.53, Recall=0.47, F1-score=0.50
Label 100: Precision=0.82, Recall=0.77, F1-score=0.79
Overall Accuracy: 0.79
```

Interpretation of Output:

For higher c values, we observe that the number of labels that do not form c clusters has increased and is now almost all the labels, because the number of dense clusters is insufficient.

Again, we observe that labels with more images usually give better results on both precision and recall, likely because smaller labels don't have enough inherent clusters to properly classify inputs. Note that due to the class imbalance of the dataset, such overarching scores may not always be the best indicators to use for evaluation.

Task 3

Specification

User specifies the classifier to be used for classifying the odd numbered images in the caltech dataset. If the user selects the m-nn classifier, the program asks the value of m from the user. Program will output the most similar label for each of the odd numbered images based on the classifiers built using all even number images.

Description

1. Firstly, the user is prompted to enter the classifier to be used for classification tasks.
2. If m-nn is selected, user is prompted to enter the value of m
3. The feature space selected for this task is the FC-layer feature space. The feature descriptors for all the images in Caltech101 dataset are fetched.
4. The even and odd numbered features are separated based on the indices
5. The classification task is done based on the user input
    a. For m-nn
        i. For m-nn, during each iteration, euclidean distances between odd numbered image features and each even numbered image feature is found.
        ii. A min-heap is used to maintain the list and image ids of the closest images.
        iii. After getting the list of m closest neighbors to the odd numbered image feature, a label with the maximum occurrences in these m images is found.
        iv. The odd numbered image in question is most similar to this label.
    b. For decision trees
        i. Information gain is calculated for each feature, and the algorithm selects the feature which maximizes the information gain. The dataset is split based on this feature.
        ii. Recursively, the best feature for left and right subtrees is found.
        iii. At the end of the process, we have the decision tree.

Design

Fully connected layer output:
We have selected the Fully Connected Layer of Resnet as our feature model as it provides features formed by deep neural networks and extract the best characteristics within an image.

We also tried with avgpool and color moments to verify our understanding but as expected, accuracy was much lower for them.

PCA for decision-trees:
Like stated earlier, we have selected the output of the fully connected layer to be the feature space for this task. FC layer output is a 1000 dimensional array, and thus poses a big bottleneck for tasks such as decision tree formation. Citing the lack of computing resources, we decided to apply the principal component analysis to the even numbered image dataset, and reduce its dimensionality to around 150 features.

Information gain for building decision trees:
Information gain method aims to minimize the disorder in the resulting subsets after a split. Information gain captures the information about the class probabilities in its computations. Furthermore, Gini impurity may result in a biased split towards features with more classes.

Outputs

Input: classifier: m-nn, m = 1

```
Image ID: 1 is similar to label 1
Image ID: 3 is similar to label 0
Image ID: 5 is similar to label 1
Image ID: 7 is similar to label 1
Image ID: 9 is similar to label 1
Image ID: 11 is similar to label 1
Image ID: 13 is similar to label 0
Image ID: 15 is similar to label 0
Image ID: 17 is similar to label 1
Image ID: 19 is similar to label 1
Image ID: 21 is similar to label 0
Image ID: 23 is similar to label 1
Image ID: 25 is similar to label 1
Image ID: 27 is similar to label 0
Image ID: 29 is similar to label 0
Image ID: 31 is similar to label 0
Image ID: 33 is similar to label 1
Image ID: 35 is similar to label 0
Image ID: 37 is similar to label 1
Image ID: 39 is similar to label 0
Image ID: 41 is similar to label 1
Image ID: 43 is similar to label 1
Image ID: 45 is similar to label 1
Image ID: 47 is similar to label 0
Image ID: 49 is similar to label 1
...
```

```
Class 0: Precision=0.5724137931034483, Recall=0.3824884792626728, F1-score=0.45856353591160226
Class 1: Precision=0.5379310344827586, Recall=0.7155963302752294, F1-score=0.6141732283464567
Class 2: Precision=0.9900990099009901, Recall=1.0, F1-score=0.9950248756218906
Class 3: Precision=0.9876237623762376, Recall=1.0, F1-score=0.9937733499377335
Class 4: Precision=0.9285714285714286, Recall=0.9629629629629629, F1-score=0.9454545454545454
Class 5: Precision=0.9731051344743277, Recall=0.995, F1-score=0.9839307787391842
Class 6: Precision=0.8888888888888888, Recall=0.38095238095238093, F1-score=0.5333333333333333
Class 7: Precision=0.75, Recall=0.8571428571428571, F1-score=0.7999999999999999
Class 8: Precision=0.9565217391304348, Recall=0.9166666666666666, F1-score=0.9361702127659574
Class 9: Precision=0.8461538461538461, Recall=0.8148148148148148, F1-score=0.830188679245283
Class 10: Precision=0.8260869565217391, Recall=0.8260869565217391, F1-score=0.8260869565217391
Class 11: Precision=0.8333333333333334, Recall=0.9375, F1-score=0.8823529411764706
Class 12: Precision=0.9344262295081968, Recall=0.890625, F1-score=0.9120000000000001
Class 13: Precision=0.9361702127659575, Recall=0.8979591836734694, F1-score=0.9166666666666666
Class 14: Precision=0.625, Recall=0.45454545454545453, F1-score=0.5263157894736842
Class 15: Precision=0.9761904761904762, Recall=0.9761904761904762, F1-score=0.9761904761904762
Class 16: Precision=0.8055555555555556, Recall=0.6304347826086957, F1-score=0.7073170731707318
Class 17: Precision=0.8888888888888888, Recall=0.96, F1-score=0.923076923076923
Class 18: Precision=0.8947368421052632, Recall=0.8095238095238095, F1-score=0.8500000000000001
Class 19: Precision=0.9841269841269841, Recall=1.0, F1-score=0.9919999999999999
Class 20: Precision=0.95, Recall=0.8260869565217391, F1-score=0.8837209302325583
Class 21: Precision=0.9666666666666667, Recall=0.9666666666666667, F1-score=0.9666666666666667
Class 22: Precision=0.8333333333333334, Recall=0.6451612903225806, F1-score=0.7272727272727272
Class 23: Precision=0.84, Recall=0.7924528301886793, F1-score=0.8155339805825242
Class 24: Precision=0.7586206896551724, Recall=0.9166666666666666, F1-score=0.830188679245283
...
Class 98: Precision=0.6944444444444444, Recall=0.8928571428571429, F1-score=0.78125
Class 99: Precision=0.5416666666666666, Recall=0.6842105263157895, F1-score=0.6046511627906976
Class 100: Precision=0.9583333333333334, Recall=0.7666666666666667, F1-score=0.8518518518518519
Accuracy: 85.61549100968188%
```

Input: classifier: m-nn, m = 5

```
Image ID: 1 is similar to label 0
Image ID: 3 is similar to label 0
Image ID: 5 is similar to label 0
Image ID: 7 is similar to label 0
Image ID: 9 is similar to label 0
Image ID: 11 is similar to label 1
Image ID: 13 is similar to label 0
Image ID: 15 is similar to label 1
Image ID: 17 is similar to label 0
Image ID: 19 is similar to label 1
Image ID: 21 is similar to label 0
Image ID: 23 is similar to label 1
Image ID: 25 is similar to label 0
Image ID: 27 is similar to label 0
Image ID: 29 is similar to label 0
Image ID: 31 is similar to label 0
Image ID: 33 is similar to label 1
Image ID: 35 is similar to label 0
Image ID: 37 is similar to label 1
Image ID: 39 is similar to label 0
Image ID: 41 is similar to label 0
Image ID: 43 is similar to label 0
Image ID: 45 is similar to label 0
Image ID: 47 is similar to label 0
Image ID: 49 is similar to label 1
...
```

```
Class 0: Precision=0.9371069182389937, Recall=0.6866359447004609, F1-score=0.7925531914893617
Class 1: Precision=0.7518248175182481, Recall=0.944954128440367, F1-score=0.8373983739837397
Class 2: Precision=0.9615384615384616, Recall=1.0, F1-score=0.9803921568627451
Class 3: Precision=0.9827586206896551, Recall=1.0, F1-score=0.9913043478260869
Class 4: Precision=0.9, Recall=1.0, F1-score=0.9473684210526316
Class 5: Precision=0.9476190476190476, Recall=0.995, F1-score=0.9707317073170733
Class 6: Precision=0.8, Recall=0.19047619047619047, F1-score=0.3076923076923077
Class 7: Precision=0.8095238095238095, Recall=0.8095238095238095, F1-score=0.8095238095238095
Class 8: Precision=0.9565217391304348, Recall=0.9166666666666666, F1-score=0.9361702127659574
Class 9: Precision=0.9090909090909091, Recall=0.7407407407407407, F1-score=0.8163265306122449
Class 10: Precision=0.9, Recall=0.782608695652174, F1-score=0.8372093023255814
Class 11: Precision=0.8888888888888888, Recall=1.0, F1-score=0.9411764705882353
Class 12: Precision=0.953125, Recall=0.953125, F1-score=0.953125
Class 13: Precision=0.8823529411764706, Recall=0.9183673469387755, F1-score=0.9
Class 14: Precision=1.0, Recall=0.18181818181818182, F1-score=0.3076923076923077
Class 15: Precision=1.0, Recall=0.9523809523809523, F1-score=0.975609756097561
Class 16: Precision=0.9375, Recall=0.6521739130434783, F1-score=0.7692307692307693
Class 17: Precision=0.9259259259259259, Recall=1.0, F1-score=0.9615384615384615
Class 18: Precision=0.9473684210526315, Recall=0.8571428571428571, F1-score=0.9
Class 19: Precision=0.9841269841269841, Recall=1.0, F1-score=0.9919999999999999
Class 20: Precision=0.9523809523809523, Recall=0.8695652173913043, F1-score=0.909090909090909
Class 21: Precision=0.9375, Recall=1.0, F1-score=0.967741935483871
Class 22: Precision=0.8333333333333334, Recall=0.4838709677419355, F1-score=0.6122448979591837
Class 23: Precision=0.8936170212765957, Recall=0.7924528301886793, F1-score=0.8400000000000001
Class 24: Precision=0.8, Recall=0.6666666666666666, F1-score=0.7272727272727272
...
Class 98: Precision=0.6, Recall=0.9642857142857143, F1-score=0.7397260273972602
Class 99: Precision=0.5789473684210527, Recall=0.5789473684210527, F1-score=0.5789473684210527
Class 100: Precision=0.92, Recall=0.7666666666666667, F1-score=0.8363636363636363
Accuracy: 87.96680497925311%
```

Input: classifier: m-nn, m = 10

```
Image ID: 1 is similar to label 0
Image ID: 3 is similar to label 1
Image ID: 5 is similar to label 0
Image ID: 7 is similar to label 0
Image ID: 9 is similar to label 0
Image ID: 11 is similar to label 1
Image ID: 13 is similar to label 0
Image ID: 15 is similar to label 0
Image ID: 17 is similar to label 0
Image ID: 19 is similar to label 1
Image ID: 21 is similar to label 0
Image ID: 23 is similar to label 1
Image ID: 25 is similar to label 0
Image ID: 27 is similar to label 0
Image ID: 29 is similar to label 0
Image ID: 31 is similar to label 0
Image ID: 33 is similar to label 1
Image ID: 35 is similar to label 0
Image ID: 37 is similar to label 1
Image ID: 39 is similar to label 0
Image ID: 41 is similar to label 1
Image ID: 43 is similar to label 0
Image ID: 45 is similar to label 0
Image ID: 47 is similar to label 0
Image ID: 49 is similar to label 0
```

```
Class 0: Precision=0.9137931034482759, Recall=0.7327188940092166, F1-score=0.813299232736573
Class 1: Precision=0.7821011673151751, Recall=0.9220183486238532, F1-score=0.8463157894736842
Class 2: Precision=0.9615384615384616, Recall=1.0, F1-score=0.9803921568627451
Class 3: Precision=0.9708029197080292, Recall=1.0, F1-score=0.9851851851851852
Class 4: Precision=0.9310344827586207, Recall=1.0, F1-score=0.9642857142857143
Class 5: Precision=0.9255813953488372, Recall=0.995, F1-score=0.9590361445783133
Class 6: Precision=1.0, Recall=0.047619047619047616, F1-score=0.0909090909090909
Class 7: Precision=0.9, Recall=0.8571428571428571, F1-score=0.8780487804878048
Class 8: Precision=0.9565217391304348, Recall=0.9166666666666666, F1-score=0.9361702127659574
Class 9: Precision=0.9523809523809523, Recall=0.7407407407407407, F1-score=0.8333333333333334
Class 10: Precision=0.85, Recall=0.7391304347826086, F1-score=0.7906976744186046
Class 11: Precision=0.9375, Recall=0.9375, F1-score=0.9375
Class 12: Precision=0.9523809523809523, Recall=0.9375, F1-score=0.9448818897637795
Class 13: Precision=0.9574468085106383, Recall=0.9183673469387755, F1-score=0.9375000000000001
Class 14: Precision=0.8333333333333334, Recall=0.22727272727272727, F1-score=0.35714285714285715
Class 15: Precision=0.975, Recall=0.9285714285714286, F1-score=0.9512195121951122
Class 16: Precision=0.9411764705882353, Recall=0.6956521739130435, F1-score=0.7999999999999999
Class 17: Precision=0.8928571428571429, Recall=1.0, F1-score=0.9433962264150945
Class 18: Precision=0.9473684210526315, Recall=0.8571428571428571, F1-score=0.9
Class 19: Precision=0.9841269841269841, Recall=1.0, F1-score=0.9919999999999999
Class 20: Precision=1.0, Recall=0.782608695652174, F1-score=0.878048780487805
Class 21: Precision=0.9090909090909091, Recall=1.0, F1-score=0.9523809523809523
Class 22: Precision=0.8, Recall=0.12903225806451613, F1-score=0.2222222222222222
Class 23: Precision=0.8823529411764706, Recall=0.8490566037735849, F1-score=0.8653846153846154
Class 24: Precision=0.7777777777777778, Recall=0.5833333333333334, F1-score=0.6666666666666666
...
Class 98: Precision=0.49122807017543857, Recall=1.0, F1-score=0.6588235294117647
Class 99: Precision=0.5555555555555556, Recall=0.5263157894736842, F1-score=0.5405405405405405
Class 100: Precision=0.8846153846153846, Recall=0.7666666666666667, F1-score=0.8214285714285715
Accuracy: 86.88335638543107%
```

Input: classifier: decision-tree

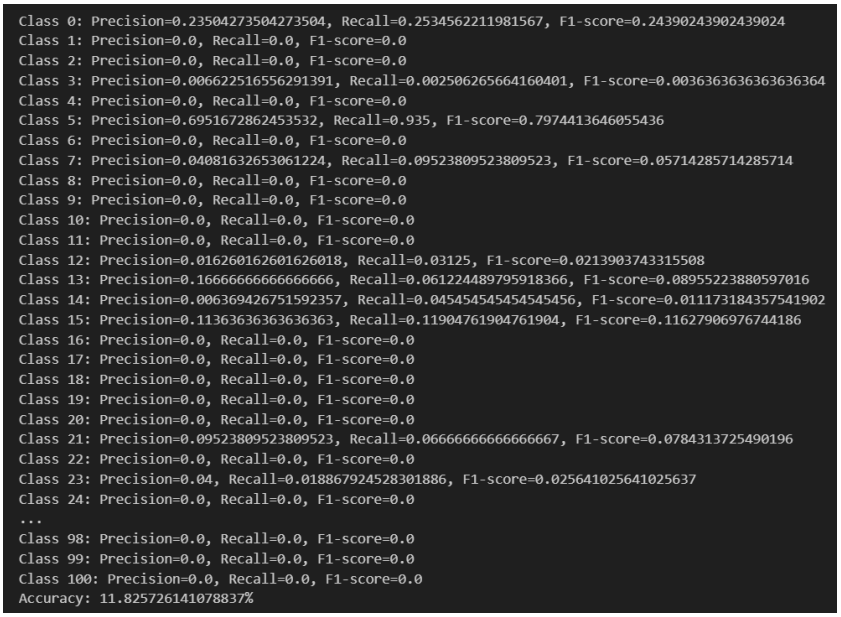


Fig. (a) Fig. (b)

Interpretation of Output:

1. The program outputs the similar labels for each of the odd numbered images in the dataset as shown in Fig. (a) above.
   Label 0 is most similar to the Image with ID 7
2. In Fig. (b), we see per label precision, recall and f1-score and the overall accuracy of the system

Observations from output:

Decision tree results:

As stated earlier, citing the logistical reasons, we decided to reduce the dimensionality of the training dataset for the decision trees to 150. This seems to have hurt the classification accuracy of the resulting classifier.
This is more clear when comparing it to the results of sklearn's decision tree classifier over the feature space of 1000 dimensions, as shown below.

```
Class 0: Precision=0.7946428571428571, Recall=0.8202764976958525, F1-score=0.8072562358276645
Class 1: Precision=0.8195121951219512, Recall=0.7706422018348624, F1-score=0.7943262411347518
Class 2: Precision=0.9361702127659575, Recall=0.88, F1-score=0.9072164948453608
Class 3: Precision=0.9897172236503856, Recall=0.9649122807017544, F1-score=0.9771573604060915
Class 4: Precision=1.0, Recall=0.9259259259259259, F1-score=0.9615384615384615
Class 5: Precision=0.9503722084367245, Recall=0.9575, F1-score=0.9539227895392279
Class 6: Precision=0.09523809523809523, Recall=0.09523809523809523, F1-score=0.09523809523809523
Class 7: Precision=0.875, Recall=0.6666666666666666, F1-score=0.7567567567567567
Class 8: Precision=0.9473684210526315, Recall=0.75, F1-score=0.8372093023255814
Class 9: Precision=0.5652173913043478, Recall=0.48148148148148145, F1-score=0.52
Class 10: Precision=0.7307692307692307, Recall=0.8260869565217391, F1-score=0.7755102040816326
Class 11: Precision=1.0, Recall=0.875, F1-score=0.9333333333333333
Class 12: Precision=0.8524590163934426, Recall=0.8125, F1-score=0.8319999999999999
Class 13: Precision=0.7272727272727273, Recall=0.8163265306122449, F1-score=0.7692307692307693
Class 14: Precision=0.32, Recall=0.36363636363636365, F1-score=0.3404255319148936
Class 15: Precision=0.5102040816326531, Recall=0.5952380952380952, F1-score=0.5494505494505494
Class 16: Precision=0.5471698113207547, Recall=0.6304347826086957, F1-score=0.5858585858585859
Class 17: Precision=0.7931034482758621, Recall=0.92, F1-score=0.851851851851852
Class 18: Precision=0.75, Recall=0.5714285714285714, F1-score=0.6486486486486486
Class 19: Precision=0.9824561403508771, Recall=0.9032258064516129, F1-score=0.9411764705882352
Class 20: Precision=0.9090909090909091, Recall=0.43478260869565216, F1-score=0.5882352941176471
Class 21: Precision=0.84, Recall=0.7, F1-score=0.7636363636363636
Class 22: Precision=0.7083333333333334, Recall=0.5483870967741935, F1-score=0.6181818181818182
Class 23: Precision=0.5254237288135594, Recall=0.5849056603773585, F1-score=0.5535714285714286
Class 24: Precision=0.64, Recall=0.6666666666666666, F1-score=0.6530612244897959
...
Class 98: Precision=0.78125, Recall=0.8928571428571429, F1-score=0.8333333333333334
Class 99: Precision=0.3125, Recall=0.2631578947368421, F1-score=0.2857142857142857
Class 100: Precision=0.6666666666666666, Recall=0.6666666666666666, F1-score=0.6666666666666666
Accuracy: 77.82%
```

Along the similar lines, the accuracy for the m-nn classifiers, which are constructed using the full 1000 dimensional feature space is around 85%, thus reinforcing our previous observations.

Task 4

Specification

User specifies the t, number of layers, number of hashes per layer and the query image id. An index structure will be created to store the even numbered images from the dataset. Program will then output the ids of t most similar images based on this index structure.

Description

Task 4a

1. User is prompted to enter the value of t, number of layers and number of hashes per layer.
2. Feature space selected for this task is the output of a fully connected layer of Resnet50 model. The feature vectors for all the images are fetched from the database.
3. Feature vectors for even numbered images are fetched to implement LSH.
4. Random projections are generated which will be used to hash the feature vectors.
5. Hash of the vector in a layer is calculated using the dot product of the projection in that layer with the feature vector.
6. If the dot product is greater than or equal to 0, '1' is appended to the code, else '0' is appended.
7. Query vector, after its hashcode is generated, is appended to the appropriate hash table.

Task 4b

1. The program iterates through each layer of the hash tables, generates the hash code for the vector into consideration and adds it to the set of visited buckets.
2. Then, each hash table in the current layer is checked.
3. If the hamming distance between the current hash code and the table key is less than or equal to 1, we consider the images from within that bucket.
4. Since most of our search space has been pruned using the hash tables, euclidean distance is used to find the similarities between the odd numbered vector and the set of similar_images formed during the search.
5. Top t images with least distances are returned.

Design

Hamming distance to check neighboring hash buckets:
Hash functions in LSH might result in hash collisions where dissimilar items are hashed into the same bucket. Similar items might not always land in the same bucket due to the nature of the hashing process. Checking the neighboring buckets expands the search space while increasing the likelihood of finding the similar items to our query image.

Sorting the candidates using euclidean distance:
After getting the candidate images from the hash buckets, we calculate the euclidean distance between the candidates and the query image vector. We then pick the top t images with the least euclidean distance as our most similar t images.

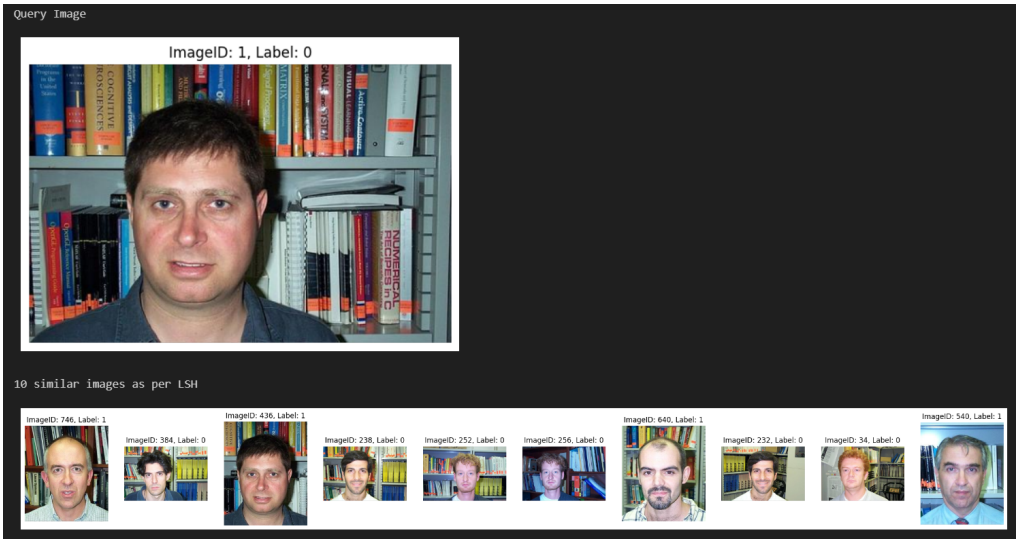
Outputs

Input:

1. L = 3, h = 3, t = 10

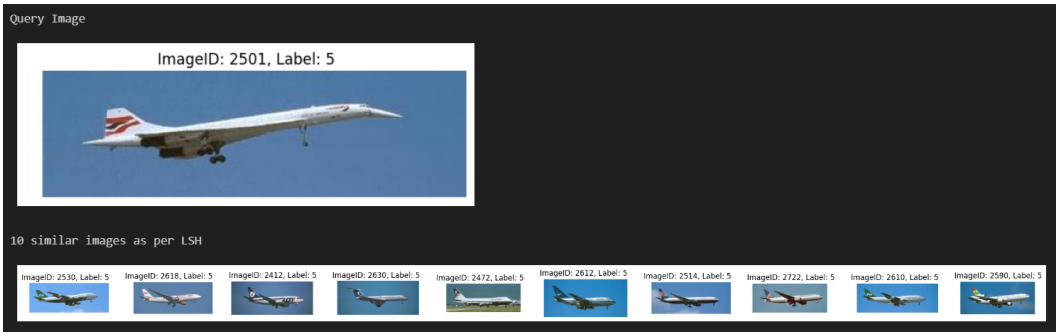
    a. Query Image ID: 1

```
Number of unique images considered: 3939
Overall number of images considered: 8340

Input image ID: 1
Input image label: 0
10 images
Label: 1, ID: 746
Label: 0, ID: 384
Label: 1, ID: 436
Label: 0, ID: 238
Label: 0, ID: 252
Label: 0, ID: 256
Label: 1, ID: 640
Label: 0, ID: 232
Label: 0, ID: 34
Label: 1, ID: 540
```



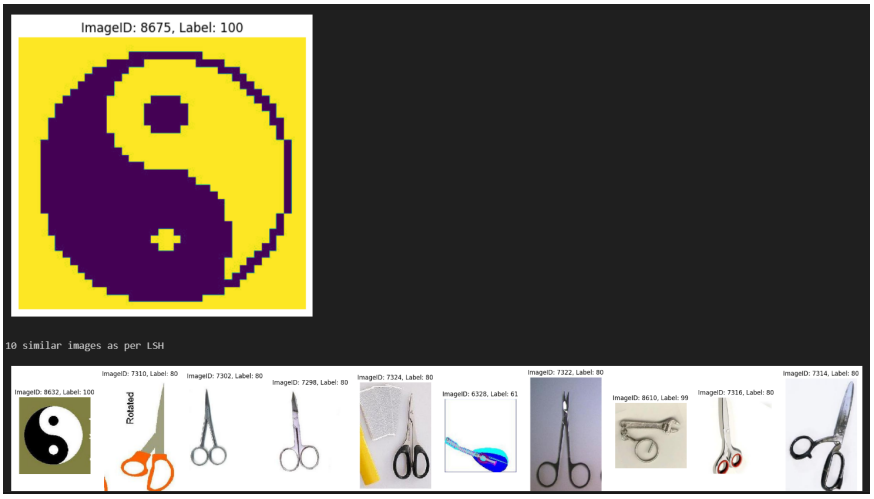
b. Query Image ID: 2501

```
Number of unique images considered: 4015
Overall number of images considered: 9276

Input image ID: 2501
Input image label: 5
10 images
Label: 5, ID: 2530
Label: 5, ID: 2618
Label: 5, ID: 2412
Label: 5, ID: 2630
Label: 5, ID: 2472
Label: 5, ID: 2612
Label: 5, ID: 2514
Label: 5, ID: 2722
Label: 5, ID: 2610
Label: 5, ID: 2590
```



c. Query Image ID: 8675

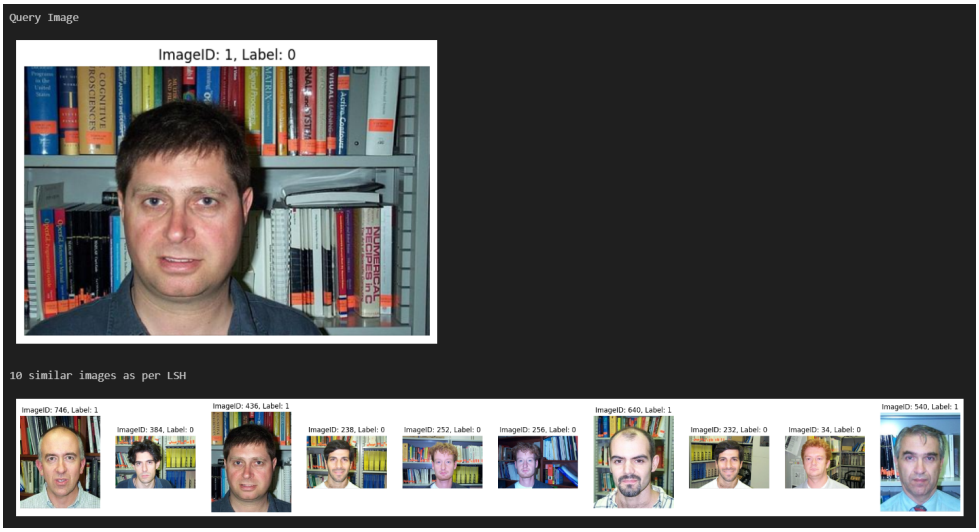
```
Number of unique images considered: 4114
Overall number of images considered: 10007

Input image ID: 8675
Input image label: 100
10 images
Label: 100, ID: 8632
Label: 80, ID: 7310
Label: 80, ID: 7302
Label: 80, ID: 7298
Label: 80, ID: 7324
Label: 61, ID: 6328
Label: 80, ID: 7322
Label: 99, ID: 8610
Label: 80, ID: 7316
Label: 80, ID: 7314
```



2. L = 10, h = 10, t = 10

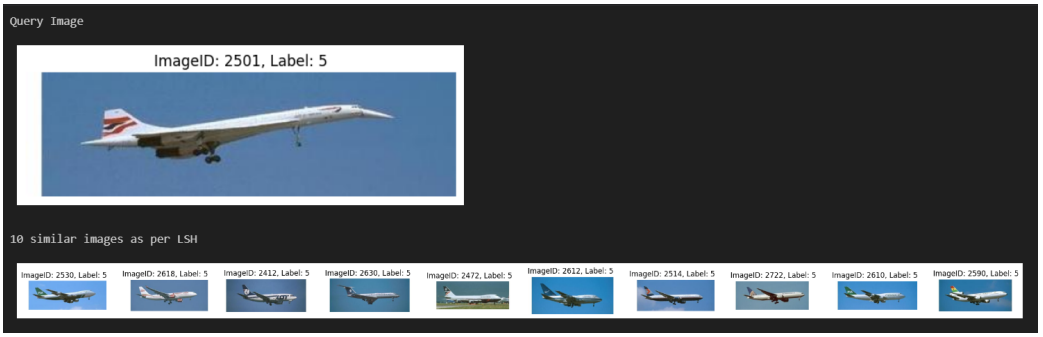
a. Query Image ID: 1

```
Number of unique images considered: 694
Overall number of images considered: 885

Input image ID: 1
Input image label: 0
10 images
Label: 1, ID: 746
Label: 0, ID: 384
Label: 1, ID: 436
Label: 0, ID: 238
Label: 0, ID: 252
Label: 0, ID: 256
Label: 1, ID: 640
Label: 0, ID: 232
Label: 0, ID: 34
Label: 1, ID: 540
```



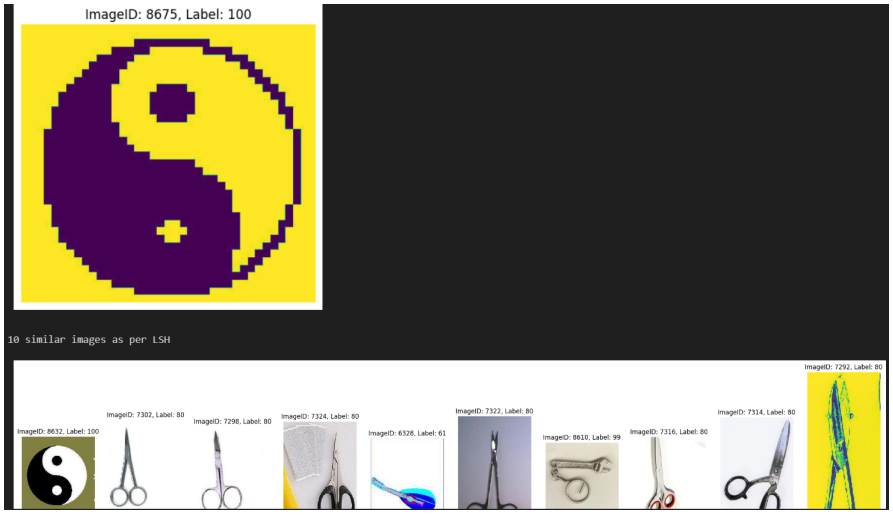
b. Query Image ID: 2501

```
Number of unique images considered: 918
Overall number of images considered: 1850

Input image ID: 2501
Input image label: 5
10 images
Label: 5, ID: 2530
Label: 5, ID: 2618
Label: 5, ID: 2412
Label: 5, ID: 2630
Label: 5, ID: 2472
Label: 5, ID: 2612
Label: 5, ID: 2514
Label: 5, ID: 2722
Label: 5, ID: 2610
Label: 5, ID: 2590
```



c. Query Image ID: 8675

```
Number of unique images considered: 2814
Overall number of images considered: 2814

Input image ID: 8675
Input image label: 100
10 images
Label: 100, ID: 8632
Label: 80, ID: 7310
Label: 80, ID: 7302
Label: 80, ID: 7298
Label: 80, ID: 7324
Label: 61, ID: 6328
Label: 80, ID: 7322
Label: 99, ID: 8610
Label: 80, ID: 7316
Label: 80, ID: 7314
```

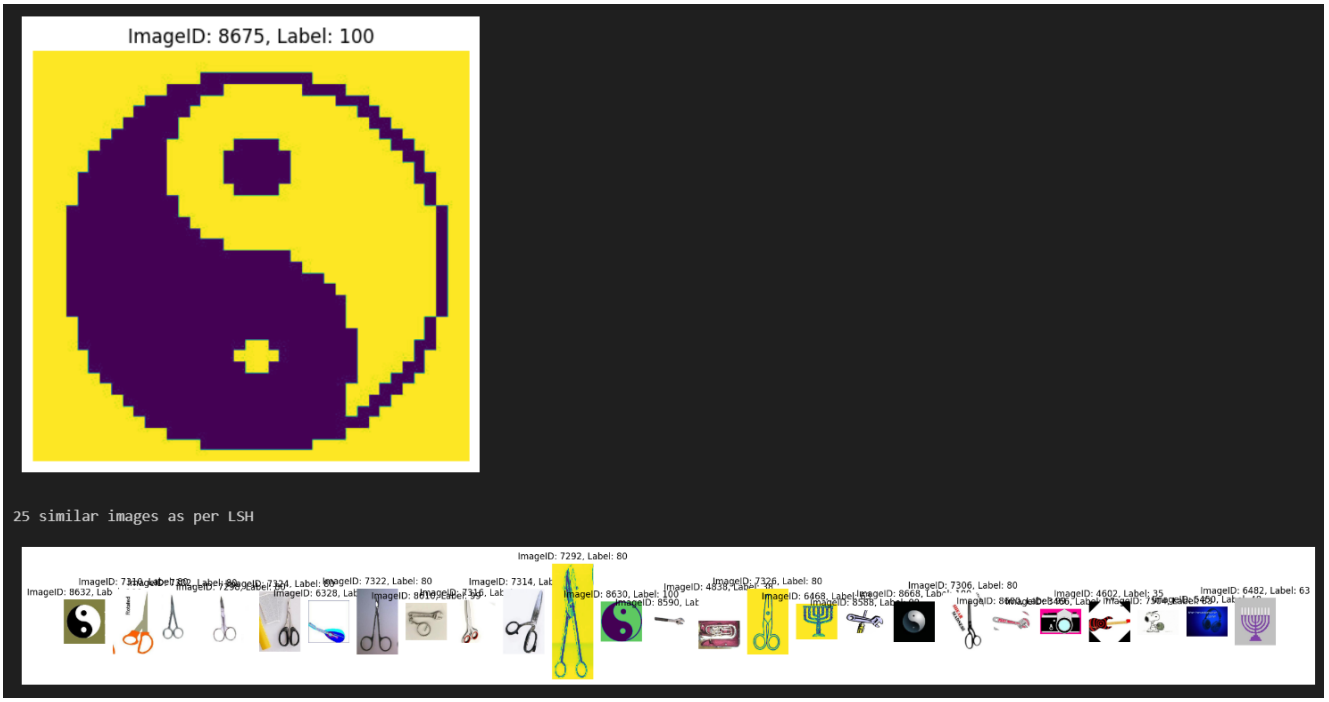


Output interpretation

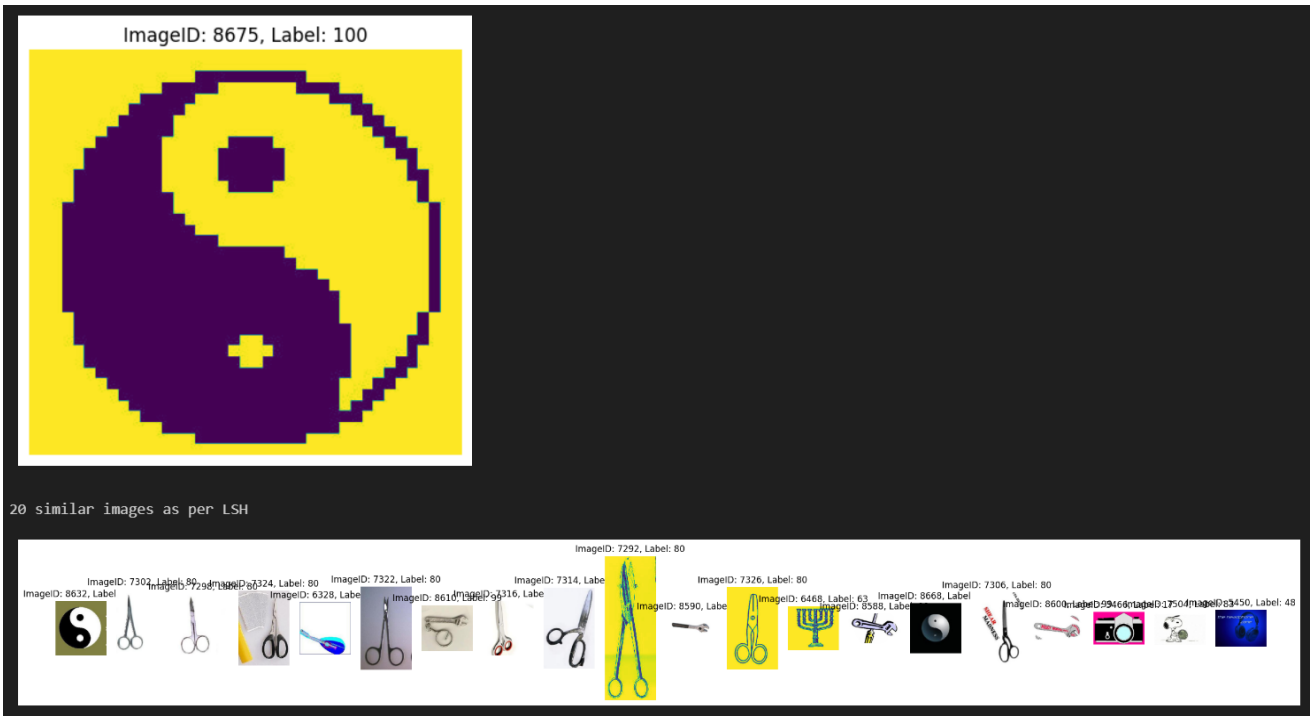
1. The program outputs the list of "t" image IDs similar to the query image ID and their associated labels. It also displays the total number of images and total number of unique images considered during search.
2. Secondly, the program visualizes similar images.

Output observations

1. One thing we observe is the number of images that are considered during the search. Across the two LSH configurations, we see a significant reduction in that number for the same images. This is expected as lesses the layers and hashes per layers, more images will be clubbed within a bucket.
2. One particular query image which does not seem to perform well is that of image ID 8675. This particular result, seems to denote some sort of shape bias within the feature descriptors. Extending this query to t = 20, we observe the similar trends of objects with circles being considered similar to the query image.



Output for t = 20, L = 3, h = 3



Output for t = 20, L = 10, h = 10

Task 5

Specification

User runs the task 4b with the desired layer and hash configuration. The user is then prompted to tag the LSH results into 4 categories, namely, very relevant (R+), relevant (R), irrelevant (I), very irrelevant (I-). The program then constructs an SVM classifier based on the feedback and finds the relevant images in the rest of the dataset.

Description

1. User is prompted to enter “t”, number of layers and number of hashes per layers. Based on this, index structure is created leveraging LSH.
2. User is then prompted to enter a query image id, and the LSH tool fetches t images similar to the query image id
3. Then the user is prompted to tag the above results into 4 categories (R+, R, I, I-).
4. This feedback is used as the training data for the classifier to be used ahead.
5. This training data is further augmented by adding the irrelevant as well as relevant data points to it.
6. SVM classifier is trained using this training data.
7. This classifier is then used to classify the rest of the data points from the dataset and relevant images are shown.

Design

Classification into two classes, relevant and irrelevant, only:

- Since one image is only relevant to one label we have taken into consideration only 2 classes - Relevant and Irrelevant for our classification.
- The Very relevant and relevant feedbacks are merged together and Irrelevant and Very Irrelevant are merged together

Augmenting the training dataset:

- The user provided feedback is very less in amount as compared to the whole dataset this results in skewed output. To compensate for this we have augmented data from the label of relevant image and irrelevant image provided by the user.
- This helps us in give more information to the classification model

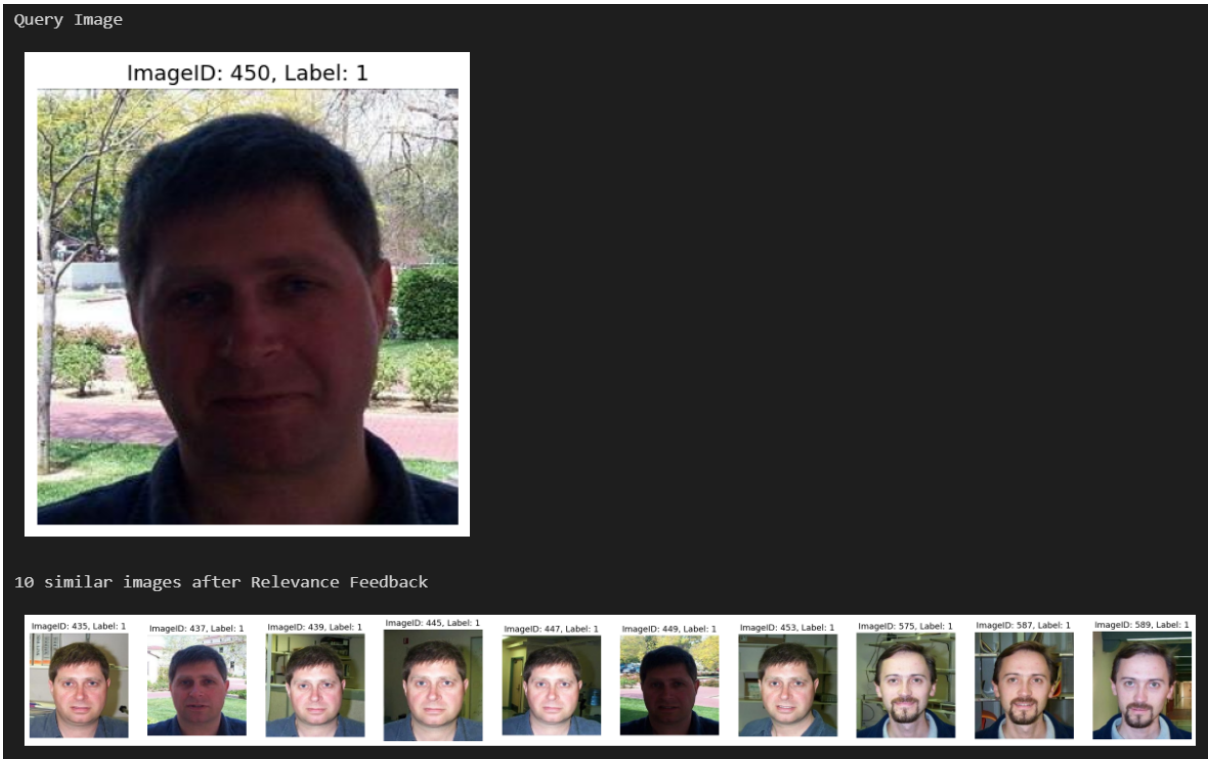
Handling No Irrelevant data scenario:

- In some of the query image id we get only relevant data through our LSH tool so when the user provides feedback with only relevant information we do not have data for irrelevant information.

- This scenario is handled by adding irrelevant information (image data which do not match our query label)

Output

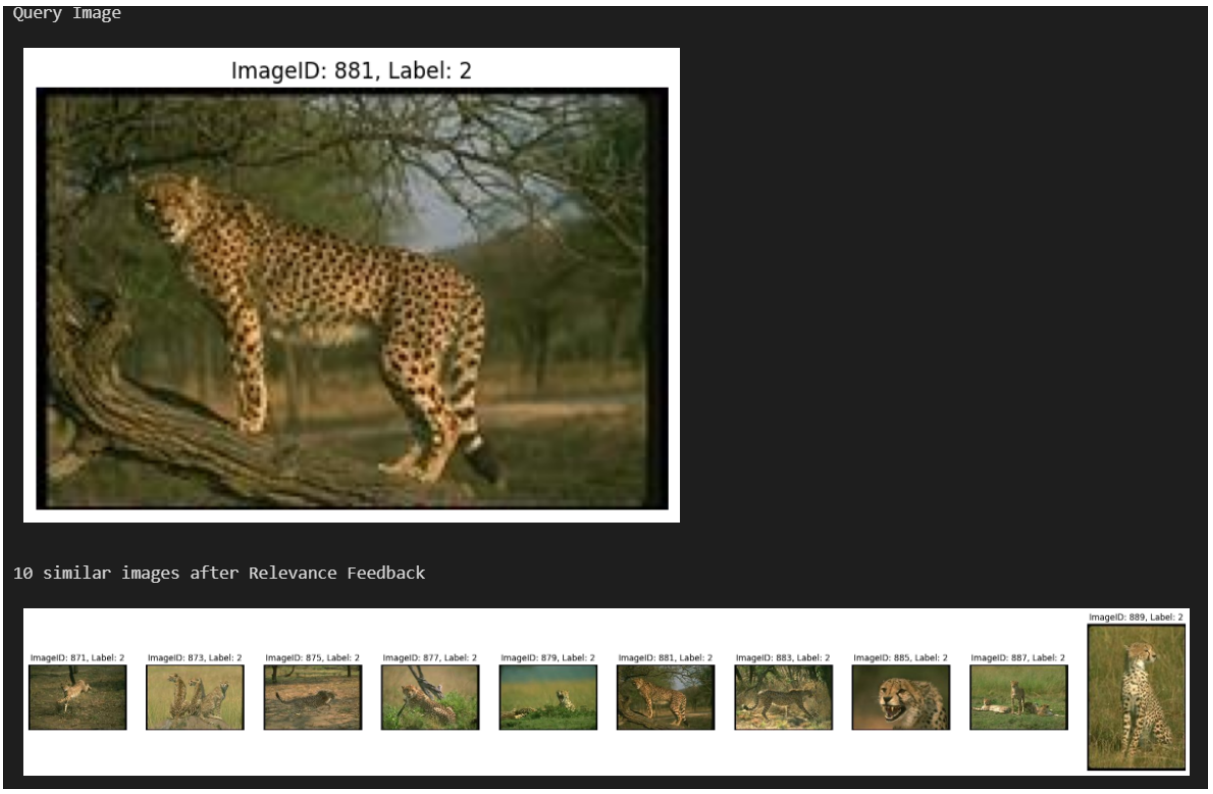
Query: t=10, L =10 h= 10, ImageID = 1



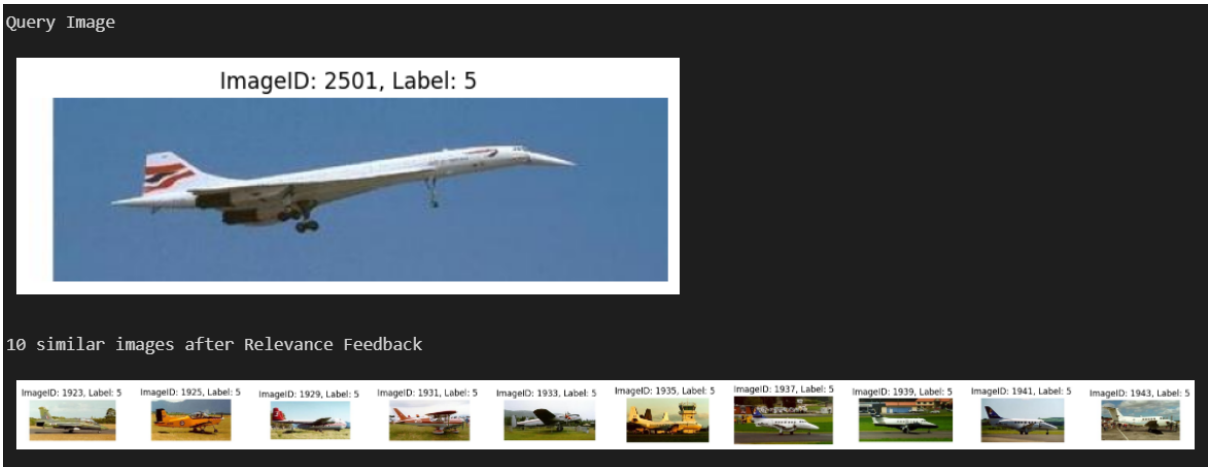
Interpretation:

- The output for this query gives perfect results as all similar images match with the query label.
- This is because we have information of both relevant and irrelevant images in balanced nature and also the overall data of label 0 is significantly high in the dataset which makes it easier to classify.
- More distinct data gives more information to the classifier.

Query: t=10, L =10 h= 10, ImageID = 881



Query: t=10, L =10 h= 10, ImageID = 2501

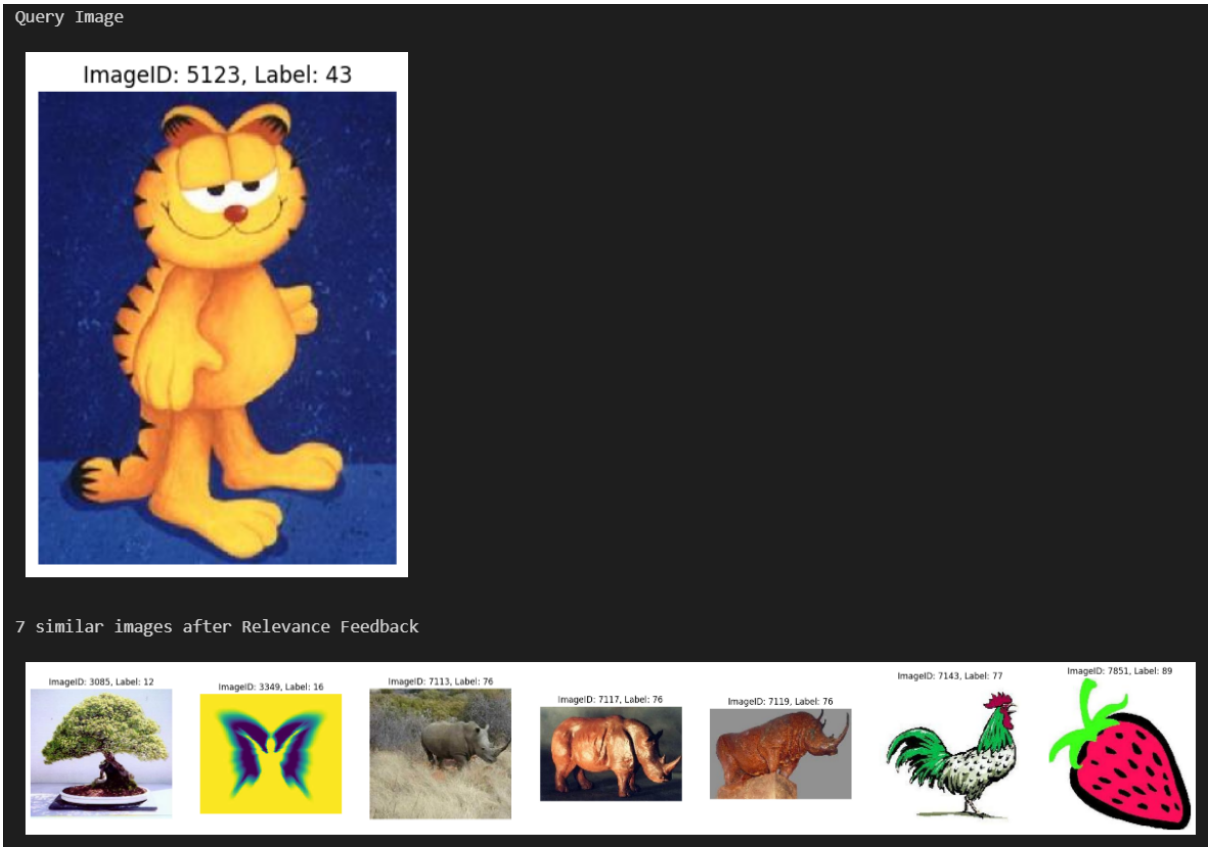


Interpretation of Output:

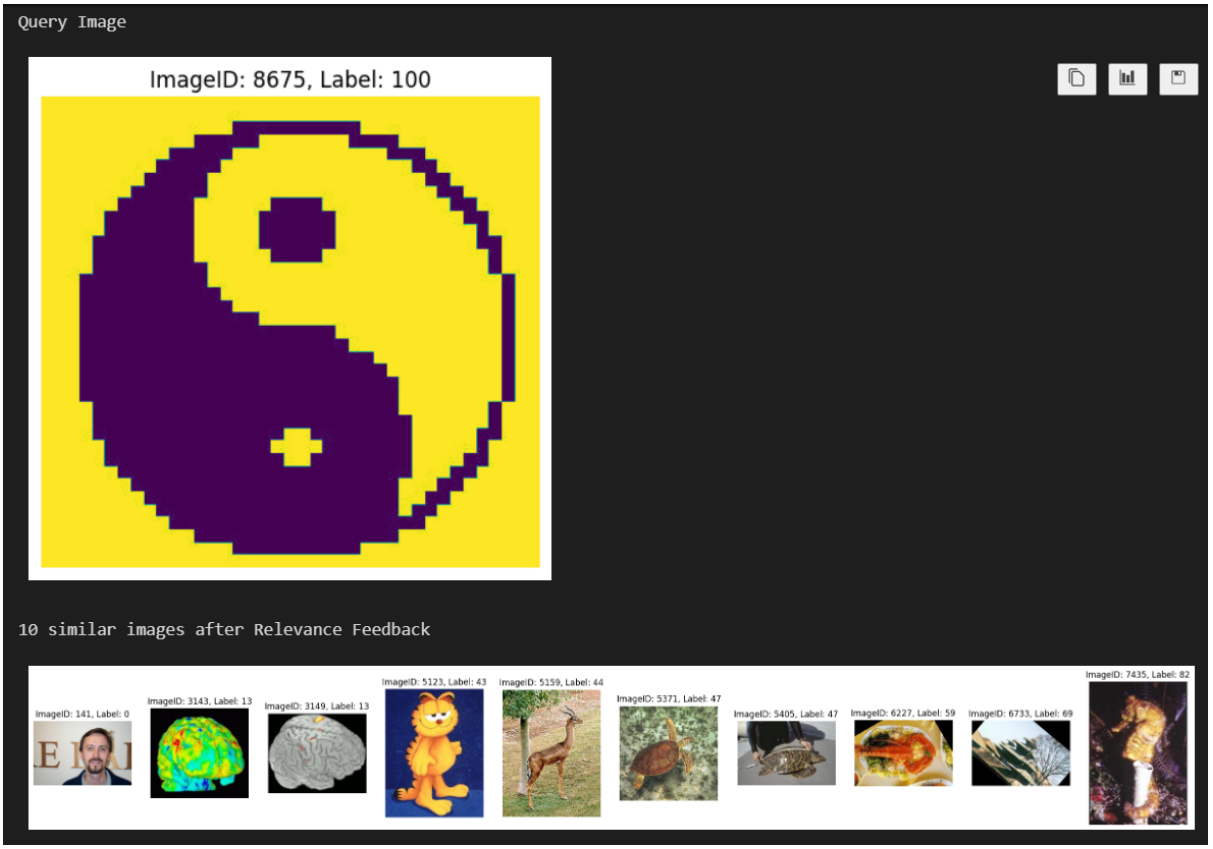
Query : t=10, L =10 h= 10, ImageID = 881 and Query : t=10, L =10 h= 10, ImageID = 2501:

- The output for this query gives perfect results as all similar images match with the query label
- This is because we get all relevant information and SVM is able to skew the results to our data
- We have logically augmented some irrelevant information to provide 2 classes to the SVM classifier

Query: t=10, L =10 h= 10, ImageID = 5123



t=10, L =10 h= 10, ImageID = 8675



Interpretation

Query : t=10, L =10 h= 10, ImageID = 5123 and Query : t=10, L =10 h= 10, ImageID = 8675:

- The output is far from accurate. This is not only because of a smaller dataset for that label, but also because of the small amount of feedback.
- Thus, the SVM classifier is not able to create a good classifier with this little information leading in discrepancy in the results
- Adding more irrelevant information also does not help as the ratio of relevant and irrelevant becomes too low which means SVM is not able to identify relevant information. A good balance between relevant and irrelevant information feedback is needed.

INTERFACE SPECIFICATIONS

The runtime is the Jupyter notebook environment, allowing for step-through execution, cell by cell. It is also easily interactive in displaying images and getting user inputs. The matplotlib library is used to visualize images.

Task 0a: Inherent Dimensionality Computation:

The program computes and prints a single value for the inherent dimensionality of the even-numbered images in the dataset.

Task 0b: Label-specific Inherent Dimensionality:

The program computes and prints the inherent dimensionality associated with each unique label using the even-numbered Caltech101 images.

Task 1: Latent Semantic Analysis and Classification:

The user is required to input the value of k, and the program outputs per-label precision, recall and F1-score for label predictions for the odd-numbered images along with the overall accuracy.

Task 2: Clustering and Visualization:

For a given value C, the program creates and visualizes clusters for the even-numbered images of each label and then outputs per-label precision, recall and F1-score for label predictions for the odd-numbered images along with the overall accuracy.

Task 3: Classifier Implementation:

The user is to select one of the three specified classifiers, and the program outputs per-label precision, recall and F1-scores for label predictions for the odd-numbered images along with the overall accuracy.

Task 4: Locality Sensitive Hashing (LSH):

The user is required to input the number of layers L and the number of hashes per layer h, a target image, and a query value t, and the program prints the number of unique and overall images considered during the process and visualizes t similar images as per LSH.

Task 5: Relevance Feedback Systems:

The user is required to run task 4 with input values of t, number of layers and number of hashes per layer and a target image.

Once we get the output of task 4, the user is prompted to enter the relevance of each image predicted by LSH. The user can tag the image in 4 categories - Very Relevant (R+), Relevant(R), Irrelevant(I) and Very Irrelevant (I-).

The program takes the feedback and outputs similar images based on this feedback by building a classifier.

SYSTEM REQUIREMENTS/ INSTALLATION AND EXECUTION INSTRUCTIONS

Operating System - Agnostic

Hardware Requirements (Recommended):
16GB RAM, 4GB GPU memory, sufficient disk space for libraries and data (5+ GB)

Software Dependencies:
Python 3.7 or above.

Configurations:
Configuration of dataset path is needed as per host machine's directory system

Library Dependencies:
Necessary Python modules are listed in the requirements.txt file

Installations:
- MongoDB installation is required, either local or hosted/Atlas. Connection string may have to be modified as needed if not local.
- Users are required to install the Python libraries mentioned above.

Execution Instructions:
Jupyter Notebooks are preferred to run the code. To execute the tasks, the user can run the respective .ipynb files created for each task.

Miscellaneous:
The code is developed keeping a GPU-available system in mind. Minor modifications would be needed in case of execution on a non-GPU system, for the ResNet50 feature extraction part.

RELATED WORK

In the realm of multimedia and web databases, several studies have prepared the ground for the advanced methodologies used in Phase 3 of our project. Research on latent semantic analysis (LSA) has demonstrated its effectiveness in revealing underlying semantic structures within diverse datasets, offering insightful information into content relationships. The utilization of clustering algorithms, such as DBScan, has been explored extensively in image analysis, offering robust solutions for identifying meaningful patterns and structures. Classification studies employing a range of models, including m-NN, decision-tree, and personalized PageRank (PPR) classifiers, contribute to the understanding of several strategies for accurate label predictions.

Locality Sensitive Hashing (LSH) tools, as implemented in Task 4a, have been widely recognized for their efficiency in approximate nearest neighbor search, with seminal work by Andoni and Indyk providing theoretical foundations for practical implementations. Image search algorithms utilizing LSH structures align with the broader context of content-based retrieval systems. Information retrieval studies serve as an inspiration for relevance feedback systems, as Task 5 demonstrates, with notable contributions from Salton and Buckley in refining search results through user interactions.

This phase integrates and builds upon a combination of prior work in multimedia analysis, clustering, classification, and retrieval, showcasing a synthesis of established methodologies and innovative approaches for a holistic exploration of the Caltech101 image dataset.

CONCLUSION

Phase 3 marks the culmination of the project, showcasing an extensive examination of the Caltech101 image dataset through advanced techniques in image analysis and retrieval. The tasks accomplished have provided valuable insights into the dataset's inherent dimensionality, latent semantics, clustering structures, and variety of classification models. The successful application of Locality Sensitive Hashing (LSH) has proven to be effective in enabling rapid and accurate image retrieval, further enhancing the project's practical utility. The integration of relevance feedback systems reflects a user-centric approach, refining search results and promoting an adaptive system.

Through meticulous evaluation and reporting of precision, recall, F1-score, and overall accuracy metrics, we have measured the effectiveness of our methodologies, enabling a comprehensive understanding of the strengths and limitations of each approach. The visual representations of clusters and image thumbnails contribute to an intuitive grasp of the dataset's intrinsic grouping. This phase shows how advanced algorithms may be used in real-world circumstances and broadens our analytical capabilities in multimedia data.

As we conclude this phase, the knowledge acquired not only helps us comprehend the Caltech101 dataset better, but it also paves the way for the development of robust multimedia and web databases systems. The seamless integration of user interaction mechanisms makes our system appear not only analytical but also adaptive and user-centric, which is an essential feature in modern multimedia applications. The journey through Phase 3 propels us in the direction of a comprehensive and impactful multimedia analysis framework.

BIBLIOGRAPHY

1. K. Selçuk Candan, Maria Luisa Sapino. 2010. Data Management for Multimedia Retrieval, Cambridge University Press.
2. Torgerson, W. S. 1952. Multidimensional scaling: I. Theory and method. Psychometrika, 17(4), 401–419. https://doi.org/10.1007/bf02288916
3. Erich Schubert, Jörg Sander, Martin Ester, Hans Peter Kriegel, and Xiaowei Xu. 2017. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. ACM Trans. Database Syst. 42, 3, Article 19 (September 2017), 21 pages. https://doi.org/10.1145/3068335
4. Alexandr Andoni and Piotr Indyk. 2008. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. Commun. ACM 51, 1 (January 2008), 117–122. https://doi.org/10.1145/1327452.1327494
5. Gerard Salton and Chris Buckley. 1990. Improving Retrieval Performance by Relevance Feedback. Journal of the American Society for Information Science. 41, pp. 288-297

APPENDIX

In this phase, other than the initial brainstorming and planning, the team worked independently on their assigned tasks for the most part, however we constantly shared our findings and insights with each other to help everyone figure out optimal strategies on their tasks. Kaushik, Madhura, Pranav and Niraj handled the coding and respective sections of the report, while Pavan and Mohan handled the bulk of the report.