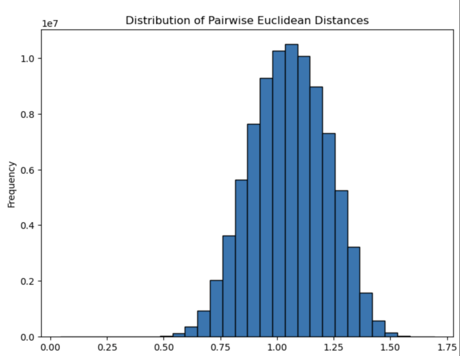
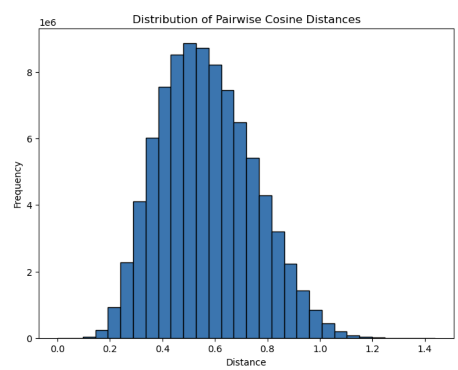
Final Report

Deep Density Clustering of Unconstrained Faces

Name: Divij Singla
Roll No: 210350

Journey of Improvements in the Algorithm

	Phase 1	Approaches Tried	Phase 2
		Feature Extraction	
1	• **DCNN** architecture was used to train CASIA-WebFace dataset for feature extraction. • Due to lack of GPU, only 300 iterations were performed instead of 750K iterations as suggested in the paper	• Tried to use **VLADCNN** architecture based on this paper • The data was huge (494414 images) and the computation was too expensive even on GPU	• Found a model pretrained on CASIA-WebFace dataset, based on Inception-ResNet Architecture which was very similar to the VLADCNN architecture • Finally the pretrained model was used for feature extraction
		Dataset Analysis	
2	Used euclidean distance measure to calculate pairwise distances and used $\epsilon = 0.23$ as suggested in paper to form $\epsilon-$neighborhoods	Dataset Analysis revealed that eucledian distance measure is not correct for forming neighborhoods as $\epsilon = 0.23$ is optimal when pairwise distances are calculated for cosine distances 	Used cosine measure while calculating distance measure. The dataset analaysis for cosine distances is given and final results were improved 

New Additions to the Algorithm

	Phase 1	Phase 2
	Clustering Algorithm	
1	Standard Agglomerative Clustering algorithm given in *sklearn.cluster* was used with a fixed distance-threshold of 50 percentile was used.	Studied another research paper and used the improved **Agglomerative Hierarchical Clustering** as described in it by implementing it by own in the code.
	Weighted Similarity Matrix	
2	While calculating density aware similarity matrix, each point's similarity was calculated according to the formula: $s(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{2}\left[\frac{\sum_{\mathbf{z}\in V(\mathbf{x}_j)} \mathcal{E}_i(\mathbf{z})}{\lvert V(\mathbf{x}_j)\rvert} + \frac{\sum_{\mathbf{z}\in V(\mathbf{x}_i)} \mathcal{E}_j(\mathbf{z})}{\lvert V(\mathbf{x}_i)\rvert}\right]$ Where $V(\mathbf{x}_p)$ define neighborhood x_p belongs to.	As mentioned in the mathematical analysis, I tried using a weighted similarity matrix to see its effect on results: $s(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{2}\left[\frac{\sum_{\mathbf{z}\in V(\mathbf{x}_j)} w_i(z)\mathcal{E}_i(\mathbf{z})}{\lvert V(\mathbf{x}_j)\rvert} + \frac{\sum_{\mathbf{z}\in V(\mathbf{x}_i)} w_j(z)\mathcal{E}_j(\mathbf{z})}{\lvert V(\mathbf{x}_i)\rvert}\right]$ Where $\mathbf{c_p}$ is the center of the neighborhood $V(x_p)$ and ϵ is a small constant to prevent division by zero and $w_p(\mathbf{z}) = \frac{1}{\lVert\mathbf{z}-\mathbf{c}_p\rVert+\epsilon}$
	Epsilon Calculation	
3	ϵ was calculated by randomly sampling 100 subjects out of CASIA dataset and computing the maximum likelihood estimator of the cosine distance between matched pairs	To speed up the process, randomly sampled 20% and 40% of the dataset were taken and 100 subjects were sampled from them.

Results

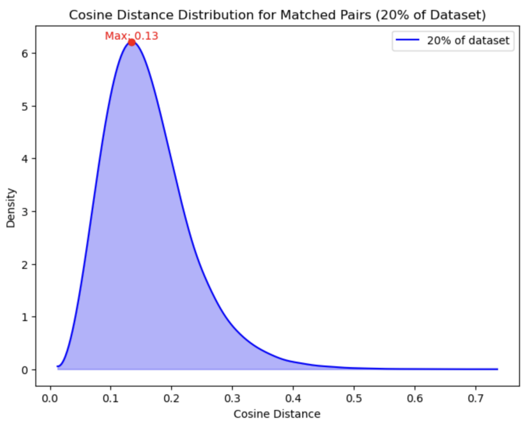
	Phase 1	Phase 2	Phase 2*	Improvement	Result in Paper (AHC)
Bcubed F-measure	0.63	**0.9268**	0.8819	47.11%	0.940
NMI	0.58	**0.8895**	0.8826	53.36%	0.987
F1 Measure	0.61	**0.9253**	0.8799	51.68%	Not Given

Table 1: Note: Phase 2* signifies results with weighted similarity matrix

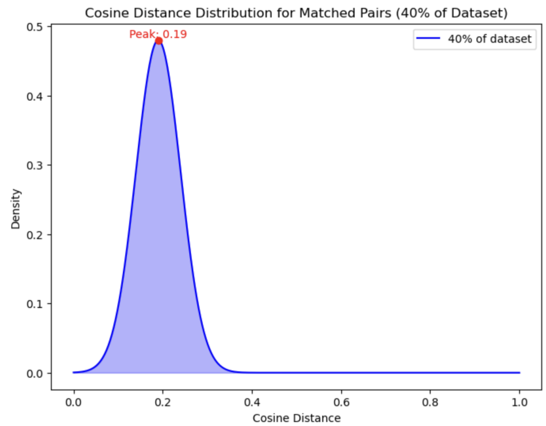
Note: As mentioned above, a pretrained resnet model similar to what described in paper was used, not the exact architecture due to computational constraints. Hence the results are not exactly same as the original paper, but very close.

20% of dataset	40% of dataset	100% of dataset
0.13	0.19	0.23

Table 2: Value of ϵ when subsets of CASIA WebFace dataset were taken



(a)



(b)

Figure 1: Figures side by side. (a) ϵ when 20% of dataset (b) ϵ when 40% of dataset

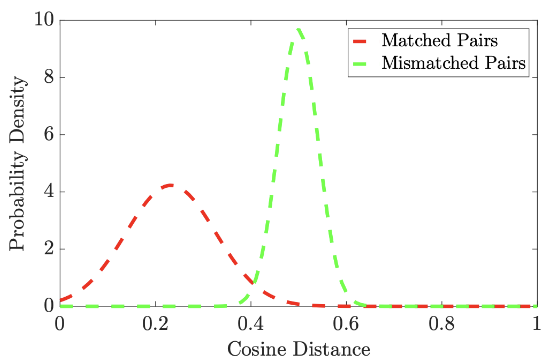


Figure 2: ϵ when 100% of dataset (Given in the paper itself)

References

- **Original Paper**: Deep Density Clustering of Unconstrained Faces Wei-An Lin Jun-Cheng Chen Carlos D. Castillo Rama Chellappa University of Maryland, College Park Link to Paper
- J. Zheng, J.-C. Chen, N. Bodla, V. M. Patel, and R. Chel- lappa. Vlad encoded deep convolutional features for uncon- strained face verification. In IEEE International Conference on Pattern Recognition, 2016. Link to Paper
- K. C. Gowda and G. Krishna. Agglomerative clustering us- ing the concept of mutual nearest neighbourhood. Pattern Recognition, 10(2):105–112, 1978. Link to Paper