Face recognition with a hyperbolic metric classification model

A. Trpin*, B. M. Boshkoska*†
* Faculty of imformation studies in Novo mesto, Ljubljanska cesta 31 a, Novo mesto, Slovenia
† Jožef Stefan Institute, Jamova 39, Ljubljana, Slovenia
alenka.trpin@fis.unm.si
biljana.mileva@fis.unm.si

Abstract—**Facial recognition systems are increasingly being used in smartphones as biometric security instead of passwords, or in airports as automated electronic passport control. It is also emerging in other forms of technology, for example in robotics. This creates large collections of photos that cannot be managed. Data mining tools and machine learning methods can be used to process these datasets and use them for prediction and classification. In such algorithms, using the most suitable distance metrics to define similarities among data has a crucial role. This paper investigates the usage of the Poincaré metric, which is used primarily in hyperbolic geometry, on the well-known k Nearest Neighbours classification algorithm. We applied this method to a database of face images. Our results indicate that the Poincaré metric is helpful with the Large Margin Nearest Neighbour learning (LMNN) method tested on an image dataset. We found that for small values of k, up to five, the algorithm using the Poincaré metric with the Edge filter gave the best results.**

Keywords*—*Poincaré metric, k Nearest Neighbours, image classification, large margin nearest neighbour classification

I. INTRODUCTION

The increasing use of computers, smartphones, apps, platforms, and cloud computing generates huge amounts of image data. While the massive amount of data may not benefit humans, the processing and analysis may lead to good prediction and classification results.

The situation is similar in the field of face recognition in various industries, such as banking. Instead of using one-time passwords, customers can authorize transactions by looking at their smartphone or computer [23]. Marketers have used facial recognition to enhance consumer experiences [18]. Hospitals use face recognition to help care for patients. This is used to access patient files, simplify patient registration, detect emotions and pain in patients, and even help identify specific genetic diseases [4].

In this paper, we focus on the face recognition task, for which we combine different images using the classification algorithm k Nearest Neighbours (kNN). In kNN, we modify the default Euclidean metric (EM) with a Poincaré metric (PM). We combined the PM model with the large margin nearest neighbour (LMNN) model. We also compare the performance of this algorithm with the Euclidean and Manhattan metrics, based on different k-values, on a data set of face images. The main contribution of this paper is that we have used a non-Euclidean metric - the Poincaré metric - for image classification. PM is used to learn the distances between the labelled instances, and by applying an appropriate linear transformation to the input data, we may improve the classification results.

We organize the paper as follows. Section 2 provides a short review of the related literature, and Section 3 explains the metric and dataset we used and the methodology. Section 4 describes the experiments used to evaluate the different filters' performance and compares results based on the Poincaré, Euclidean and Manhattan metrics. Section 5 provides our conclusions.

II. RELATED WORK

Facial recognition systems are prevalent because they are visible and valuable in security, especially in banks, shops, train stations, and airports [10]. Face and object recognition are increasingly used in self-driving vehicles [6].

Most of the work in the literature focuses on learning with Euclidean or Mahalanobis metrics. In addition to the Euclidean metric, many researchers have also pursued non-Euclidean metric learning, as it can address more complex intra-class and inter-class variations [3].

Different methods of measuring the distance between facial features for emotion recognition and identifying the optimal neural network structure to classify six facial ex- impressions have been used on small datasets [7]. However, this database was very small.

Recently, deep learning has been recognized to be very useful for image recognition, particularly for describing the trend toward autonomous driving [6]. However, the feature classification application uses only the CNN method.

Henaff et al. have improved the unsupervised target for learning the variability of natural signals. They applied Contrast Predictive Coding with linear classification on the ImageNet dataset [13].

Some authors proposed using the Viola-Jones algorithm for face detection and Histogram of Oriented Gradients (HOG) for feature extraction from face images. They used Principal Component Analysis (PCA) to reduce the dimensionality of the features and thus ex- tract the most important features. They used three different classifiers to classify facial expressions: SVM, kNN and multilayer perceptron neural network (MLPNN) [5]. While they used only one feature descriptor, in our work, unlike their study, we used several filters.

III. METHODOLOGY

Some authors have concluded that hyperbolic spaces with negative curvature are often better suited for learning image embedding and that hyperbolic spaces can thus be valuable for improving the performance of computer vision systems [9], [19].

The rapid development of computer vision techniques requires learning distance or similarity metrics between objects. LMNN (Large Margin Nearest Neighbour) is a technique for learning Mahalanobis metrics in a kNN classification environment using semidefinite programming [25]. Since the kNN algorithm is metric-based, we intuitively used the Poincaré metric in combination with the LMNN instead of the Euclidean metric.

A. Distance Metric

Distance metrics are a key part of several machine learning methods. These distance metrics are used to calculate the similarity between data points. Effective distance metrics improve the performance of a machine learning model, whether for classification tasks. The primary metrics are the Euclidean and Manhattan metrics, defined as follows.

We have two n-dimensional points $p, q \in \mathbb{R}^n$. Then we can define Manhattan or L_1 metric as:

$$d(p,q) = ||p-q||_1 = \sum_{i=1}^{n} |x_i - y_i|. \tag{1}$$

And well-known Euclidean metric is defined as:

$$d(p,q) = ||p-q||_2 = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}. \tag{2}$$

Unlike Euclidean geometry, where it is known that exactly one parallel to a given line can be drawn through a given point, in hyperbolic geometry, there are infinitely many parallels to a given line through a given point [21]. In this study, we use the Poincaré disc model, defined in n-dimensional space as $B^n = \{x \in \mathbb{R}^n, ||x|| < 1\}$. If P is the point that lies in the unit disk, then Poincaré distance between origin and point P is equal to

$$d(0,P) = \frac{1}{2} \ln \left(\frac{1+|P|}{1-|P|} \right). \tag{3}$$

where $|P|$ means Euclidean distance from the origin to P. The Poincaré distance between two different points P, Q is defined as:

$$d(P,Q) = \operatorname{arcosh} \left(1 + \frac{2||P-Q||^2}{(1-||P||^2)(1-||Q||^2)} \right) \tag{4}$$

[22], [14].

Fig. 1: Some images in the database

After conversion, we get

$$\begin{aligned} d(P,Q) = \ln \Bigg(1 + \frac{2||P-Q||^2}{(1-||P||^2)(1-||Q||^2)} + \\ + \sqrt{\left(1 + \frac{2||P-Q||^2}{(1-||P||^2)(1-||Q||^2)} \right)^2 - 1} \Bigg) \end{aligned} \tag{5}$$

[11].

B. Datasets of images

The database contains 5343 greyscale and colour images and 321 classes obtained from the following datasets:

- the Japanese Female Facial Expression (JAFFE) Database [15],
- grayscale MIT-CBCL face recognition database (/http://cbcl.mit.edu/software-datasets/heisele/facerecognition-database.html) [26],
- Psychological Image Collection at Stirling (PICS) (http://pics.stir.ac.uk/, 2020), which are greyscale and coloured images,
- Tom Mitchell image database, which contains 640 greyscale images, each image describable by pose, expression, eyes, and size (https://archive.ics.uci.edu/ml/datasets.php), and
- the Yale Face Database which contains 165 greyscale GIF images of 15 people [2].

The number of instances and classes for each dataset are summarized in Table I.

TABLE I: Characteristic and experiment settings of the database

Dataset	instances	classes	greyscale/colour
JAFFE Database	217	10	grey
MIT-CBCL	2209	10	grey
PICS	2182	266	colour, grey
UCI MLR	640	20	grey
Yale Face Database	165	15	grey

Some examples from the database are shown in the Fig. 1.

C. Image classification

The objective of the LMNN technique is that k nearest neighbours are always in the same class, while instances from different classes are separated by a large margin [25]. Our approach for image recognition is based on the LMNN model with the Poincaré metric to measure the distances between attributes. We used five different filters:

- Auto Colour Correlogram - AutoCC [8],
- Colour layout filter - Colour [17],
- PHOG [1],
- Edge histogram filter - Edge [16],
- Local binary patterns pyramid - BPP [20].

Some of the filters are based on the colour spectrum (Colour, AutoCC), some on texture analysis (BPP) and some on gradient (Phog) and some on edges (Edge). These filters extract each photo into a numerical attribute. The filters are based on different concepts of pixel capture, gradients and orientations of pixels, histograms, and their distribution. The filtering was used to achieve various purposes: to remove noise or unwanted elements from the image, add a particular look, use a filter for texture analysis and capture the spatial distribution of colours in the image. Data are represented with numerical attributes describing the image's appearance on a small scale. In our method, we have carried out the following steps.

- A normalization of the numerical features of each image to unify the pixel range. In addition, it is used to convert raw data into cleaner data and improves data efficiency. To avoid overfitting, we used the resample function.
- Feature standardization scales the vector values to have zero mean and uniform variance.
- The metric learning algorithm (LMNN) maintains k-nearest neighbours from the same class, while a large margin separates instances from different classes.
- Train, test and validate a classifier with kNN and PM. For the database, we randomly split into a test training set and a validation set for which we took a split of 60%, 20%, and 20%. The distance between a pair of images is computed using a distance metric derived from attributes. In this paper, we extend the LMNN classification method to exploit the prior information. Such data is ready for testing and learning with a kNN classifier that has a built-in PM.

IV. Results

In this experiment, we used filters from the Weka data mining tool and available Python implementations of included algorithms (sci-kit–learn for machine learning). The results in the tables below show the classifier's performance, which was evaluated with the train test split function. The variable k is the number of nearest neighbours searched by the linear nearest neighbour search algorithm in the kNN classifier.

The obtained results presented in Table II show that the algorithm's accuracy with Euclidean distance depends on the choice of filter and the selected value of the parameter k. The best results were obtained with the BPP filter.

Table III shows the accuracy of different filters and nearest neighbours for the Manhattan metric. The best results were obtained with the Phog filter and for lower k.

TABLE II: Results with Euclidean metric

k/filters	BPP	Edge	Phog	AutoCC	Colour
1	100%	100%	100%	99,76%	100%
2	100%	99,95%	100%	92,38%	99,52%
3	100%	99,95%	100%	93,85%	99,42%
4	100%	100%	100%	91,78%	99,52%
5	100%	99,95%	100%	91,43%	99,52%
6	100%	99,91%	100%	90,48%	99,27%
7	100%	99,95%	100%	90,07%	99,32%
8	100%	99,91%	100%	89,13%	98,84%
9	100%	99,86%	99,95%	89,36%	98,74%
10	100%	99,72%	99,91%	89,30%	98,55%
11	100%	99,68%	99,91%	88,89%	97,87%
12	100%	99,49%	99,86%	87,71%	97,43%
13	100%	99,44%	99,91%	87,41%	97,05%
14	100%	99,26%	99,95%	87,00%	96,95%
15	100%	99,26%	99,95%	86,64%	96,27%
16	100%	99,17%	99,91%	86,23%	96,18%
17	100%	99,98%	99,91%	86,23%	95,21%
18	100%	99,98%	99,91%	85,46%	95,01%
19	100%	98,89%	99,91%	85,05%	94,68%
20	100%	98,52%	99,86%	84,52%	94,58%

TABLE III: Results with different filters and Manhattan metric

k/filters	BPP	Edge	Phog	AutoCC	Colour
1	100%	100%	100%	99,82%	100%
3	100%	99,95%	100%	92,55%	99,37%
3	100%	99,91%	100%	92,67%	99,42%
4	100%	99,91%	100%	91,31%	99,13%
5	100%	100%	100%	91,13%	99,13%
6	100%	99,95%	100%	90,07%	98,98%
7	100%	99,91%	100%	89,95%	98,84%
8	99,95%	99,86%	100%	88,77%	98,55%
9	99,91%	99,95%	100%	88,83%	98,31%
10	99,91%	99,68%	100%	87,88%	98,16%
11	99,91%	99,63%	100%	87,23%	97,68%
12	99,91%	99,44%	99,95%	87,23%	97,19%
13	99,95%	74,40%	99,95%	86,11%	97,34%
14	99,95%	99,26%	99,95%	85,17%	96,95%
15	99,95%	99,26%	99,95%	85,28%	96,52%
16	99,95%	99,17%	99,95%	84,99%	96,08%
17	99,95%	99,03%	99,95%	84,46%	95,50%
18	99,95%	99,17%	99,86%	84,04%	95,06%
19	99,95%	98,98%	99,86%	83,81%	93,85%
20	99,95%	98,80%	99,86%	83,10%	94,00%

In the last table, Table IV below, we can see that the best results with PM were obtained with the Edge filter. Good results are also obtained for Colour filter and k less than 9. EM deviates with the BPP filter, Manhattan with the Phog filter, and PM with the Edge filter. Comparing the metrics with each other, on average EM returns the best results.

V. Conclusion

From the machine learning perspective, filters and used metric are essential for high-quality image classification. The first stage of this research deals with image classification on five benchmark image datasets developed in recent years.

TABLE IV: Results with different filters and PM

k/filters	BPP	Edge	Phog	AutoCC	Colour
1	100%	100%	6,76%	99,76%	100%
2	84,37%	99,54%	1,25%	92,14%	99,52%
3	81,31%	99,58%	1,30%	93,50%	99,23%
4	78,61%	99,35%	1,34%	91,37%	98,84%
5	77,32%	99,07%	1,30%	90,90%	98,64%
6	76,43%	98,66%	1,25%	90,37%	98,35%
7	75,58%	98,66%	3,94%	89,89%	98,35%
8	74,08%	98,33%	3,98%	89,60%	97,58%
9	73,23%	98,10%	2,87%	88,89%	97,29%
10	71,77%	97,78%	3,84%	88,77%	96,61%
11	71,49%	97,55%	3,98%	88,30%	96,27%
12	70,84%	97,36%	1,39%	87,53%	95,98%
13	70,03%	96,94%	3,33%	87,41%	95,55%
14	69,30%	96,67%	3,33%	86,76%	95,45%
15	68,45%	96,48%	3,33%	86,94%	95,16%
16	67,56%	96,06%	3,33%	86,05%	94,87%
17	66,71%	96,46%	3,33%	85,70%	94,14%
18	66,42%	95,05%	6,57%	85,11%	94,05%
19	65,37%	94,86%	6,06%	84,87%	93,22%
20	65,05%	94,49%	5,42%	83,92%	92,93%

We used LMNN and Poincaré metric, which prepared the image data to apply the kNN method. This method applies to both colour and greyscale images of faces. Compared EM and Manhattan metric we found found out that on average, there are no major deviations from the metric, except for the BPP, Edge, and Phog filters. On the other hand, PM get different results - the best results were with Edge, AutoCC and Colour.

This study is limited to images of faces. Future work includes testing this method on different types of images, such as pupils, tumours or other non-image datasets. Future work includes testing the performance of the proposed classifier based on the Poincaré metric when applying clustering methods. It can also be explored with other classification methods such as SVM, decision trees, neural networks, and other filters.

ACKNOWLEDGMENT

This work was supported by Slovenian Research Agency (ARRS project No. P1-0383) and Faculty of imformation studies in Novo mesto, Ljubljanska cesta 31 a, Novo mesto, Slovenia.

REFERENCES

[1] Bai, Y., Guo, L., Jin, L., and Huang, Q., A novel feature extraction method using Pyramid Histogram of Orientation Gradients for smile recognition. 2009 16th IEEE International Conference on Image Processing (ICIP), 3305-3308, 2009.

[2] Belhumeur, P. N., Hespanha, J. P., and Kriegman, D. J., Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. Lecture Notes in Computer Science, Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, 1064(7), 45–58, 1996.

[3] Bi, Y., Fan, B., and Wu, F., Multiple Cayley-Klein metric learning. 1–15,1999, 2017.

[4] Bohr, A., and Memarzadeh, K., The rise of artificial intelligence in healthcare applications. Artificial Intelligence in Healthcare, 25–60, 2020.

[5] Dino, H. I. and Abdulrazzaq, M. B., Facial Expression Classification Based on SVM, KNN and MLP Classifiers. 2019 International Conference on Advanced Science and Engineering, ICOASE 2019, 70–75, 2019.

[6] Fujiyoshi, H., Hirakawa, T. and Yamashita, T., Deep learning-based image recognition for autonomous driving. IATSS Research, 43(4), 244–252, 2019.

[7] Greche, L., Jazouli, M., Es-Sbai, N., Majda, A. and Zarghili, A., Comparison between Euclidean and Manhattan distance measure for facial expressions classification. 2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems, WITS 2017, 2–5, 2017.

[8] Huang, J., Kumar, S. R., Mitra, M., Zhu, W. J., and Zabih, R., Image indexing using color correlograms. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 762–768, 1997.

[9] Khrulkov, V., Mirvakhabova, L., Ustinova, E., Oseledets, I. and Lempitsky, V., Hyperbolic image embeddings. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 6418–6428, 2020

[10] Kortli, Y., Jridi, M., Falou, A. A., and Atri, M. (2020). Face Recognition Systems: A Survey. Sensors (Basel, Switzerland), 20(2), 342, 2020.

[11] Krantz, G. S.: Complex analysis: the geometric viewpoint. The mathematical association of America, USA, 2004.

[12] Han, J., Kamber, M., and Pei, J., Data Mining: Concepts and Techniques, 2012.

[13] Henaff, O. J., Srinivas, A., Fauw, J. De, Razavi, A., Doersch, C., Eslami, S. M. A., and Eslami, A. V. O., Data-Efficient image recognition with contrastive predictive coding. 37th International Conference on Machine Learning, ICML 2020, 4130–4140, 2020.

[14] Nielsen, F., Hadjeres, G., Approximating covering and minimum enclosing balls in hyperbolic geometry. An information geometry problem in mathematical finance, pp. 586–594, 2016.

[15] Lyons, M., Akamatsu, S., Kamachi, M., and Gyoba, J., Coding facial expressions with Gabor wavelets. Proceedings - 3rd IEEE International Conference on Automatic Face and Gesture Recognition, FG 1998, May, 200–205, 1998.

[16] Popova, A., and Neshov, N. N., Image Similarity Search Approach Based On The Best Features Ranking. 37(1), 51–65, 2013.

[17] Sadique, M. F., and Haque, S. M. R., Content-Based Image Retrieval Using Color Layout Descriptor, Gray-Level Co-Occurrence Matrix and K-Nearest Neighbors. International Journal of Information Technology and Computer Science, 12(3), 19–25, 2020.

[18] Spivak, I., Krepych, S., Faifura, V. and Spivak, S., Methods and Tools of Face Recognition for the Marketing Decision Making, IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology, pp. 212-216, 2019.

[19] Park, J., Cho, J., Chang, H. J., and Choi, J. Y., Unsupervised hyperbolic representation learning via message passing auto-encoders. In 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition.). IEEE. 2020.

[20] Pietikäinen, M., Hadid, A., Zhao, G., and Ahonen, T., Local Binary Patterns for Still Images, 2011.

[21] Thorgeirsson, S., Hyperbolic geometry: history, models, and axioms, 2014.

[22] Trpin, A., Boshkoska, B.M., Boškoski, P., Poincaré Metric in Algorithms for Data Mining Tools. Beyond Databases, Architectures and Structures. Paving the Road to Smart Data Processing and Analysis. BDAS 2019. Communications in Computer and Information Science, vol. 1018. Springer, 2019.

[23] Vishnuvardhan, G., Ravi, V. (2021). Face Recognition Using Transfer Learning on Facenet: Application to Banking Operations. In: Gunjan, V.K., Zurada, J.M. (eds) Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough. Studies in Computational Intelligence, vol 956. Springer, Cham.

[24] Wang, M., and Deng, W., Deep face recognition: A survey. Neurocomputing, 429, 215–244, 2021.

[25] Weinberger, K. Q., and Saul, L. K., Distance metric learning for large margin nearest neighbor classification. Journal of Machine Learning Research, 10, 207–244, 2009.

[26] Weyrauch, B., Heisele, B., Huang, J., and Blanz, V., Component-based face recognition with 3D morphable models. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2004.

[27] Yue, M., Face recognition. Proceedings - IEEE International Symposium on Circuits and Systems, 2000.