

FeatSet: A Compilation of Visual Features Extracted from Public Image Datasets

**Mirela T. Cazzolato¹, Lucas C. Scabora¹, Guilherme F. Zabet¹,
Marco A. Gutierrez², Caetano Traina Jr.¹, Agma J. M. Traina¹**

¹ Institute of Mathematics and Computer Sciences
University of São Paulo (ICMC-USP), São Carlos, Brazil

²Heart Institute (InCor) – Clinical Hospital of Faculty of Medicine
University of São Paulo (HC-FMUSP), São Paulo, Brazil

mirelac@usp.br, agma@icmc.usp.br

Abstract. *In this paper, we present FeatSet, a compilation of visual features extracted from open image datasets reported in the literature. FeatSet has a collection of 11 visual features, consisting of color, texture, and shape representations of the images acquired from 13 datasets. We organized the available features in a standard collection, including the available metadata and labels, when available. We also provide a description of the domain of each dataset included in our collection, with visual analysis using Multidimensional Scaling (MDS) and Principal Components Analysis (PCA) methods. FeatSet is recommended for supervised and non-supervised learning, also widely supporting Content-Based Image Retrieval (CBIR) applications and complex data indexing using Metric Access Methods (MAMs).*

1. Introduction

Images acquired from different application scenarios have been the focus of numerous studies in the last decades. Users generate large amounts of images, sharing them over the internet, mainly due to the widespread use of social networks, blogs, public repositories, and cooperative research. Mobile cameras, microscopes, Magnetic Resonance Imaging (MRI) machines, and X-Rays are just a few examples of simple and specialized acquisition equipment capable of generating images in different contexts. Data management systems usually compare scalar data, such as small strings, numbers, and dates, based on order and equality operators (*i.e.*, $<$, \leq , $=$, \geq , and $>$). On the other hand, images are considered complex data that order operators cannot compare, and checking if a pair of images is equal or different brings too little semantics to the analysis. Accordingly, complex objects are usually compared by their content, using similarity operators such as the *Range* and *k*NN queries for Content-Based Retrieval (CBR).

Feature Extraction Methods (FEMs) generate feature vectors as low-level representations of images according to their visual content. The generated features can describe, among others, the color distribution of an image, the grayscale variation, the texture patterns, and the objects' salient edges. Similarity operators compare pairs of complex objects by employing distance functions to the corresponding feature vectors, measuring how dissimilar they are. Also, machine learning algorithms employ feature vectors to train their model in supervised scenarios (for instance, when we have a label

for every image in the dataset) and in exploratory analyses. All of these analyses depend on extracting the features from the image datasets, which can be laborious due to the necessity of implementing and/or setting up FEMs, and also time-consuming.

Many works from the literature have employed visual features for image analysis in different contexts. For instance, in the work [de Sousa Fogaça and Bueno 2020], the authors mapped color-based features into the multidimensional space to estimate the trajectory of objects by simulating their evolution over time. In [Pereira and Ribeiro 2021], the authors employed visual features extracted from mammograms for the semantic annotation and classification of images using ontologies. Low-level features have been widely employed to validate the indexing capabilities of Metric Access Methods [Zabot et al. 2019b, Moriyama et al. 2021]. Also, in [Maheshwari et al. 2021] the authors employed various features to identify COVID-19 in images.

Motivated by the potential applicability of image features and the difficulties of employing FEMs, we propose the *FeatSet* dataset in this work. *FeatSet* is a compilation of widely-used visual features extracted from public image datasets of different application scenarios. The contributions of *FeatSet* are two-fold:

- The curation of 13 open image databases, organizing their main information in a single repository;
- Color, texture, and shape visual features extracted from the images with 11 distinct FEMs, widely employed in the literature, including those from the MPEG7 Standard [Manjunath et al. 2002]. The feature vectors are organized into a standard model and openly available.

We cast the possibilities of use for *FeatSet*, and show examples of analysis of the available data.

Previous use of data. A small part of *FeatSet* has been employed in the previous studies [Zabot et al. 2019a, Zabot et al. 2019b]. In that works, the authors explored different visual features to validate a novel Multi-Metric Access Method, aiming at indexing complex objects based on images’ visual characteristics and the correlation among the distance spaces.

In this work, we present an extended and complete version of the data used in [Zabot et al. 2019a, Zabot et al. 2019b]. *FeatSet* is a new dataset, composed of diverse visual features extracted from various public image datasets of different application scenarios. It allows analysts to deeper evaluating machine learning approaches, CBIR strategies, and related techniques.

Paper outline. The remaining sections of this paper are organized as follows. Section 2 describes *FeatSet*. Section 3 discusses application scenarios and challenges for *FeatSet*. Section 4 details the steps to download *FeatSet*, and describes the data organization and description of the dataset’s public repository. Finally, Section 5 concludes this work.

2. *FeatSet*: A Collection of Visual Features from Image Datasets

In this section, we detail the process of acquiring the original images, extracting the visual features, and composing *FeatSet*, as Figure 1 illustrates. Firstly (Step i), we looked for open image datasets of different sizes, with photos acquired from different applica-

tion scenarios. We looked for literature papers proposing or using image data from open repositories and websites for this task.

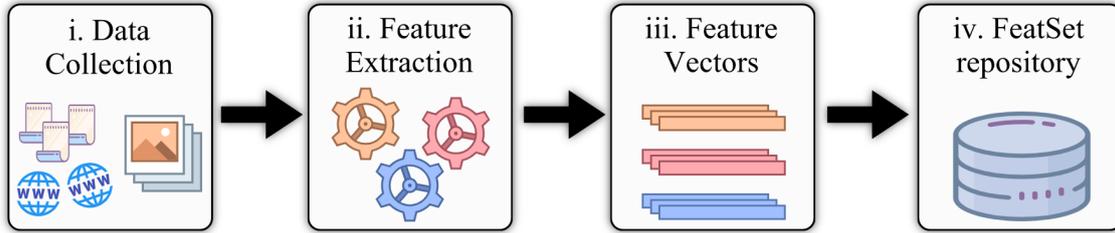


Figure 1. Steps carried to compose *FeatSet*.

2.1. Data Collection and Preprocessing

Table 1 lists the datasets collected, as well as the specific reference, the number of available images, and a brief description of each. All 13 datasets were manually collected from their original sources, which are referenced in the corresponding folders in the repository, with a “*read me*” file containing the original source, URL, date of collection, license, a brief description, and the corresponding reference. In particular, the *ds-MNIST* dataset provides the images as multidimensional matrices. We converted each file to *.png*, which is one of the image formats accepted by the implementation of FEMs employed in this work. Figure 2 shows examples of images acquired from each dataset.

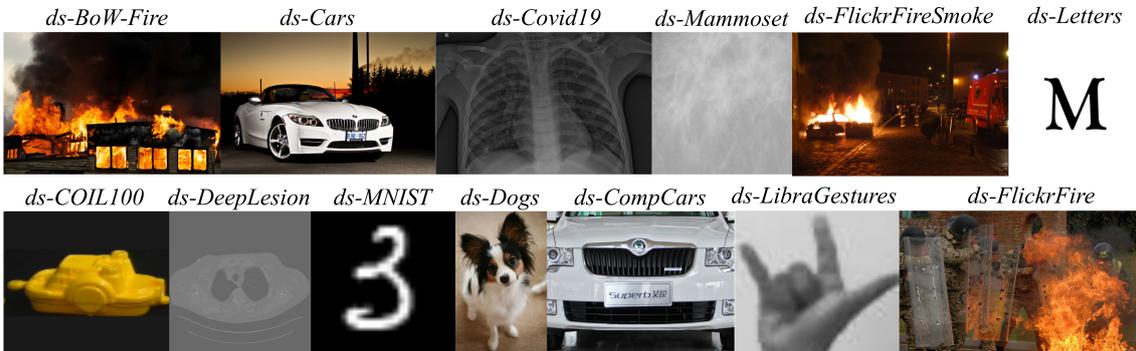


Figure 2. Examples of images from the public datasets considered.

2.2. Feature Extraction

After curating every dataset, organizing the files and available metadata, we employed feature extraction methods (FEMs) to obtain the visual features from the acquired images. Figure 1 (Steps ii and iii) illustrates this task. Each FEM receives as input an image file and generates a d -dimensional vector, where the features are represented as an array of floats.

We employed the FEMs listed in Table 2 with the corresponding acronyms, number of dimensions and types. Notice that every employed FEM generates a specific number of features (dimensions), and corresponds to a specific type T of visual feature, where $T \in \{Color, Texture, Shape\}$. In the case of NCH, we generated six histogram variations, with 8, 16, 32, 64, 128, and 256 features. As a result, we have 11 distinct FEMs, which can generate 16 different feature configurations.

Table 1. Datasets composing *FeatSet*.

Dataset	#Images	Application Scenario
<i>ds-BoWFire</i> [Chino et al. 2015]	226	Depicts images of fire incidents in emergency situations, labeled fire and non-fire.
<i>ds-Flickr-Fire</i> [Bedo et al. 2015]	1,984	Acquired from Flickr, using tags related to fire to filter the information.
<i>ds-Mammoset</i> [Oliveira et al. 2017]	3,457	Regions of Interest obtained from mammograms, with tissue labels such as benign and malignant.
<i>ds-LibraGestures</i> [Bastos et al. 2015]	4,800	Depicts images of hand gestures representing the Brazilian Sign Language (Libras).
<i>ds-FlickrFireSmoke</i> [Cazzolato et al. 2017]	5,556	Acquired from Flickr, using tags related to fire and smoke to filter the information.
<i>ds-Covid19</i> [Cohen et al. 2020]	5,933	Chest X-Rays and Computed Tomographies, taken from patients which are positive or suspected of COVID-19 or other viral and bacterial pneumonias.
<i>ds-COIL100</i> [Nene et al. 2020]	7,200	Objects depicted at angles in a 360 rotation, at every 5 degrees.
<i>ds-Letters</i> [Hajder 2020]	15,340	Standard Windows fonts with each letters organized in classes by typeface.
<i>ds-Cars</i> [Krause et al. 2013]	16,185	Images of cars from 196 classes, including annotations.
<i>ds-Dogs</i> [Khosla et al. 2011]	20,580	Images of dogs of 120 breeds from around the world.
<i>ds-DeepLesion</i> [Yan et al. 2017]	33,334	Images Slices extracted from Computed Tomographies.
<i>ds-MNIST</i> [Lecun et al. 1998]	70,000	Images of 10 handwritten digits (0 to 9).
<i>ds-CompCars</i> [Yang et al. 2015]	164,344	Depicts images of cars, taken from two scenarios: web-nature and auto parts.

The implementations of FEMs used in this work are from the Arboretum library, available at the Database and Image Group (GBDI) website¹. Most of the Arboretum’s available FEMs are from the MPEG-7 Standard [Manjunath et al. 2002], proposed by ISO/IEC JTC1, which aims to define an efficient way to search, filter, and identify multimedia content, defining the expected representations for the images in terms of color, texture, and shape. In this article, we employ the following MPEG-7 extractors: Color Layout, Color Structure, Scalable Color, Color Temperature, Edge Histogram, and Texture Browsing. Briefly, they work as the following:

- **Color Layout:** Describes the image color distribution considering spatial location [Kasutani and Yamada 2001]. It splits the image into squared sub-regions and labels each square with a few nonlinear quantized DCT coefficients of grid-based average colors.

¹The Arboretum library is available at <https://gbdi.icmc.usp.br/>, under the “Projects” menu.

Table 2. Feature Extraction Methods FEMs employed, the corresponding acronyms, dimensionality and feature type.

FEM	Acronym	#Dimensions	Type
Color Temperature	CT	3	Color
Texture Spectrum	TS	8	Texture
Color Layout	CL	16	Color
Haralick	Hr	24	Texture
Color Structure	CS	128	Color
Edge Histogram	EH	150	Shape
Local Binary Pattern	LBP	177	Texture
Scalable Color	SC	256	Color
BIC Histogram	BIC	512	Color
Total Color Histogram	TCH	768	Color
Normalized Histogram	NCH	8, 16, 32, 64, 128, 256	Color

- **Scalable Color:** A color histogram in the HSV color space, which is encoded by a Haar transform and is intended to capture the prominent color distribution [Manjunath et al. 2001].
- **Color Structure:** Aims to capture both the color content and information about the spatial arrangement of that color content [Sikora 2001]. It works such a histogram that counts how many times a color is present in structures with fixed-size windows. Each fixed-size window selects equally spaced pixels to represent the local color structure. The window size and the number of local structures are parameters of the Color Structure descriptor.
- **Edge Histogram:** Represents the spatial distribution of five types of edges, Vertical, Horizontal, 45 degree, 135 degree and non-directional, regarding $N \times N$ blocks, where N is an extractor parameter [Park et al. 2000]. Each block is constructed by partitioning the original image into squared regions and consists of local histograms of these edge directions, which may optionally be aggregated into global or semi-global histograms.
- **Texture Browsing:** This extractor is obtained from the same parameters used in the Gabor filters applied to the images [Lee and Chen 2005]. Its vector consists of 12 positions: 2 for regularity, 6 for directionality, and 4 for coarseness.
- **Color Temperature:** Is based on the hypothesis that there is a correlation between the illumination properties of the image and its “feeling of temperature”. CT represents the feature vector as the linearized pixels in the XYZ color space, discarding the luminance of Y channel that is above a given threshold parameter. CT averages the color coordinate in XYZ and converts it to UCS. Finally, CT calculates the two-color isotherm lines from the color diagrams [Bedo et al. 2015].

BIC (Border/Interior Pixel Classification) [Stehling et al. 2002], TCH, and Normalized histograms describe the grayscale color distribution of the pixels in the image. Haralick is texture FEM that computes the image dimensions as variances and moments based on co-occurrence matrices. Texture Spectrum and LBP describes the local correlation among grayscale values within pixels.

2.3. Data Description

After curating the public datasets and extracting the visual features, we organized the *FeatSet* repository (Step iv of Figure 1). Figure 3 shows the generic *FeatSet* schema, which is similar to each one of the 13 datasets. Figure 3(a) is the metadata table, with the object identifier (OID) for each complex object in the dataset, employed as primary key (PK), the filename of the image, and the set of classes if any. The image filename is the same as the original dataset, which allows reproducibility. Table 3 details the existing set of classes on *FeatSet*. The datasets *ds-BoWFire*, *ds-Flickr-Fire*, *ds-LibraGestures*, *ds-Letters*, *ds-Dogs*, and *ds-MNIST* have a single set of classes, represented by column `class` in the metadata table. Dataset *ds-FlickrFireSmoke* has two sets of classes, one to denote the presence or absence of fire in the image (column `class_0`), and the other to determine if the image has or not smoke in it (column `class_1`). The remaining datasets do not have classes. Figure 3(b) illustrates the set of FEM tables originated from Section 2.2. Each FEM table has the OID column as a foreign key (FK) to the respective metadata table, and every dimension of the feature vector is stored in a column named `feature_i`, for $0 \leq i \leq d$, where d is the number of dimensions of the current FEM, given by Table 2.

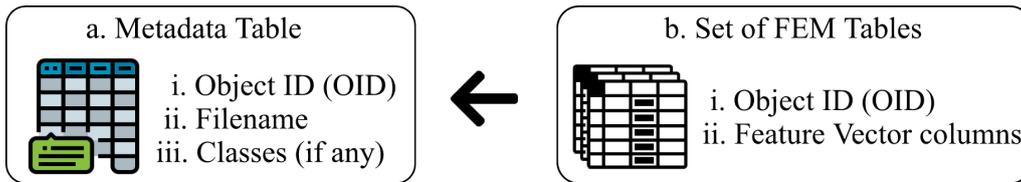


Figure 3. Schema for *FeatSet*.

Table 3. Existing set of labels (classes) in *FeatSet*.

Dataset	Set of Labels (Classes)
<i>ds-BoWFire</i> [Chino et al. 2015]	Presence/Absence of Fire
<i>ds-Flickr-Fire</i> [Bedo et al. 2015]	Presence/Absence of Flame
<i>ds-LibraGestures</i> [Bastos et al. 2015]	Gesture translation
<i>ds-FlickrFireSmoke</i> [Cazzolato et al. 2017]	Presence of Fire and/or Smoke
<i>ds-Letters</i> [Hajder 2020]	Letter's Font
<i>ds-Dogs</i> [Khosla et al. 2011]	Dog's breed
<i>ds-MNIST</i> [Lecun et al. 1998]	A digit from 0 to 9

Both metadata and FEM tables in *FeatSet* are organized in separated Structured Query Language (SQL) scripts to create and populate those tables. We decided to maintain the separated tables for each dataset within *FeatSet* because this organization allows users to select only the scripts from the data they want to work with. We also provide Comma-Separated Values (CSV) files for every dataset and FEM, allowing users to input the data into machine learning libraries, managing the features outside the Database Management System (DBMS), and also concatenate the desired files whenever necessary.

3. Applicability and Challenges for *FeatSet*

FeatSet opens research opportunities regarding various computer science tasks, such as evaluating CBIR systems, machine learning methods, data visualization, and information

prediction. As *FeatSet* constitutes a compilation of public datasets acquired from various application contexts, the visual features can also be employed in multidisciplinary studies. In this section, we discuss potential research opportunities.

3.1. Visual Analysis

The Multidimensional Scaling (MDS) method represents the (dis)similarity among objects onto a projection of the data in a low-dimensional space [Borg and Groenen 2005]. Here, we employed MDS to show the advantage of representing the image datasets using diverse visual representations. Figure 4 shows the distance space distribution formed by an image sample from *ds-BoWFire*, using the different visual features provided by *FeatSet*. We observe major differences in the data dispersion from the generated distance spaces. For instance, CL shows objects dispersed almost homogeneously, while CT depicts the objects in a “line-shaped” dispersion.

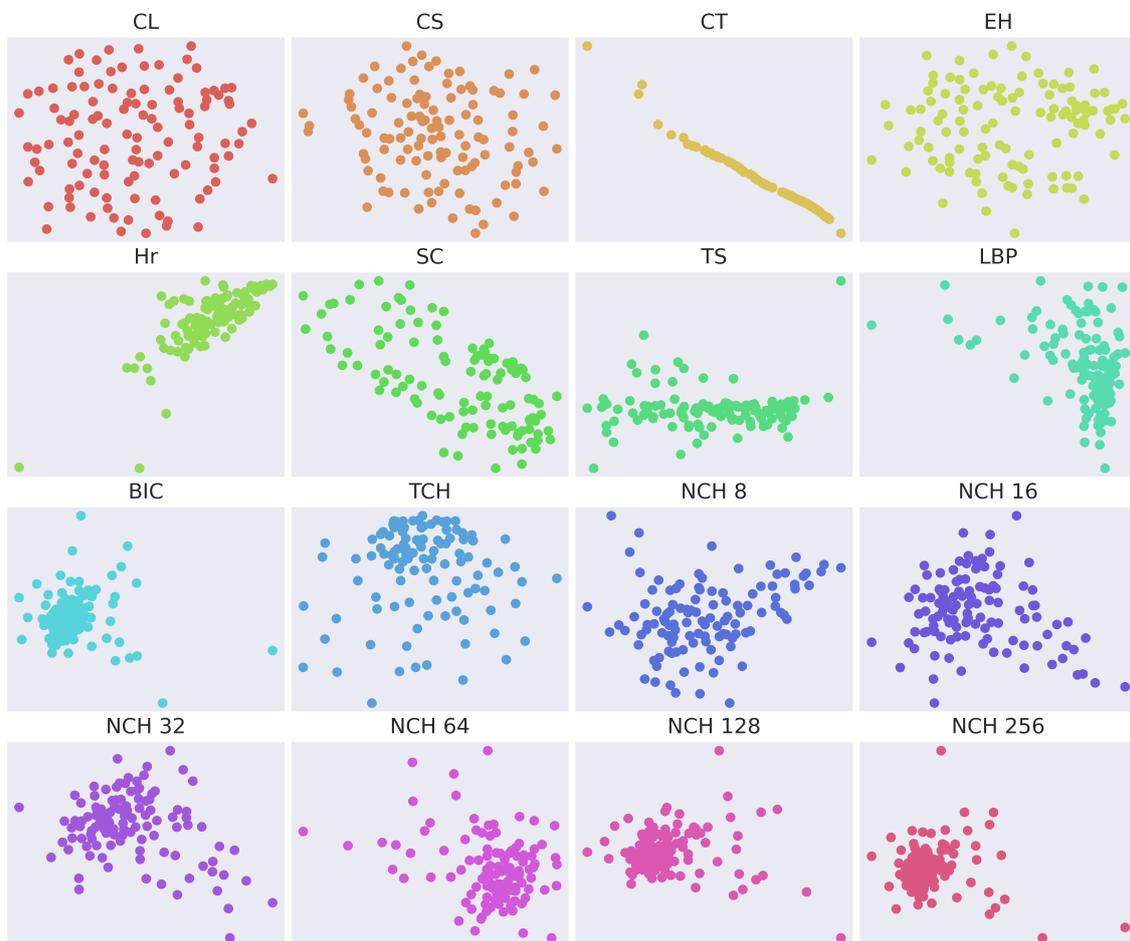


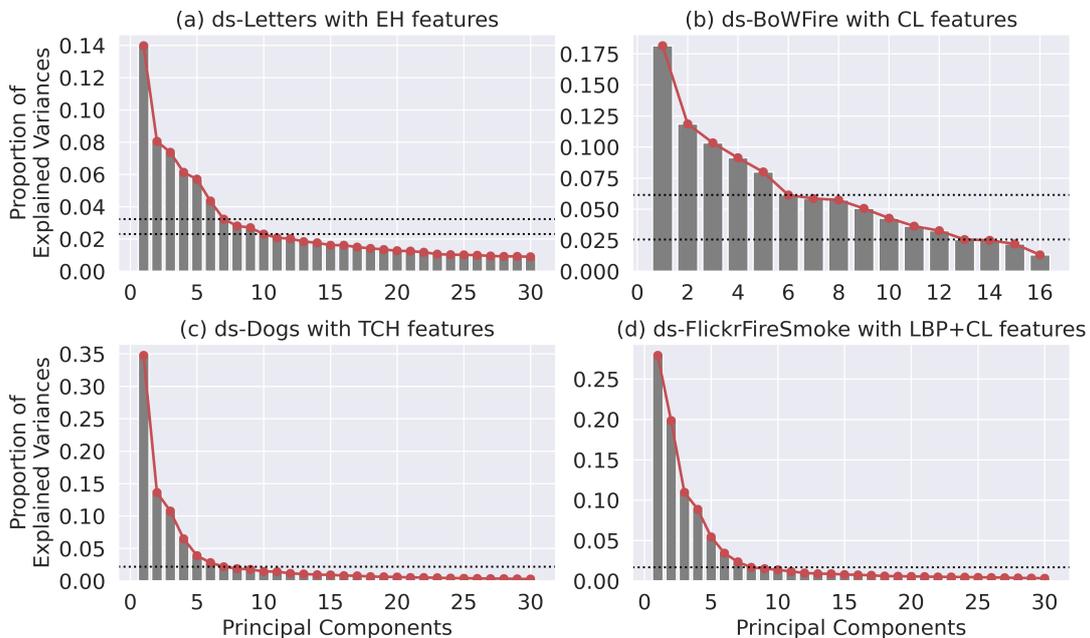
Figure 4. MDS plot depicts the objects’ dispersion within the diverse distance space distribution of *ds-BoWFire*, generated with different visual features.

3.2. Principal Components Analysis

Many of the employed FEMs, such as BIC, TCH, and NCH, produce high-dimensional feature vectors (see Table 2). Also, in many application scenarios, the analyst may opt to combine features of different visual characteristics to improve the semantics of the image

representation. For example, if we consider *ds-Mammoset*, microcalcifications can show different color and texture patterns that, when combined, allow a more profound pattern recognition. Feature concatenation can improve data representation but has the cost of increasing the data dimensionality, which can be approached in different manners. One example is the application of the Principal Component Analysis (PCA) to reduce data dimensionality.

Figure 5 shows the proportion of explained variances according to the principal components generated by PCA. In the examples, we selected four datasets from *FeatSet*, and explored three single feature representations (a, b, and c) and the combination of LBP with CL (d). To improve the visualization, we plotted a maximum of 30 principal components. The dashed horizontal lines represent curve elbows, visually observed in the plots. The curve elbow can be used as a heuristic to select the number of principal components to employ in data analysis, selecting the position where the curve stops decreasing and flattens out. Using this criterion, in (a) we could use 7 or 10 principal components, in (b) 6 or 13, in (c) 7, and in (d) 8 principal components. One can also consider the dimensions which sum is at least a threshold, for example, 70% of the entire variance. In this case, in plot (a) we would select the first 18 principal components, in (b) the first 8, in (c) the first 6, and in (d) the first 5 principal components. Regardless of the employed heuristic, the selected visual features can be used to further analyze the complex objects by employing CBIR or machine learning methods. Although PCA was used here as a dimensionality reduction technique when concatenating feature vectors, it is also an example of how feature analysis layers can be stacked to develop deep learning models for image feature engineering.



The dashed horizontal lines indicates the visual curve elbows.

Figure 5. The scree plots show the proportion of explainable variances according to the principal components generated by PCA.

3.3. Challenges and Opportunities of Analysis

FeatSet comprises small and large datasets (ranging from 226 to 164,344 objects), and FEMs of low and high dimensionality (from 3 to 768 dimensions). The variation in size and dimensionality can support the validation of techniques focused on content-based retrieval and diversity, complex data indexing with metric access methods, and similar methods. Similarity-based comparisons of complex objects can also support the identification of near-duplicate images through feature-matching.

The dataset organization allows straightforwardly employing machine learning methods since they already are in the input format of many existing analysis libraries, such as Scikit Learn² for Python. Users can perform classification and clustering methods, compute correlations among the different data representations, perform object recognition, among others. *FeatSet*'s schema also allows users to work with the visual features inside the Database Management System (DBMS) by loading the available database files provided in the Git repository.

FeatSet can be further extended by extracting new visual features with other FEMs reported in the literature. Finally, users can include new image datasets into *FeatSet*, extracting their features using the FEMs reported in this work, which are openly available at the Arboretum library.

4. Public Repository and Citation Request

FeatSet is publicly available for research use under the Creative Commons license. All data and additional information are organized in the Git repository <https://github.com/mtcazzolato/featset/>. The repository is organized as follows:

- *FeatSet/*
 - *README.md*: Read me file with *FeatSet* description, citation instructions for every part of the dataset, and other information.
 - *SQL-Scripts-Link*: Link to download the SQL scripts used to load the data.
 - *CSV-File-Link*: Link to download the CSV files with the data.
 - *python-scripts/*: Folder with the two python scripts that generate the plots depicted in Figure 4 and Figure 5.
 - *schema.png*: The database schema.

Each one of the 13 datasets presented in Table 1 follows the schema illustrated on Figure 3, where its metadata is stored in an SQL file with the dataset name (e.g., *ds-CompCars.sql*), and every FEM is inside of a file with the same name plus its respective acronym as a suffix (e.g., *ds-CompCars-CL.sql* for Color Layout). All SQL scripts start with the CREATE TABLE statement, followed by the INSERT INTO statements to populate those tables.

Furthermore, *FeatSet*'s repository also provides a Comma-Separated Values (CSV) file alternative for every single table from the 13 datasets. *FeatSet* is available for researchers and data scientists, and, in case of publication or any work derived from any of the datasets from *FeatSet*, we ask to acknowledge the image dataset owners and us by citing both the image dataset reference and this paper, following the instructions of the provided README .md file.

²The Scikit Learn Python library: <https://scikit-learn.org/stable/>

5. Conclusion

In this work, we proposed the *FeatSet* dataset, a compilation of visual features extracted from public image datasets of different application scenarios. *FeatSet* is composed of 13 datasets and has 11 visual features representing the images. We provided examples of how the feature vectors can be explored for different tasks. Also, the public datasets inside *FeatSet* are from diverse application domains, which can help analysts evaluate their techniques' semantics in a wide range of examples. *FeatSet* is organized in a public repository and is available in SQL scripts to load the database in a DBMS and in CSV files to be used directly in existing data analysis libraries.

Acknowledgment

This research was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, by the São Paulo Research Foundation (FAPESP, grants No. 2016/17078-0, 2020/07200-9, 2020/11258-2), and the National Council for Scientific and Technological Development (CNPq).

References

- Bastos, I. L. O., Angelo, M. F., and Loula, A. C. (2015). Recognition of static gestures applied to brazilian sign language (libras). In *28th SIBGRAPI*.
- Bedo, M. V. N., Blanco, G., Oliveira, W. D., Cazzolato, M. T., Costa, A. F., Rodrigues-Jr., J. F., Traina, A. J. M., and Traina Jr., C. (2015). Techniques for effective and efficient fire detection from social media images. In Hammoudi, S., Maciaszek, L. A., and Teniente, E., editors, *ICEIS 2015 - Proceedings of the 17th International Conference on Enterprise Information Systems, Volume 1, Barcelona, Spain, 27-30 April, 2015*, pages 34–45. SciTePress.
- Borg, I. and Groenen, P. (2005). *Modern Multidimensional Scaling: Theory and Applications*. Springer Series in Statistics. Springer New York.
- Cazzolato, M. T., Avalhais, L. P. S., Chino, D. Y. T., Ramos, J. S., Souza, J. A., Rodrigues-Jr, J. F., and Traina, A. J. M. (2017). Fismo: A compilation of datasets from emergency situations for fire and smoke analysis. In *SBB2017 - SBB2 Proceedings of Satellite Events of the 32nd Brazilian Symposium on Databases - DSW (Dataset Showcase Workshop)*, pages 213–223. SBC.
- Chino, D. Y. T., Avalhais, L. P. S., Rodrigues-Jr., J. F., and Traina, A. J. M. (2015). Bowfire: Detection of fire in still images by integrating pixel color and texture analysis. In *28th SIBGRAPI Conference on Graphics, Patterns and Images, SIBGRAPI 2015, Salvador, Bahia, Brazil, August 26-29, 2015*, pages 95–102. IEEE Computer Society.
- Cohen, J. P., Morrison, P., Dao, L., Roth, K., Duong, T. Q., and Ghassemi, M. (2020). Covid-19 image data collection: Prospective predictions are the future. *CoRR*, abs/2006.11988.
- de Sousa Fogaça, I. C. O. and Bueno, R. (2020). Temporal evolution of complex data. In *Anais do XXXV Simpósio Brasileiro de Bancos de Dados, SBB2020, online, September 28 - October 1, 2020*, pages 25–36. SBC.
- Hajder, S. (2020). *Letters organized by typefaces*. Last accessed in October, 2020.

- Kasutani, E. and Yamada, A. (2001). The mpeg-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval. In *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, volume 1, pages 674–677 vol.1.
- Khosla, A., Jayadevaprakash, N., Yao, B., and Fei-Fei, L. (2011). Novel dataset for fine-grained image categorization. In *First Workshop on Fine-Grained Visual Categorization, IEEE Conference on Computer Vision and Pattern Recognition*, Colorado Springs, CO.
- Krause, J., Stark, M., Deng, J., and Fei-Fei, L. (2013). 3d object representations for fine-grained categorization. In *2013 IEEE International Conference on Computer Vision Workshops, ICCV Workshops 2013, Sydney, Australia, December 1-8, 2013*, pages 554–561. IEEE Computer Society.
- Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.
- Lee, K.-L. and Chen, L.-H. (2005). An efficient computation method for the texture browsing descriptor of mpeg-7. *Image and Vision Computing*, 23:479–489.
- Maheshwari, S., Sharma, R. R., and Kumar, M. (2021). Lbp-based information assisted intelligent system for COVID-19 identification. *Comput. Biol. Medicine*, 134:104453.
- Manjunath, B., Ohm, J., Vasudevan, V., and Yamada, A. (2001). Color and texture descriptors. *Circuits and Systems for Video Technology, IEEE Transactions on*, 11:703 – 715.
- Manjunath, B. S., Salembier, P., and Sikora, T. (2002). *Introduction to MPEG-7: multimedia content description interface*. John Wiley & Sons.
- Moriyama, A., Rodrigues, L. S., Scabora, L. C., Cazzolato, M. T., Traina, A. J. M., and Traina, C. (2021). Vd-tree: how to build an efficient and fit metric access method using voronoi diagrams. In Hung, C., Hong, J., Bechini, A., and Song, E., editors, *SAC '21: The 36th ACM/SIGAPP Symposium on Applied Computing, Virtual Event, Republic of Korea, March 22-26, 2021*, pages 327–335. ACM.
- Nene, S. A., Nayar, S. K., and Murase, H. (2020). Columbia object image library (coil-100). Technical report, Technical Report CUCS-006-96. Last accessed in October, 2020.
- Oliveira, P. H., Scabora, L. C., Cazzolato, M. T., Bedo, M. V. N., Traina, A. J. M., and Traina-Jr., C. (2017). MAMMOSET: An Enhanced Dataset of Mammograms. In *Proceedings of the Satellite Events of the 32nd Brazilian Symposium on Databases*, pages 256–266. SBC.
- Park, D. K., Jeon, Y. S., and Won, C. S. (2000). Efficient use of local edge histogram descriptor. In Ghandeharizadeh, S., Chang, S., Fischer, S., Konstan, J. A., and Nahrstedt, K., editors, *Proceedings of the ACM Multimedia 2000 Workshops, Los Angeles, CA, USA, October 30 - November 3, 2000*, pages 51–54. ACM Press.
- Pereira, J. W. and Ribeiro, M. X. (2021). Semantic annotation and classification of mammography images using ontologies. In Almeida, J. R., González, A. R., Shen, L., Kane, B., Traina, A., Soda, P., and Oliveira, J. L., editors, *34th IEEE International*

- Symposium on Computer-Based Medical Systems, CBMS 2021, Aveiro, Portugal, June 7-9, 2021*, pages 378–383. IEEE.
- Sikora, T. (2001). The mpeg-7 visual standard for content description-an overview. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(6):696–702.
- Stehling, R. O., Nascimento, M. A., and Falcão, A. X. (2002). A compact and efficient image retrieval approach based on border/interior pixel classification. In *Proceedings of the 2002 ACM CIKM International Conference on Information and Knowledge Management, McLean, VA, USA, November 4-9, 2002*, pages 102–109. ACM.
- Yan, K., Wang, X., Lu, L., and Summers, R. M. (2017). Deeplesion: Automated deep mining, categorization and detection of significant radiology image findings using large-scale clinical lesion annotations. *CoRR*, abs/1710.01766.
- Yang, L., Luo, P., Loy, C. C., and Tang, X. (2015). A large-scale car dataset for fine-grained categorization and verification. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015*, pages 3973–3981. IEEE Computer Society.
- Zabot, G. F., Cazzolato, M. T., Scabora, L. C., Façal, B. S., Traina, A. J. M., and Traina Jr., C. (2019a). UCORM: indexing uncorrelated metric spaces for concise content-based retrieval of medical images. In *32nd IEEE International Symposium on Computer-Based Medical Systems, CBMS 2019, Cordoba, Spain, June 5-7, 2019*, pages 306–311. IEEE.
- Zabot, G. F., Cazzolato, M. T., Scabora, L. C., Traina, A. J. M., and Traina Jr., C. (2019b). Efficient indexing of multiple metric spaces with spectra. In *IEEE International Symposium on Multimedia, ISM 2019, San Diego, CA, USA, December 9-11, 2019*, pages 169–176. IEEE.