# Readme for the Corel Dataset

Mete Ozay, *Student Member, IEEE*, Fatos T. Yarman Vural, *Member, IEEE*

## 1 How to cite the dataset and the paper?

If you use this database in your own research, please cite the following paper:
@techreport{FSG,
author = {Ozay, Mete and Vural, Fatos T Yarman},
title = {A New Fuzzy Stacked Generalization Technique for Deep learning and Analysis of its Performance},
number = {arXiv:1204.0171},
month = {Apr},
year = {2012},
note = {Comments: Submitted to IEEE TNN}
}

## 2 Properties of the Corel Dataset and Feature Sets

In this document, properties of the Haar and MPEG-7 features extracted from the samples in Corel Dataset, that are used in the paper "A New Fuzzy Stacked Generalization Technique and Analysis of its Performance" [1], are given.

The dataset is available on the following website:
https://github.com/meteozay/Corel$_Dataset$

In the experiments, we have used a subset of the benchmark Corel Dataset consisting of $599$ classes each of which contains $97-100$ samples. To be precise, $238^{th}$ class contains $97$ samples, $342^{th}$ and $376^{th}$ classes contain $99$ samples, and rest of the classes contain $100$ samples. The employed Haar and 7 of MPEG-7 visual features (descriptors) [2], [3] are summarized as follows

- **Color Structure** (Feature Space Dimension $D_j=32$) gives information on the spatial statistics and structural properties of colors in the image. A color structure feature is a vector of eight bit-quantized histogram values $\overline{hist}_m(q)$, where $m$ is the scale of a square structuring element and $q \in \{1,2,\ldots,Q\}$, where $Q \in \{32,64,128,256\}$. For instance, a feature extracted using $m=8$ gives information about the number of times a color is present in an $8\times 8$ windowed neighborhood which is traversed over the image. Moreover, if $Q=32$, then each variable of the vector represents the relative spatial frequency of each quantized color value in the image.
- **Color Layout** (Feature Space Dimension $D_j=12$) captures the spatial arrangement of colors in an image, similar to Color Structure, but the features are invariant to scale unlike Color Structure. In order to achieve scale invariance, an image is partitioned into $8\times 8=64$ blocks. Then, a representative color is selected from each block by using a pooling method such as average pooling which computes the average of the pixel colors in a block. Three pooled

- *M. Ozay and F. T. Yarman Vural are with the Department of Computer Engineering, Middle East Technical University, Ankara, Turkey. E-mail: mozay@metu.edu.tr, vural@ceng.metu.edu.tr*

images are constructed for luminance, the blue and red chrominance values. Next, a Discrete Cosine Transform (DCT) is employed on the pooled images and three sets of DCT coefficients are constructed. In order to group the low frequency coefficients of the $64$ blocks, a zigzag scanning is performed with these three sets of DCT coefficients. Therefore, $3$ zigzag scanned matrices is obtained for each pooled images. In the experiments, a total of $12$ coefficients, $6$ for luminance and $3$ for each chrominance, are used.

- **Edge Histogram** (Feature Space Dimension $D_j = 80$) represents the spatial distribution of five types of edges, namely four directional edges (horizontal, vertical, left and right diagonal edges) and a non-directional edge for $16$ local regions in the image. Since there are five types of edges for each local region, $16 * 5 = 80$ histogram bins are computed.
- **Region-based Shape** (Feature Space Dimension $D_j = 35$) considers all the pixels which represent a shape, such as the boundary and interior pixels. Therefore, shapes of the objects consisting of a single connected region or multiple regions with holes can be represented with that feature. First, a shape is decomposed into a number of orthogonal complex-valued 2-D basis functions using Angular Radial Transform which is a 2D complex transform defined on a unit disk in polar coordinates. Then, the normalized and quantized magnitudes of coefficients are used to describe the shape. In the implementations, twelve angular and three radial functions are used to extract $35$ features, which are concatenated into a feature vector of Region-based Shape descriptor.
- **Dominant Color** (Feature Space Dimension $D_j = 16$) provides a compact representation of descriptive colors in an image region or a whole image. A dominant color feature vector is

$$\big((\overline{comp}_1, p_1, \sigma_1), (\overline{comp}_2, p_2, \sigma_2) \ldots, (\overline{comp}_N, p_N, \sigma_N), homog\big),$$

  where $N$ is the number of dominant colors in the image, $\overline{comp}_i$ is a vector of color space component values, $p_i$ is the fraction of pixels in the image or image region corresponding to the $i^{th}$ color such that $\sum_{i=1}^{N} p_i = 1$, $\sigma_i$ describes the variation of the color values of the pixels in a cluster around the $i^{th}$ color, and $homog$ is a single number that represents the overall spatial homogeneity of the dominant colors in the image. In the experiments, $3$ dominant colors are computed in $RGB$ space.
- **Scalable Color** ($D_j = 64$) computes a color histogram of an image in the $\mathcal{HSI}$ color space and applies a Haar transform-based encoding scheme across values of the histogram. In the algorithm, first the histogram values are computed, normalized and nonlinearly mapped into a four-bit integer representation, giving higher significance to small values. Then, the Haar transform is applied to the four-bit integer values across the histogram bins. In the experiments, $64$ coefficients of the Haar transform are used, which are equivalent to histograms with $64$ bins.
- **Homogenous Texture** ($D_j = 62$) describes the region texture in an image using the mean energy and the energy deviation from a set of frequency channels which are computed using Gabor filters computed on the image. Gabor features are computed for each angular direction $30° \times a$, where $a \in \{0, 1, 2, 3, 4, 5\}$ is the angular index and radial direction $B_b = B_0 \cdot 2^{-b}$, where $b \in \{0, 1, 2, 3, 4\}$ and $B_0$ is the largest bandwidth specified by $\frac{1}{2}$. Therefore, $30$ frequency channels are computed using Gabor filters. For each $i^{th}$ channel, the energy $en_i$ and the energy deviation $dev_i$, which are defined as as the log-scaled sum and standard devitation of the square of the Gabor-filtered Fourier transform coefficients of an image, are computed. Concatenating $en_i$ and $dev_i$, $\forall i = 1, 2, \ldots, 30$, with the mean and the standard deviation of the image, a $62$ dimensional vector is constructed as a Homogenous Texture feature vector.
- **Haar Descriptor** ($D_j = 195$) employs Haar wavelets to hierarchically decompose an image in order to obtain information about the coarseness of the image at different scales. At each scale, three different orientations of Haar wavelets, each of which responds to variances in

intensities across different axes, are used. Therefore, the information about how intensity varies in each color channel in the horizontal, vertical and diagonal directions is obtained. In the experiments, each image is scaled into a $128 \times 128$ pixel and $4$-layer two dimensional Haar wavelet transform is applied. Concatenation of the wavelet coefficients computed at each layer constructs the feature vector.

# 3 PROPERTIES OF THE MATLAB FILES

## 3.1 corel_dataset_all_features_labels.mat

Each feature set and the set of class labels of the samples are given in individual matlab ”mat” files.

In order to load the feature sets, you can use;
`load('feature_name.mat')`
In order to load the class labels, you can use;
`load('labels.mat')`
Then, the matlab workspace contains the following variables:

- **color_structure**: **Color Structure** features;
  - $N$ by $D_j$ feature matrix, where $N$ is the number of samples and $D_j$ is the dimension of the feature space.
- **color_layout**: **Color Layout** features;
  - $N$ by $D_j$ feature matrix, where $N$ is the number of samples and $D_j$ is the dimension of the feature space.
- **edge_histogram**: **Edge Histogram** features;
  - $N$ by $D_j$ feature matrix, where $N$ is the number of samples and $D_j$ is the dimension of the feature space.
- **region_shape**: **Region-based Shape** features;
  - $N$ by $D_j$ feature matrix, where $N$ is the number of samples and $D_j$ is the dimension of the feature space.
- **dominant_color**: **Dominant Color** features;
  - $N$ by $D_j$ feature matrix, where $N$ is the number of samples and $D_j$ is the dimension of the feature space.
- **scalable_color**: **Scalable Color** features;
  - $N$ by $D_j$ feature matrix, where $N$ is the number of samples and $D_j$ is the dimension of the feature space.
- **homogenous_texture**: **Homogenous Texture** features;
  - $N$ by $D_j$ feature matrix, where $N$ is the number of samples and $D_j$ is the dimension of the feature space.
- **edge_histogram**: **Haar Descriptor** features;
  - $N$ by $D_j$ feature matrix, where $N$ is the number of samples and $D_j$ is the dimension of the feature space.
- **labels**: Labels of the samples;
  - $N$ by $1$ feature vector, where $N$ is the number of samples.

## 3.2 class_name_and_index.mat

”class_name_and_index.mat” is C by 2 matlab cell, where C= $599$ is the number of classes.

In order to load the dataset, you can use;
`load('class_name_and_index.mat')`

The first column of the cell contains the category or class names, and the second column contains the category or class identity numbers (i.e. ids). For instance, **class_name_index**$\{$**1,1**$\}$ provides the class name *africa* with the associated class id **class_name_index**$\{$**1,2**$\}$ which is $1$. Therefore, the class name of the $1^{st}$ class is *africa*.

## REFERENCES

[1] M. Ozay and F. T. Yarman-Vural, "A new fuzzy stacked generalization technique and analysis of its performance," *CoRR*, vol. abs/1204.0171, 2012.

[2] H. Eidenberger, "Statistical analysis of content-based mpeg-7 descriptors for image retrieval," *Multimedia Systems*, vol. 10, no. 2, pp. 84–97, 2004.

[3] P. Salembier and T. Sikora, *Introduction to MPEG-7: Multimedia Content Description Interface*, B. Manjunath, Ed. New York, NY, USA: John Wiley & Sons, Inc., 2002.