Robust Face Recognition via Sparse Representation

Summary written by Sanika Phatak
February, 22, 2020

Summary: The paper introduces the concept of Sparse Representation Classification (SRC) using l_1 norm for face recognition. The paper tackles two major issue in face recognition - feature extraction and robustness to occlusion. It is proposed that if the sparsity is correctly computed and the number of features is sufficiently large then the choice of features is not important. The SRC framework introduced in the paper can also handle occluded images taking advantage of occlusion errors being sparse and also predicts how much occlusion can be tackled by the algorithm and suggest the choice for training images to maximise the robustness to occlusion.

Related work: The concept of sparse linear representation originated from statistical signal processing in [5]. The algorithm also builds on the concept of samples from each class lying on a linear subspace [1]. The optimisations for the algorithm are similar to Lasso i statistics [6] which penalize the l^1-norm of coefficients in linear combination.

Approach: The paper exploits the discriminative nature of sparse representation to perform classification by representing the test samples in an overcomplete dictionary whose base elements are the training samples. An overcomplete dictionary is defined as a concatenation of n training samples for k classes: $A = [A_1, A_2, ..., A_k] = [v_{1,1}, v_{1,2}, ...v_{k,nk}]$ where each vector v of length m represents an image generated by stacking all its columns together. Now, the test image y can be represented linearly : $y = Ax_o \epsilon \mathbb{R}^m$ where the coefficient vector, $x_o = [0, 0, ..., \alpha_{i,1}, \alpha_{i,2}, ...\alpha_{i,ni}, ...0, 0]$, α being a scalar value. All other values except for the i^{th} class are zero hence introducing sparsity. The sparsest solution for y is $argmin||x||_0$ (l^0-norm) which counts the number of non zero entries in a vector. According to the paper [4] if the solution is sparse enough, $argmin||x||_0 = argmin||x||_1$ and hence can be easily found by linear programming methods. When the sparse solution is computed by minimising l^1-norm, the classification is done based on how well each associated coefficient of each training class reproduces the test image y. If $\delta_i : \mathbb{R}^n - > \mathbb{R}^n$ is the characteristic function, then y is classified on the basis of the residuals: $min\, r_i(y) = ||y - A\delta_i(\hat{x_1})||_2$. SRC also tackles two major issues in face recognition: i) Feature extraction: Most feature transformations are linear and hence the image space can be represented by the feature space instead and still the same procedure of SRC can be followed for classification with the only difference being the matrix A being replaced by a lower (d) dimensional random feature matrix $RA \,\epsilon\, \mathbb{R}^{dxn}$. By the property of ”blessing of dimensionality” proved in [3, 2], if the solution is sparse enough, SRC will always give the same classification result regardless of the features used. ii) Robustness to Occlusion: Incase of occlusion, we can represent $y = Ax_o + e_o$ with e being the error vector. The occluded pixels cover a small part of the image and hence error vector is also sparse. So, $y = [A, I] \begin{pmatrix} x_o \\ e_o \end{pmatrix} = Bw_o$ where the identity matrix handles general classes of corruption and hence the SRC algorithm can be extended to the matrix B to tackle occlusion.

Experiments and results: When compared against Nearest Neighbors, Linear SVM and Nearest Subspace, SRC gave consistently better recognition rates on Yale database. For random pixel corruption, SRC significantly performs better than the other classifiers performing almost perfectly till 60% corruption.

Strengths: SRC exploits the concept of sparsity to make the face recognition feature space independent along with tackling occlusion issues. Many algorithms need a complex feature extraction algorithm for classification but SRC makes the classification feature independent. Simple linear mathematics is used to tackle the problem of occlusion holistically. The training images used need not be very specific as long as they represent all the classes.

Weaknesses: Computationally very intensive and take very long time to compute the training coefficient matrix. The dimensions of the data can be reduced only to a specific value and needs to be above a threshold to make the data sparse. The paper only uses frontal images and extending the same algorithm to varying pose would require a very high training database or non linear models.

Reflections: SRC is a very robust algorithm to tackle face recognition issues like occlusion and feature selection. The algorithm performs significantly well than existing algorithms but is computationally intensive. There is a huge scope of improvement in the current model to cover holistic issues like pose and variation in lighting in face recognition. The robustness of the algorithm lies in its simplicity with linear mathematics.

References

[1]

[2] E. Candes. Compressive sampling. 2006.

[3] D. Donoho. High dimensional data analysis: The curses and blessings of dimensionality. 2000.

[4] D. Donoho. For most large indetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution. 59(6):797–829, 2006.

[5] K. Huang and S. Aviyente. Sparse representation for signal classification. 2006.

[6] P. Zhao and B. Yu. On model selection consistency of lasso. (7):2541–2567, 2006.