A Novel Locally Linear KNN Model for Visual Recognition

Qingfeng Liu, Chengjun Liu
Department of Computer Science, New Jersey Institue of Technology.

For most previous methods [1], [3], [4], [5], representation and classification are developed independently, which violates the need that the representation methods should serve and facilitate the subsequent classification methods for visual recognition. In addition, these methods bring other issues such as classifier restriction, computational complexity etc..

To address these issues, this paper presents a novel locally linear KNN model with the goal of not only developing efficient representation and classification methods, but also establishing a relation between them to approximate some classification rules, e.g. the Bayes decision rule for minimum error.

First, the proposed model represents the test sample as a linear combination of all the training samples and derives a new representation by learning the coefficients considering the reconstruction, locality and sparsity constraints as equation 1.

$$\min_{\mathbf{v}} ||\mathbf{x}-\mathbf{B}\mathbf{v}||^2 + \lambda||\mathbf{v}||_1 + \alpha||\mathbf{v}-\beta\mathbf{d}||^2 \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the test sample, $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_m] \in \mathbb{R}^{n\times m}$ is the training sample matrix and coefficients vector $\mathbf{v} \in \mathbb{R}^m$ is the derived representation. The vector $\mathbf{d} = [d_1, d_2, ..., d_m]^t \in \mathbb{R}^m$, and $d_i = exp\{-\frac{1}{2\sigma^2}||\mathbf{x}-\mathbf{b}_i||^2\}$. The parameter σ is used for adjusting the decay speed.

The theoretical analysis shows that the new representation has the **grouping effect of the nearest neighbors (GENN)**, which is able to approximate the "ideal representation".

Theorem 0.1 ***Grouping effect of the nearest neighbors***. *Given a L_2 normalized test sample $\boldsymbol{x}$ ($||\boldsymbol{x}||^2 = 1$), the L_2 normalized training sample matrix $\boldsymbol{B}$ ($||\boldsymbol{b}_i||^2 = 1, i = 1, 2, ..., m$) and the vector $\boldsymbol{d} = [d_1, d_2, ..., d_m]^t$, let $\boldsymbol{v}^* = [v_1^*, v_2^*, ..., v_m^*]^t$ be the solution to the LLKNN model defined in equation 1. Define the sample correlation ρ of two training samples $\boldsymbol{b}_i$ and $\boldsymbol{b}_j$ as $\rho = \boldsymbol{b}_i^t\boldsymbol{b}_j$ and the difference between the coefficients v_i^* and v_j^* $(i, j = 1, 2, ..., m)$ as*

$$M(i, j) = |v_i^* - v_j^*| \tag{2}$$

Then, if the signs of v_i^* and v_j^* are the same, we have

$$M(i, j) \le \frac{C}{\alpha}\sqrt{2(1-\rho)} + \beta|d_i - d_j| \tag{3}$$

where $C = \sqrt{(1+\alpha\beta^2||\boldsymbol{d}||^2)}$, which is a constant.

And then the locally linear KNN model based classifier (LLKNNC) (see equation 4) and the locally linear nearest mean classifier (LLNMC) (see equation 5), whose relation is just like the KNN classifier to the nearest mean classifier, are derived.

$$c^* = \arg\max_c \sum_{\mathbf{b}_i \in \mathbf{B}_c} v_i \tag{4}$$

$$\begin{aligned} c^* &= \arg\min_c ||\mathbf{x}-\mathbf{m}_c||_2^2 \\ &= \arg\min_c ||\mathbf{x} - \sum_{\mathbf{b}_i \in \mathbf{B}_c} v_i\mathbf{b}_i||_2^2 \end{aligned} \tag{5}$$

The power of the proposed LLKNNC is guaranteed by establishing its connection to the Bayes decision rule for minimum error in the view of kernel density estimation as stated in theorem 0.2.

Theorem 0.2 *Given the test sample $\boldsymbol{x}$, the corresponding representation $\boldsymbol{v}$, the two transformations $v_i = \frac{v_i - v_{min}}{v_{max}-v_{min}}$ and $v_i = \frac{v_i}{\sum_{i=1}^m v_i}$ are applied first, where v_{min} and v_{max} is the minimal and maximal value among all the elements of the vector $\boldsymbol{v}$.*

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

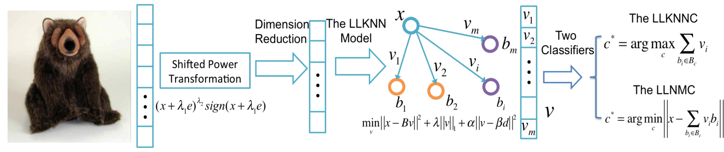

Figure 1: The system architecture of the proposed Locally Linear KNN model.

Then, if the prior distribution $p(c)$ is equal for all the classes, the Bayes decision rule is approximated by the proposed LLKNN model based classifier in the sense of kernel density estimation.

$$\begin{aligned} c^* &= \arg\max_c \sum_{\boldsymbol{b}_i \in \boldsymbol{B}_c} v_i \\ &\approx \arg\max_c \sum_{\boldsymbol{b}_i \in \boldsymbol{B}_c} \beta d_i + const \\ &\propto \arg\max_c p(c|\boldsymbol{x}) \end{aligned} \tag{6}$$

Furthermore, theorem 0.2 tells us that the power of LLKNNC comes from the kernel density estimation. Then there are two issues of the kernel density estimation that should be resolved to improve the reliability, namely the sensitiveness to the global window width denoted as the value of σ and the adverse impact of distant neighbors.

As for the sensitiveness to the value of σ, the shifted power transformation [2] (see equation 7) is first applied to transform data to a near Gaussian shape so that the new data can be well estimated. Then the L_2 normalization is further applied to the vector $\mathbf{d}$.

$$T(\mathbf{x}) = |\mathbf{x}+\lambda_1\mathbf{e}|^{\lambda_2} sign(\mathbf{x}+\lambda_1\mathbf{e}) \tag{7}$$

As for the adverse impact of distant neighbors, a coefficients cut-off method is applied. The LLKNNC thus is defined as follows

$$c^* = \arg\max_c \sum_{\substack{(\mathbf{b}_i \in \mathbf{B}_c)\wedge \\ (v_i \in T(k))}} v_i \tag{8}$$

where T(k) is the set of top k largest values of v_i for each class. Similarly, the LLNMC is defined as follows:

$$c^* = \arg\min_c ||\mathbf{x} - \sum_{\substack{(\mathbf{b}_i \in \mathbf{B}_c)\wedge \\ (v_i \in T(k))}} v_i\mathbf{b}_i||_2^2 \tag{9}$$

Experimental analysis on many representative databases such as MIT-67 indoor scenes dataset and the Caltech 256 dataset demonstrate that our method can achieve competitive results to the state-of-the-art methods.

[1] Zhuolin Jiang, Zhe Lin, and Larry S. Davis. Label consistent k-svd: Learning a discriminative dictionary for recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(11):2651–2664, 2013.

[2] Ruppert D Wand M.P., Marron J.S. Transformations in density estimation. *Journal of the American Statistical Association*, 86(414):343–353, 1991.

[3] John Wright, Allen Y. Yang, Arvind Ganesh, Shankar S. Sastry, and Yi Ma. Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(2):210–227, 2009.

[4] Meng Yang, Lei Zhang, Xiangchu Feng, and David Zhang. Sparse representation based fisher discrimination dictionary learning for image classification. *International Journal of Computer Vision*, pages 1–24, 2014.

[5] Qiang Zhang and Baoxin Li. Discriminative k-svd for dictionary learning in face recognition. In *CVPR*, pages 2691–2698, 2010.