

# Plant Disease Recognition: A Visual Region and Loss Reweighting Approach

Keerthi K S<sup>1</sup> and Bincy Louis<sup>2</sup>

M.Tech Student, Department of Electronics and Communication<sup>1</sup>

Assistant Professor, Department of Electronics and Communication<sup>2</sup>

Loudes Matha College of Science and Technology, Thiruvananthapuram, Kerala, India

**Abstract:** *Plant disease diagnosis is very important for agriculture because of its importance in increasing crop production. Nowadays the advances in image processing gives a new way. One of the newest way to solve this issue via visual plant disease analysis. In this paper, we discuss the problem of plant disease recognition. Here we tackle plant disease recognition via reweighting both visual regions and loss to emphasize diseased parts. We first compute the weights of all the divided patches from each image based on the cluster distribution of these patches to indicate the discriminative level of each patch. Then we allocate the weight to each loss for each patch-label pair during weakly supervised training to enable discriminative disease part learning. We finally extract patch features from the network trained with loss reweighting, and utilize the LSTM network to encode the weighed patch feature sequence into a comprehensive feature representation. Extensive evaluations on this dataset and another public dataset demonstrate the advantage of the proposed method. We expect this research will further the agenda of plant disease recognition in the community of image processing.*

**Keywords:** Long Sort Term Memory, Feature Integration, Cluster

## I. INTRODUCTION

Plant diseases cause threats to food security by reducing crop production. 20%-40% of all crop losses globally are due to plant diseases. Therefore, plant disease diagnosis is critical to the prevention of spread of plant diseases and reduction of economic losses in agriculture. Nowadays image-based technologies are being widely applied. The recent advances in machine learning, especially deep learning we assert that plant image analysis and recognition can also provide a new way for plant disease diagnosis[1]. The applications in visual plant disease diagnosis conversely promote the development of image processing technologies. There are mainly three different characteristics for plant disease images taken. The first one is Randomly distributed lesions where the foliar lesions probably randomly occur in the plant leaves. The cherry fungal shot hole disease is distributed in many different parts of the leaf, including the top, left and right positions. Because deep convolutional neural networks trained with image level labels only tend to focus on the most discriminative parts while missing other object parts. The second one is Diverse symptoms here even for the same plant disease, there are probably various visual symptoms on the plant leaves at different time periods. The appearances vary considerably in different infected stages, leading to large intra-class variations. The third one is Complex backgrounds. There usually exist various background clutters in real-world scenarios. There are dense leaves and any other irrelevant objects in the background. The disease symptoms are not salient, making plant disease recognition more difficult. Taking the characteristics of the plant disease image into consideration, we tackle visual plant disease recognition via reweighting both visual regions and loss. Particularly, considering randomly distributed lesions, we explore the multi-scale strategy by dividing the plant disease images into non-overlapping patches, and compute the weights of these patches according to the cluster distribution of these patches in order to indicate the discriminative level of each patch. By setting different weights to different patches, we enhance the influence of patches with diseased symptoms and reduce the interference of irrelevant patches. We further allocate the weight to each loss for each patch-label pair during weakly-supervised training for diseased parts learning. Finally, we extract patch features from the network trained with loss reweighting and adopt a LSTM network to encode the weighted patch feature sequence into a comprehensive feature representation Long short-term memory is an artificial neural network used in the fields of artificial intelligence and deep learning. LSTMs are predominantly used to learn, process, and

classify sequential data because these networks can learn long-term dependencies between time steps of data. Common LSTM applications include sentiment analysis, language modeling, speech recognition, and video analysis.

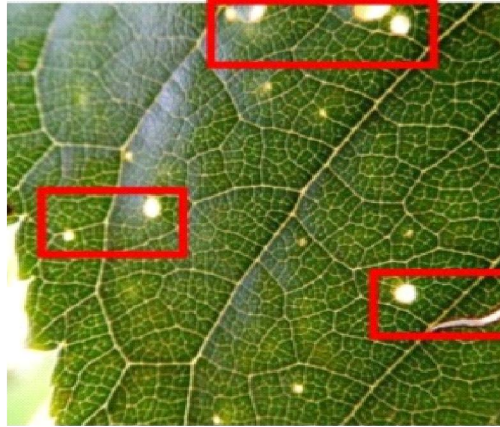


Fig 1. Randomly distributed lesions.



Fig 2. Diverse symptoms



Fig 3. Complex backgrounds

## II. RELATED WORKS

Plant disease diagnosis is critical to the prevention of spread of crop diseases and reduction of economic losses in agriculture [1]. Most of traditional methods rely on the molecular technologies [2], [3] that are complicated, time-consuming and constrained to centralized laboratories. Therefore, some works adopt traditional computer vision methods for plant disease recognition, such as hyperspectral image analysis [4], artificial bee colony algorithm [5], and image segmentation [6]. Recently, there have been more attempts to utilize deep learning in plant disease recognition

for its powerful capability of discriminative feature learning [7]–[10]. For example, Wang et al. [10] finetuned the VGG, ResNet50, and GoogleNet directly on the leaf disease set and Ferentinos et al. [9] finetuned the AlexNet and GoogleNet directly to identify 14 crop species and 26 diseases. However, most of them directly extract deep features without considering the characteristics of the plant disease image. Besides, most of the works conduct their evaluations on small-scale datasets. Table I summarizes the most common plant disease and crop pest datasets. We can see that PlantVillage Dataset [11] is the largest plant disease dataset, but only contains 38 plant disease categories. In addition, the images from this dataset are taken on the table, and not in the real-world scenarios. We show some samples of these leaf datasets in Fig. 4. Different from these works, we systematically analyze the problem of plant disease recognition and propose a multi scale method to reweight visual regions and the loss to emphasize discriminative diseased parts for plant disease recognition based on the characteristics of the plant disease image. Furthermore, we collect a large-scale plant disease dataset PDD271, which not only has the advantage in data volume and category coverage, but also is collected in real-world scenarios with complex background. In particular, there is another agricultural dataset IP102 [12], which is relevant to crop pest. This dataset contains more than 75,000 images belonging to 102 categories for insect pest recognition. Fine-grained image recognition aims to distinguish sub ordinate categories, such as birds and food. In the early stage, researchers [13], [14] based on deep learning first used strong supervised mechanisms with part bounding box annotations to learn to attend on discriminative parts. Recent researches [15], [16], [17]–[22] focused on weakly-supervised recognition methods without high-cost object part locations or attribute annotations. For example, Yang et al. [22] initialized many anchors randomly and extracted their features as their informativeness using the RPN method, and finally chose the informative region to improve the classification performance. There are also several attention-based methods proposed for Fine-Grained Visual Classification. For example, Hu et al. [23] used attention maps to guide the data augmentation, Peng et al. [24] proposed the object-part attention model to select discriminative regions subjecting to the object-part spatial constraint, and SeNet154 [25] enhance the recognition performance with spatial-channel attention. However, attention-based methods probably focus on the most discriminative parts while missing other parts for the whole image. recognition method via reweighting both visual regions and the loss to emphasize the diseased parts for the plant disease recognition. Our method is heuristic with many possibilities to improve. For example, our method and attention-based methods can work together in one framework to further enhance the recognition performance, such as gate-attention deep networks [26]. We can also conveniently combine our framework with the spatial context through discriminative spatio-appearance kernels [27] to promote the performance.

III. SYSTEM STUDY

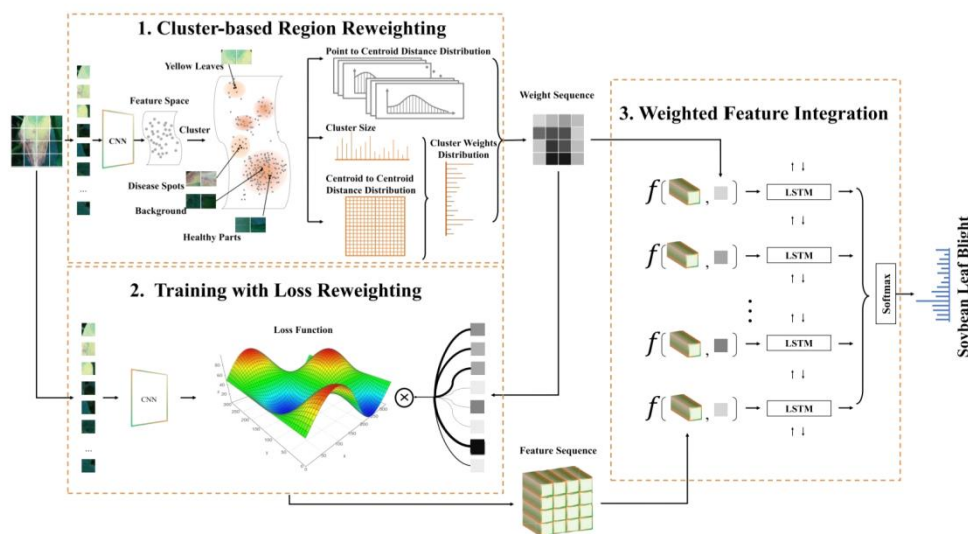


Fig 4. The proposed plant disease recognition framework

There are mainly three different stages which include Cluster Based Region Reweighting, Training with Loss Reweighting and Weighted Feature Integration . We introduce the proposed framework which explores a multi-scale



strategy and reweights both visual regions and the loss during the weakly-supervised learning to emphasize discriminative diseased parts for the purpose of the plant disease recognition. CRR takes all the divided patches from plant disease images as input and sets the weight of each patch according to the cluster distribution of the visual features of these patches. For each patch-label pair, TLR allocates the corresponding weight to each loss during weakly supervised training in order to enable the discriminative disease part learning. Based on extracted patch features from TLR and corresponding weights from CRR, WFI utilizes the LSTM network to encode the weighed patch feature sequence into a comprehensive feature representation.

IV. METHODS

4.1 Cluster-Based Region Reweighting

Many diseases present small and scattered lesions. The deep convolutional neural networks trained with image level labels often overlook these lesions while focusing on more salient parts. Considering these situations, we explore a multi-scale strategy by dividing the images into non-overlapping patches and enlarging every patch to avoid missing diseased patches. However, the disease-independent patches, such as the complex backgrounds and the healthy parts, are enhanced even more in the above process, which could lead to severe unbalance between the diseased patches and the irrelevant ones. To address this problem, we attempt to use the visual similarity among the same disease to cluster the patches of the same disease. Afterwards, we reweight the patches based on the clustering result and indicate the discriminative level of each patch. Formally, all patches from all the original training images form a new training set. Let X ∈ R^{m×N} denotes the visual features of these patches, where m is the dimension of the visual feature and N is the number of training patches. We then have these patches clustered into k cluster classes c with their centroids being {μ 1 , μ 2 , . . . , μ k } ∈ R m . To compute the weight w x , x ∈ X, the weights of the clusters w c and the probability distribution p x of x belonging to over all clusters are computed.

Then, w x is computed as:

w x = p x · w c (1)

where w c = [w c 1 , . . . , w c i , . . . , w c k ] and w c i denotes the weight of the cluster c i . Normally, the patches containing similar visual symptoms are likely to be assigned to the same clusters. In case of small distance among clusters, the visual phenotypes of different diseases are similar and hard to distinguish by the deep model. Therefore, these clusters are given higher weights to enhance their influence in follow-up feature learning and integration. The size of cluster is also an important indicator. There is a highly skewed distribution of different disease patches. For example, the number of non-diseased patches containing complex backgrounds and foliar healthy parts is very large, but the number of patches containing cotton eye spot disease is small due to the concentrated symptom of this disease leading to the poor classification performance. Meanwhile, the distance between two clusters indicates their visual difference. If one cluster is far from the other clusters, we can easily obtain discriminative features for this cluster, thus assign a small weight to it. Hence we assign these clusters suitable weights to make their influences as balanced as possible. Given all these, we assign the cluster weights according to the following rule: the larger size the cluster and the farther away from the others, the smaller its weight. We use a monotone decreasing function F = e^{x/(x-1)} to model this change. According to the size of the cluster and the distance distribution among the cluster centroids, we compute the weights of the cluster c i as follows,

w\_{c\_i} = F(N\_{c\_i}) \times F(\sum\_{j \neq i, j \in 1, \dots, k} d(\mu\_i, \mu\_j)), (2)

where N\_{c\_i} is the number of patches in cluster c i and d(\mu\_i, \mu\_j) is the distance between the centroid \mu\_i and \mu\_j .To compute the probability distribution p\_x , we use a soft assignment strategy based on the distances between a patch and the cluster centroids. The assignment probability distribution p\_x is computed as following,

p\_x = F(d(x, \mu\_i)), i \in 1, \dots, k. (3)

The weight w\_x is then computed using Equation 1. For the patches from validation and testing datasets, we compute the probability distributions p\_x based on the distance to the centroids learned from the training patches. The cluster weights







have been computed and the patch weights are computed via Equation 1. The patches and their corresponding weights are used for the model training with loss reweighting.

4.2 Training with Loss Reweighting

To extract more discriminative regional features for the given patches, we train the network with a reweighted loss function. An observed input patch x shares the same label l with the original image. The model computes an observation o for this patch. The score can be interpreted as an estimation of a class posterior probability p\_theta(o|x), where theta is the model parameters. Given labeled training data {(x\_n, l\_n) : n = 0, . . . , N - 1}, the original cross entropy loss is defined as:

Loss\_ori(x\_n, l\_n; theta) = - sum\_{n=0}^{N-1} sum\_{l=1}^M y\_{o\_n, l\_n} log(p\_theta(o\_n|x\_n)) (4)

y is one binary indicator defined as:

y\_{o\_n, l\_n} = { 1 if o\_n = l\_n, 0 if o\_n != l\_n } (5)

However, this loss treats every patch equally. As a result, patches irrelevant to the disease symptoms distract the optimization of network. To solve this, we propose a new reweighted loss to enhance the influence of patches with discriminative diseased symptoms and to reduce the interference of irrelevant patches. For the observed input patch x, we define the reweighted loss function as:

Loss\_rew(x\_n, w\_{x\_n}, l\_n; theta) = - sum\_{n=0}^{N-1} sum\_{l=1}^M w\_{w\_n} y\_{o\_n, l\_n} log(p\_theta(o\_n|x\_n)). (6)

We allocate the weight to each loss for each patch-label pair. This loss forces the model to focus on the patches with discriminative diseased parts and to ignore the irrelevant patches as much as possible. This trained model can be used to extract visual features from all the patches. The patch features from the same image form a sequence as the input for the following weighted feature integration. For clarity, x\_hat denotes patch feature from the same image.

4.3 Weighted Feature Integration

The combination of diseased and healthy patches in plant images constitutes the complex and diverse visual patterns. We try to model the semantic correlation from the combination of local patches. Specifically, we propose a feature integration model with reweighting patch features as the inputs to induce the BiLSTM network to model the semantic correlativity among patches by end-to-end training. Given a feature sequence S = [ x\_hat\_1, . . . , x\_hat\_t ] extracted from the network with TLR and its corresponding weight sequence W = [ w\_hat\_1, . . . , w\_hat\_t ] obtained from CRR, where t denotes the number of patches for one image, we combine the feature sequence with the weight via a following function A(S, W),

A(S, W) = [ f(x\_hat\_1, w\_hat\_1), . . . , f(x\_hat\_t, w\_hat\_t) ] (7)

There is common two-layer stacked LSTM is adopted to fuse weighted patch feature sequences into the final representation. The hidden state of the first LSTM is fed into the second LSTM layer which follows the reversed order of the first one. The dimension of hidden states from both layers is 4,096.

The output o' = L(A(S, W); theta'), where theta' is the model parameters. We use softmax to generate the class probability vector for each image Si', denoted as phi(L(A(Si', Wi'); theta')) in R^M x 1. The final loss function is defined as:



$$Loss_{Lstm}(S_{i'}, W_{i'}; \theta') = - \sum_{i'=0}^{N_i-1} \sum_{l=1}^M y_{o'_{i'}, l_{i'}} \log(\phi(L(A(S_{i'}, W_{i'}); \theta'))). \quad (8)$$

By optimizing this loss function, we obtain a weighted encoded patch feature sequence into a comprehensive feature representation for plant disease recognition.

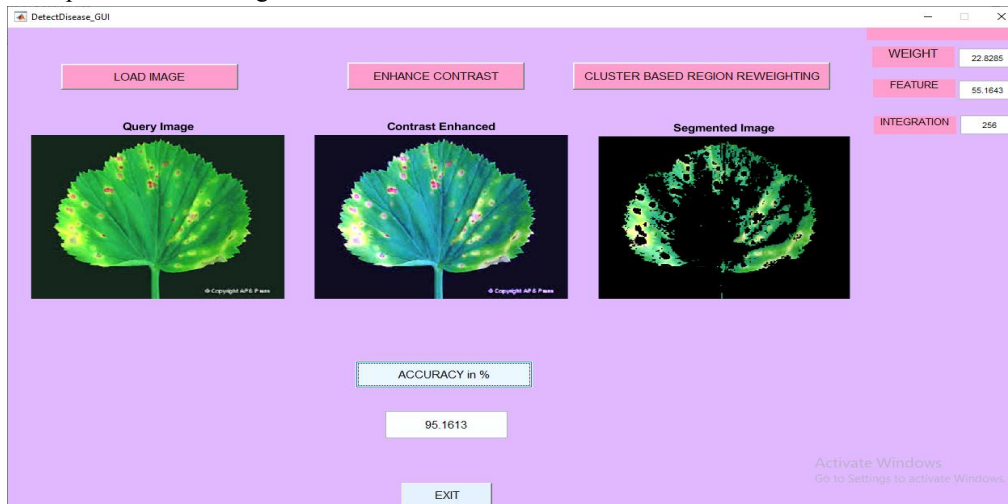


Fig 5. Proposed system result

## V. CONCLUSION

Plant disease recognition is an interesting and practical topic. However, this problem has not been sufficiently explored due to the lack of systematical investigation and large-scale dataset. The most challenging step in constructing such a dataset is providing a reasonable structure from both the agriculture and image processing perspective. In this paper, we systematically investigate the problem of plant disease recognition in the community of image processing. With the help of agriculture experts, we construct the first large-scale plant disease dataset with 271 plant disease categories and 220,592 images. Furthermore, we present a plant disease oriented framework for plant disease recognition based on their distinctive characteristics. We design a strategy to compute patch weights based on the cluster distribution of patch features and then use learned weights to reweight both patch features to highlight diseased patches and the loss to guide the model optimization.

## REFERENCES

- [1]. Z. Li et al., "Non-invasive plant disease diagnostics enabled by smartphone-based fingerprinting of leaf volatiles," *Nature Plants*, vol. 5, no. 8, pp. 856–866, Aug. 2019.
- [2]. A. J. Wakeham, G. Keane, and R. Kennedy, "Field evaluation of a competitive lateral-flow assay for detection of alternaria brassicae in veg- etable Brassica crops," *Plant Disease*, vol. 100, no. 9, pp. 1831–1839, Sep. 2016
- [3]. A. K. Lees, L. Sullivan, J. S. Lynott, and D. W. Cullen, "Development of a quantitative real-time PCR assay for phytophthora infestans and its applicability to leaf, tuber and soil samples," *Plant Pathol.*, vol. 61, no. 5, pp. 867–876, Oct. 2012.
- [4]. C. H. Bock, G. H. Poole, P. E. Parker, and T. R. Gottwald, "Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging," *Crit. Rev. Plant Sci.*, vol. 29, no. 2, pp. 59–107, Mar. 2010.
- [5]. F. Ahmad and A. Airuddin, "Leaf lesion detection method using artificial bee colony algorithm," in *Advances in Computer Science and its Applications*, vol. 279. Beijing, China: Springer, 2014, pp. 989–995.

- [6]. S. Prasad, P. Kumar, and A. Jain, "Detection of disease using block- based unsupervised natural plant leaf color image segmentation," in Swarm, Evolutionary, and Memetic Computing. Beijing, China: Springer, 2011, pp. 399–406.
- [7]. S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–11, May 2016.
- [8]. J. Wang, L. Chen, J. Zhang, Y. Yuan, M. Li, and W. Zeng, "CNN transfer learning for automatic image-based classification of crop disease," in *Image and Graphics Technologies and Applications*. Beijing, China: Springer, 2018, pp. 319–329.
- [9]. K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Comput. Electron. Agricult.*, vol. 145, pp. 311–318, Feb. 2018.
- [10]. G. Wang, Y. Sun, and J. Wang, "Automatic image-based plant disease severity estimation using deep learning," *Comput. Intell. Neurosci.*, vol. 2017, pp. 1–8, Jul. 2017.
- [11]. D. P. Hughes and M. Salathé, "An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing," 2015, arXiv:1511.08060. [Online]. Available: <https://arxiv.org/abs/1511.08060>
- [12]. X. Wu, C. Zhan, Y.-K. Lai, M.-M. Cheng, and J. Yang, "IP102: A large-scale benchmark dataset for insect pest recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.(CVPR)*, Jun. 2019, pp. 8787–8796.
- [13]. J. Donahue et al., "Decaf: A deep convolutional activation feature for generic visual recognition," in *Proc. Int. Conf. Mach. Learn.*, vol. 2014, pp. 647–655.
- [14]. N. Zhang, J. Donahue, R. Girshick, and T. Darrell, "Part-based R-CNNs for fine-grained category detection," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 834–849.
- [15]. [15] W. Min, S. Jiang, L. Liu, Y. Rui, and R. Jain, "A survey on food computing," *ACM Comput. Surv.*, vol. 52, no. 5, pp. 92:1–92:36, 2019.
- [16]. W. Ge, X. Lin, and Y. Yu, "Weakly supervised complementary parts models for fine-grained image classification from the bottom up," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3034–3043.
- [17]. T. Xiao, Y. Xu, K. Yang, J. Zhang, Y. Peng, and Z. Zhang, "The application of two-level attention models in deep convolutional neural network for fine-grained image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 842–850.
- [18]. M. Lam, B. Mahasseni, and S. Todorovic, "Fine-grained recognition as HSnet search for informative image parts," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6497–6506.
- [19]. T.-Y. Lin, A. RoyChowdhury, and S. Maji, "Bilinear convolutional neural networks for fine-grained visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 6, pp. 1309–1322, Jun. 2018.
- [20]. S. Jiang, W. Min, L. Liu, and Z. Luo, "Multi-scale multi-view deep feature aggregation for food recognition," *IEEE Trans. Image Process.*, vol. 29, pp. 265–276, 2020.
- [21]. W. Min, L. Liu, Z. Luo, and S. Jiang, "Ingredient-guided cascaded multi- attention network for food recognition," in *Proc. 27th ACM Int. Conf. Multimedia*, Oct. 2019, pp. 99–107.
- [22]. Z. Yang, T. Luo, D. Wang, Z. Hu, J. Gao, and L. Wang, "Learning to navigate for fine-grained classification," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 438–454.
- [23]. T. Hu and H. Qi, "See better before looking closer: Weakly supervised data augmentation network for fine-grained visual classification," *CoRR*, vol. abs/1901.09891, 2019.
- [24]. Y. Peng, X. He, and J. Zhao, "Object-part attention model for fine- grained image classification," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1487–1500, Mar. 2018.
- [25]. J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.
- [26]. G.-J. Qi, "Hierarchically gated deep networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2267–2275.

- [27]. G.-J. Qi, X.-S. Hua, Y. Rui, J. Tang, and H.-J. Zhang, "Image classification with kernelized spatial-context," IEEE Trans. Multimedia, vol. 12, no. 4, pp. 278–287, Jun. 2010.