ECCV'20 ONLINE

23-28 AUGUST 2020

16TH EUROPEAN CONFERENCE ON

COMPUTER VISION

WWW.ECCV2020.EU

Patch-based CNN evaluation for bark classification

Debaleena Misra, Carlos Crispim-Junior, Laure Tougne

Univ Lyon, Lyon 2, LIRIS, F-69676 Lyon, France

Outline

- Background
  - Motivation
  - Objective
- Methodology
  - Patch preparation
    - Image up-scaling
    - Patch extraction
  - Transfer learning with patches
  - Compute image-level labels
    - Majority voting
    - Tie-breaking strategies
- Experimental details
- Results
- Conclusion

Motivation

Tree recognition from only bark images is a challenging and interesting problem.

The organ bark has several advantages :

- Is a consistent organ found round the year
- Has a slow aging process
- Easily accessible compared to higher-level fruits, flowers or leaves

Challenges of bark recognition :

- Texture properties impacted by environmental factors and diseases
- Branch shadow clutter and occlusion
- Lack of large bark datasets

Objective

Evaluate the feasibility of using deep learning for bark recognition from challenging datasets having

- **High number of classes**
- **Very few samples for some classes**
- **High intra-class variance & low inter-class variance**
- **Huge variation in image dimensions**

Dataset overview

Bark-101 dataset [1]

- 2587 images from 101 classes
- 50-50 train-test split





Related work

Traditional statistical methods using hand-crafted features

- Textual analysis methods with Gray level run-length method (RLM), Concurrence matrices (COMM) & Histogram inspection [4]
- Spectral methods using Gabor filters and descriptors SURF or SIFT [5]
- LBP-inspired texture descriptors (LCoLBP, SMBP, CLBP) with SVM & KNN classifiers [1]
- Gabor wavelets with radial basis probabilistic neural networks (RBPNN) [6]

Recent deep learning based approaches

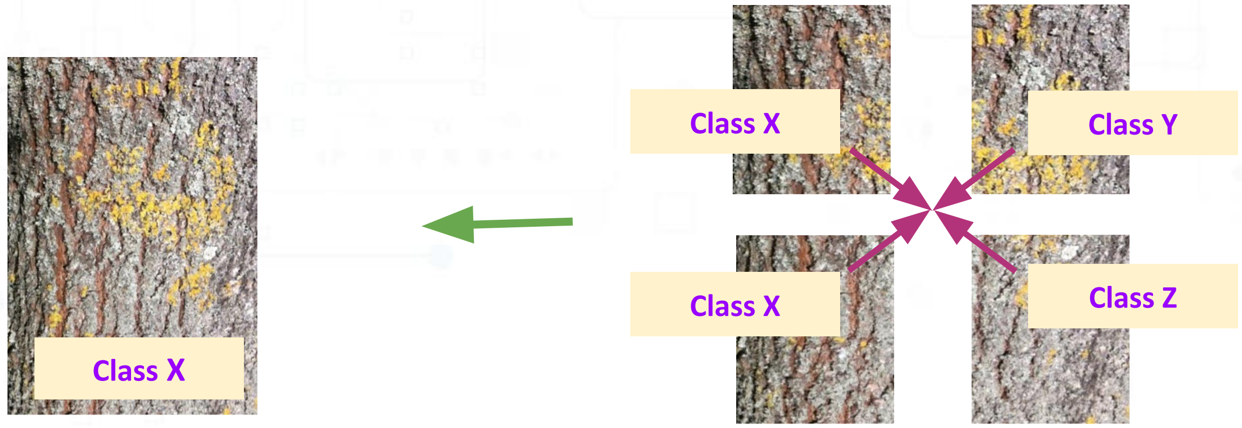
- AlexNet to study depth images of bark from 2 plant species [7]
- ResNet for BarkNet dataset with 23 classes [8]

Proposed methodology

Step 1 : Image up-scaling

Step 2 : Patch extraction after up-scaling

Step 3 : CNN fine-tuning & inference over patches



Final label for the image

Step 4 : Majority voting with tie-breaking

- **Steps 1 & 2** : Patch preparation
- **Step 3** : Transfer-learning with patches
- **Step 4** : Image-level label assignment by majority voting of patch-level results

ECCV'20 ONLINE
23-28 AUGUST 2020

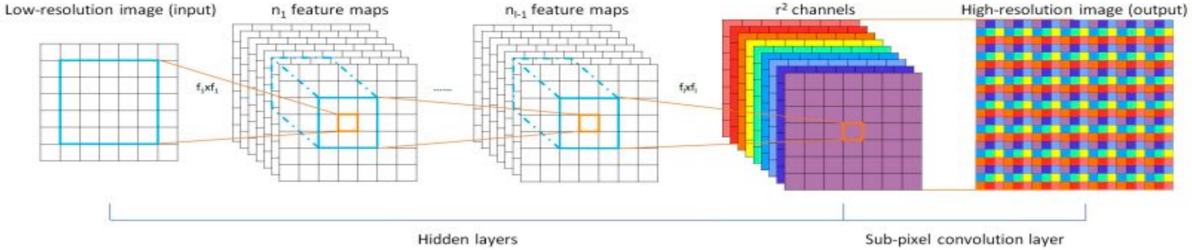
Image Up-scaling

Method 1 : Traditional Bicubic interpolation

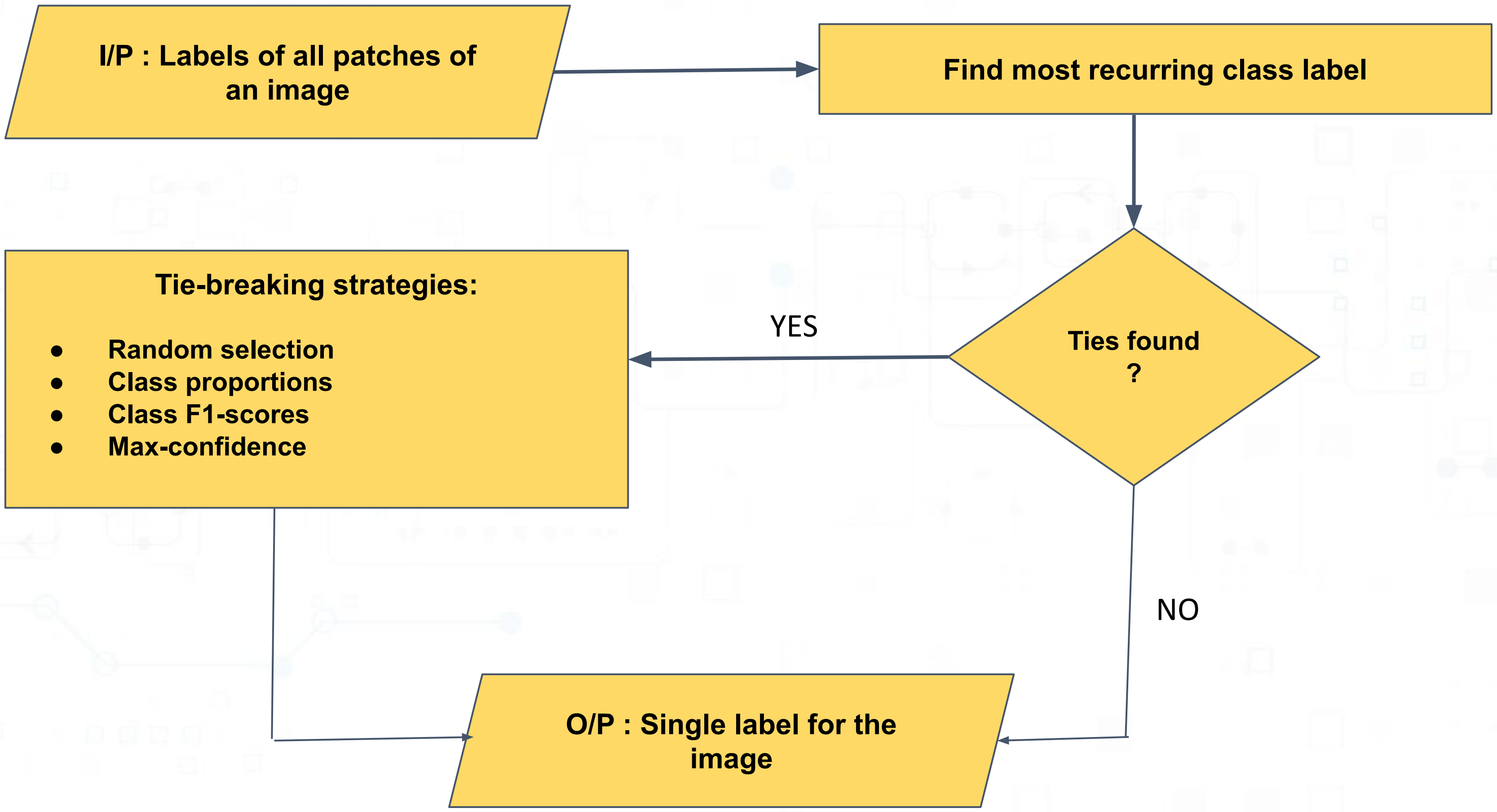
- Classical image upsampling algorithm
- Uses geometrical transformations in 2D images with Lagrange polynomials, cubic splines or cubic convolutions
- Output pixel value is computed as weighted sum of pixels in 4-by-4 pixel neighborhood

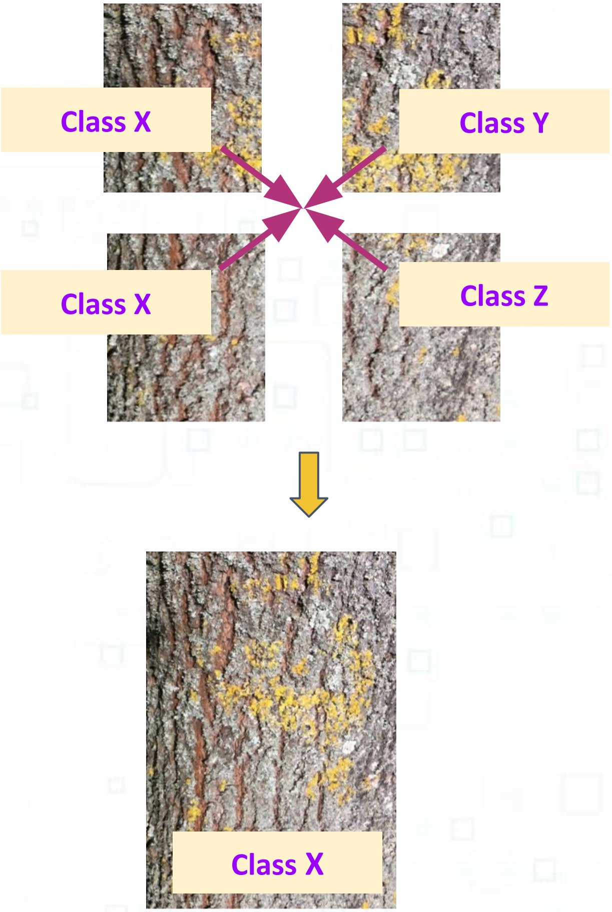
Method 2 : Efficient Sub-pixel Convolutional Neural Network (ESPCN) [2]

- 2 convolution layers for feature extraction in LR space
- 1 sub-pixel convolution layer for learning an array of upscaling filters to aggregate LR feature maps into SR image in single step



Majority Voting





Tie-breaking strategies

In case of occurence of ties, following strategies are applied on the set of tied classes :

- **Random selection :** Any one class is randomly selected
- **Class proportions :** Selects class having the highest number of training samples
- **Class F1-score :** Selects class having the best prediction accuracy (F1-score)
- **Max-confidence :** Selects class having the highest soft-max accumulation of all its votes [3]

Experimental details

Data

Bark-101 original image count

- **1292 train**
- **1295 test**

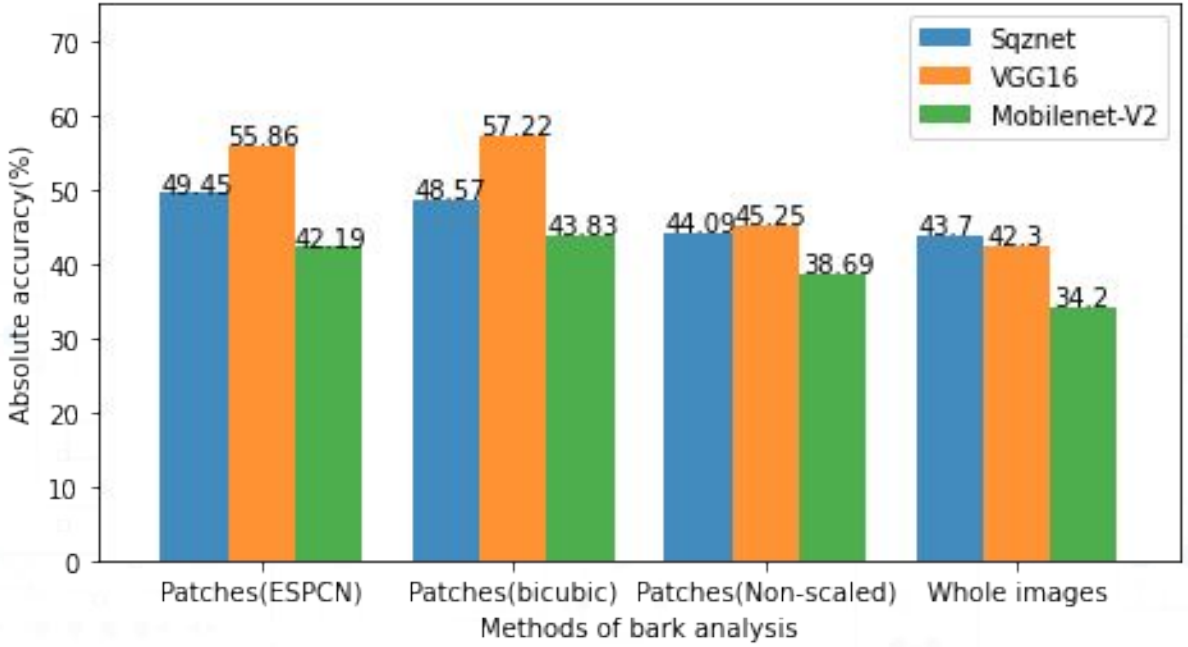
Count of patches (25% train as validation)

Source image	Train	Test	Validation
Non-scaled	3156	1051	4164
Up-scaled	74799	24932	99107

Evaluation Metrics

- **Patch-level accuracy : Estimates how many test patches are correctly classified**
- **Absolute accuracy : Estimates how many test images (1295) are correctly classified after majority voting**

Results



- **Patch-based CNN classification outperforms traditional methods (prior work on Bark-101 using LBP-like filters & SVM classifiers achieved a best accuracy of 41.9 %) & baseline settings with whole images**
- **Both methods of image up-scaling have comparable performance in our study**

Comparison of tie-breaking strategies

- **Differences among tie-breaking strategies substantial only when there is a high number of ties**
- **When count of ties is high, breaking ties using *Maximum Confidence sum* gives the best accuracy**

CNN model	Patch-level accuracy(%)	Absolute accuracy(%) by Majority Voting			
		Random selection	Max confidence	Class proportions	Class F1-scores
Squeezenet	47.84	43.47	**44.09**	43.17	43.63
VGG16	47.48	44.40	**45.25**	44.32	44.09
MobilenetV2	41.83	37.61	**38.69**	36.68	37.22

1 With patches from non-scaled images

Patch Method	Squeezenet	VGG16	MobilenetV2
Non-Scaled Original	217	283	274
Upscaled by Bicubic	52	45	45
Upscaled by ESPCN	50	46	63

Count of ties encountered

CNN model	Patch-level accuracy(%)	Absolute accuracy(%) by Majority Voting			
		Random	Max confidence	Class proportions	Class F1-scores
Squeezenet	35.69	48.32	**48.57**	48.11	48.19
VGG16	41.04	57.21	56.99	**57.22**	57.14
MobilenetV2	33.36	43.73	43.60	**43.83**	43.60

2 With patches from images upscaled by bicubic interpolation

CNN model	Patch-level accuracy(%)	Absolute accuracy(%) by Majority Voting			
		Random	Max confidence	Class proportions	Class F1-scores
Squeezenet	34.85	49.40	49.38	**49.45**	49.38
VGG16	39.27	**55.86**	55.75	55.76	55.75
MobilenetV2	32.12	**42.19**	41.78	41.93	41.85

3 With patches from images upscaled by ESPCN

Conclusions & future work

- **Patch-based CNN classification outperforms traditional methods**
- **Image up-scaling can be useful, particularly when there is huge variety of image dimensions in the dataset as ours**
- **Incorporating tie-breaking strategies in majority voting is important when there are several ties**
- **In future, the proposed methodology would be applied to other plant organs (leaf, flower, fruit) and extended to develop a multi-modal tree recognition model**
- **The feasibility of the approach would be tested on mobile platforms**

References

[1] Ratajczak, Rémi, et al. "Efficient Bark Recognition in the Wild." International Conference on Computer Vision Theory and Applications (VISAPP) 2019.

[2] Shi, Wenzhe, et al. "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[3] Kokkinos, Y., & Margaritis, K. G. (2014, September). “Breaking ties of plurality voting in ensembles of distributed neural network classifiers using soft max accumulations.” In IFIP International Conference on Artificial Intelligence Applications and Innovations (pp. 20-28). Springer, Berlin, Heidelberg.

[4] Wan, Y.Y., Du, J.X., Huang, D.S., Chi, Z., Cheung, Y.M., Wang, X.F., Zhang, G.J. “Bark texture feature extraction based on statistical texture analysis.” In: Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, 2004. pp. 482-485. IEEE (2004)

[5] Bertrand, S., Ameur, R.B., Cerutti, G., Coquin, D., Valet, L., Tougne, L.. “Bark and leaf fusion systems to improve automatic tree species recognition.” Ecological Informatics 46, 57-73 (2018)

[6] Huang, Z.K., Huang, D.S., Du, J.X., Quan, Z.H., Guo, S.B. “Bark classification based on gabor filter features using rbpnn neural network.” In: International conference on neural information processing. pp. 80-87. Springer (2006).

[7] Mizoguchi, T., Ishii, A., Nakamura, H., Inoue, T., Takamatsu, H. “Lidar-based individual tree species classification using convolutional neural network.” In: Videometrics, Range Imaging, and Applications XIV. vol. 10332, p. 103320O. International Society for Optics and Photonics (2017).

[8] Carpentier, M., Giguere, P., Gaudreault, J. “Tree species identification from bark images using convolutional neural networks.” In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 1075-1081. IEEE (2018).