Dealing with challenging images using Noisy Student and Focal loss in detection of trees near electric wires.

Artur Andre A. M. Oliveira[1,2] and **Roberto Hirata Jr.**[1] and **Zhangyang Wang**[2]

arturao@ime.usp.br - hirata@ime.usp.br - atlaswang@utexas.edu

[1]Institute of Mathematics and Statistics, University of São Paulo
[2]Visual Informatics Group, University of Texas at Austin

IME INSTITUTO DE MATEMÁTICA E ESTATÍSTICA UNIVERSIDADE DE SÃO PAULO

TEXAS The University of Texas at Austin

Motivation

- Electrical wires and tall vegetation may lead to problems like wildfires and disruption of the energy distribution [2].
- A citizen can request to manage the vegetation but it should not depend on her/him.
- Automatically detecting trees close to wires enable a faster management of the trees and reduce the risk of hazardous problems.
- The urban scene is complex and can be hard to determine there is a need for vegetation management or not.

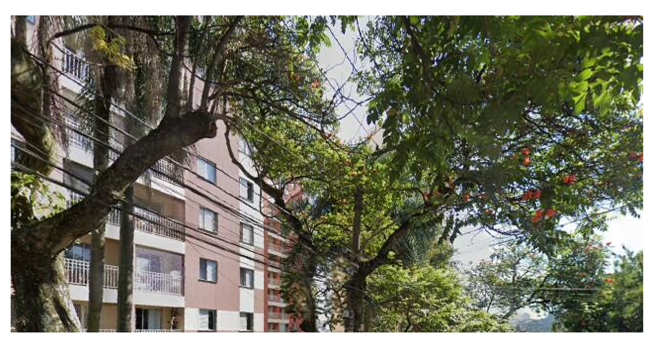

Fig. 1: Image from Google Street View with trees in contact with electrical wires.

Fig. 2: Image from Google Street View with bad visibility due to sun glare.

Methodology

- We collected 50k urban images at the street-level using Google Street View.
- 11k images were labeled and used for training, validation and test.
- The other images were left unlabeled and used to train an EfficientNetB2 network with the Noisy Student training protocol [3] and Focal loss [1] (FL).
- We introduce an auxiliar label to characterize challenging images.

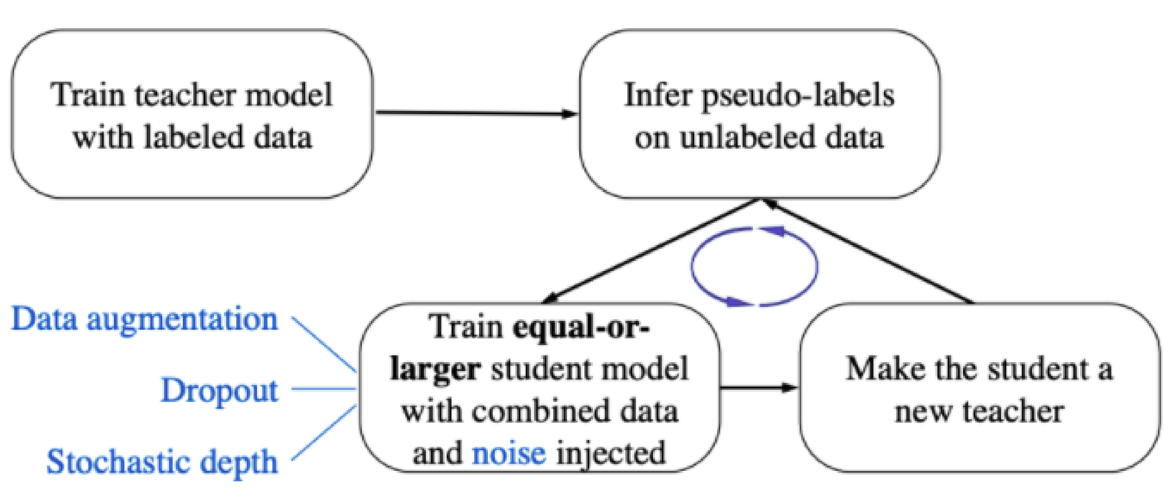

Fig. 3: The Noise Student protocol. Image from [3]

Results

- The confidence of the model is characterized by the maximum probability in the prediction vector given a sample.
- The following graphs show the number of images (in)correctly classified with a given confidence. Notice that the y-axis is log-scaled for better visualization.

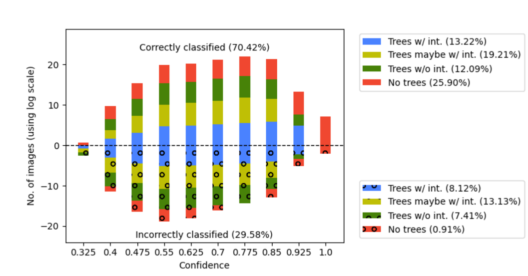

Fig. 4: Cross-entropy loss (CE) results for the training partition.

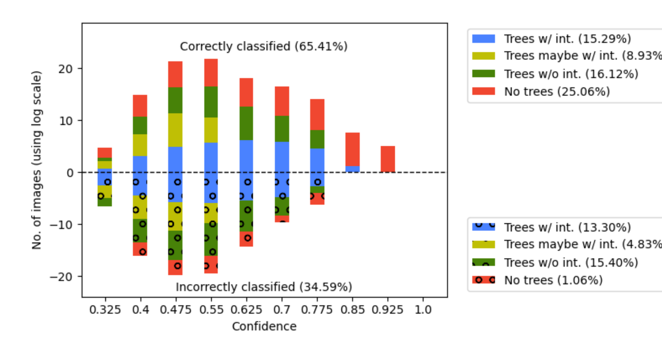

Fig. 5: Focal loss results for the training partition.

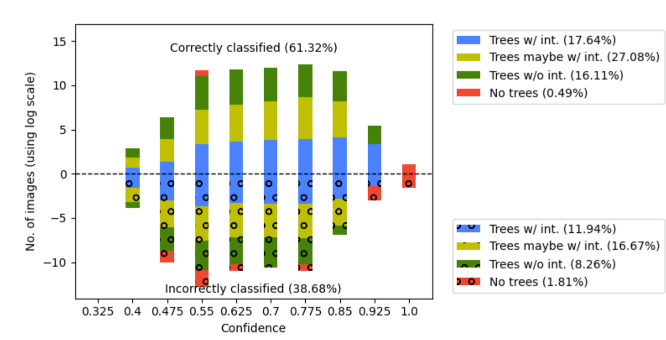

Fig. 6: Focal loss results for the validation partition.

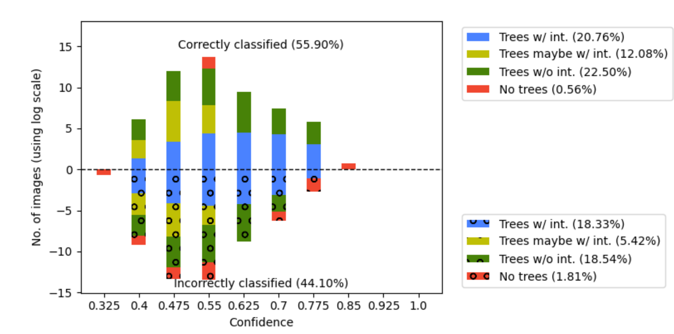

Fig. 7: Cross-entropy loss for the validation partition.

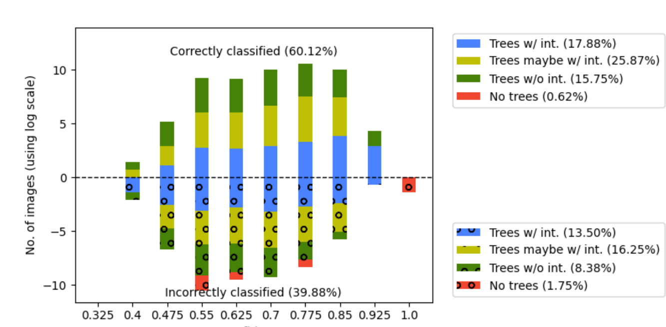

Fig. 8: Cross-entropy loss results for the test partition.

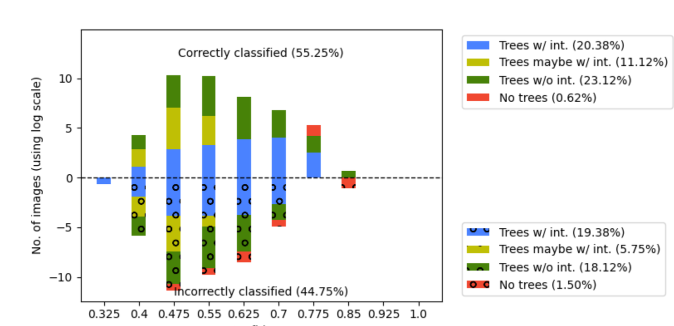

Fig. 9: Focal loss results for the test partition.

- Fig. 10 shows the confusion matrix over the test dataset with Focal Loss: (0) Trees w/ an Intersection, (1) Trees maybe w/ Intersection, (2) Trees w/o Intersection and (3) No trees.
- True positive rate for class (0): 80.3%.
- True negative rate with respect to the union of the classes (2) and (3) is 71.7%.

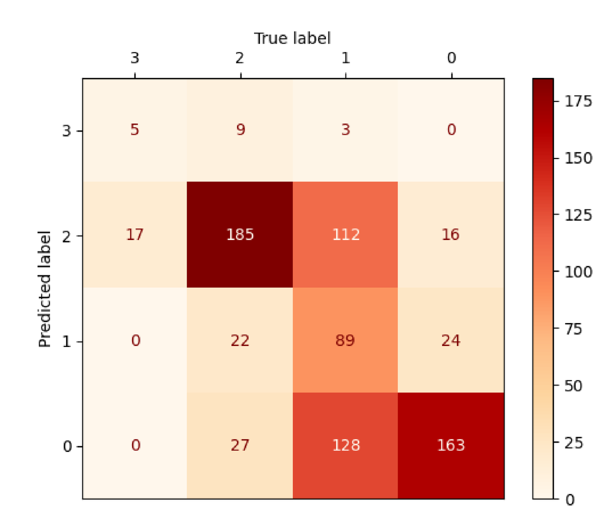

Fig. 10: Confusion matrix over the test dataset

Discussion

- Every challenging image has a low confidence (i.e. equal or below 55%) under FL,
- Furthermore, most of the incorrect cases also have a low confidence score.
- An active learning system based on the confidence level can collect additional images for a given location and increase the overall accuracy.

Fig. 11: Image from Google Street View. The confidence for this false negative dropped from 59% to 49% with FL.

Fig. 12: Image from Google Street View. CE Misclassified as maybe with intersection and corrected by FL as without intersection by FL.

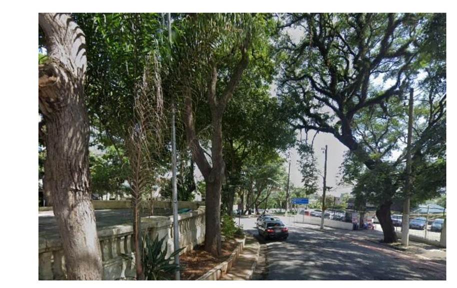

Fig. 13: Image from Google Street View. FL misclassification as without intersection but with low confidence (40%).

Conclusion

- The method presented here has high true positive and negative rates,
- Despite the low overall accuracy, based on the low confidence, an active learning system could be used to collect alternative images.

Acknowledgements

This research is part of the INCT of the Future Internet for Smart Cities funded by CNPq 465446/2014-0, FAPESP 14/50937-1 and 15/24485-9. The authors acknowledge São Paulo Research Foundation (FAPESP) 18/10767-0 and 15/22308-2. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

References

[1] Tsung-Yi Lin et al. “Focal loss for dense object detection”. In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2980–2988.

[2] Jun Ma et al. “Real-time detection of wildfire risk caused by powerline vegetation faults using advanced machine learning techniques”. In: *Advanced Engineering Informatics* 44 (2020), p. 101070.

[3] Qizhe Xie et al. “Self-training with noisy student improves imagenet classification”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 10687–10698.