Artificial Intelligence for Tree Risk Protection

This transcript is auto-generated. Please direct any questions to isa@isa-arbor.com.

Music: [00:00:00] In every country, you know, we can work together and learn what we need to beat the challenge.

Tinitia Price: Hello, and welcome to the ISA Conference rewind video podcast series I'm Tinitia Price Instructional Designer with the International Society of Arboriculture. Today, ISA is proud to bring you a presentation by Ryan Suttle and Nasko Apostolov about their study that explores how artificial intelligence can be used to predict treatment.

This talk was originally given at the 2021 ISA Virtual Conference. So the view seen here are those of the presenters. If you're fascinated by how technology can be used to enrich the study of arboriculture, you'll find something interesting here. Now, sit back and enjoy. Alright.

Ryan Suttle: Right, hello, everybody. Thank you for coming to our talk. My name's Ryan Suttle, Certified Arborist from the Department of Environmental Conservation at UMass Amherst.

Nasko Apostolov: And I'm Nasko Apostolov, a masters students in Department of Civil and Environmental Engineering at, here at UMass as well.

Ryan Suttle: And we're going to talk about the use of artificial intelligence for tree risk prediction today. So let's dive into it. Just give us a brief outline so we have kinda the bones of our talk today. I'll start with an introduction into kind of the motivation for this study as well as some broad use of artificial intelligence, and then I'll go over on the data collection and description for what we used. And at that point, Nasko will take over and really get into kinda the nitty gritty of the artificial intelligence [00:02:00] side, talk about some of our results, which I think are, are pretty exciting. And we'll get into some discussion, and hopefully have time for questions at the end.

All right. So, to start us off, why study trees at all? I'm sure I'm preaching to the choir here, today. But trees serve a lot of benefits in urban areas. They can decrease ambient temperature through evapotranspiration cooling building energy use reduction, carbon sequestration and storage all three of which are only gonna become more important as our climate changes. They remove air pollutants and can mitigate stormwater runoff, and raise property values where they're planted. Less tangibly and quantifiably, they also have huge aesthetic and cultural value. Any of you watchers from Massachusetts may be familiar with this tree on the left it's The Buttonball Tree in Sunderland. It's the largest tree in the state. And it's just gorgeous. It has, you know, tremendous aesthetic value and cultural value to that town. It's a historical landmark and all of these benefits are, are maximized due to the size of the tree.

However, with all these good things that trees do, there's also risk that they present especially in proximity to overhead utility lines. A large portion of power outages are caused by tree failures. Trees kinda cause outages in two ways, falling into the lines and growing into them failures of individual trees coming down on the lines are by far the largest proportional cause of these outages. And mitigating it has a high direct cost to utility companies in terms of funding tree pruning and removal programs and infrastructure repair. Typically, [00:04:00] utilities will spend millions of dollars annually even small, rural companies. And these outages also present a high indirect cost to the community the customers that utilities serve.

There was a really interesting study done a couple of years ago estimating $8.3 billion of lost revenue in a 10 year period in the State of Connecticut due to tree caused outages alone. Between 2005 and 2015, so this was before we were doing everything online, I suspect that that cost has gone up dramatically in the last couple of years. It could be a, it could be an interesting future study. And all these things can lead to a negative perception of trees from the public and utility companies.

So that kinda frames our research question for today. Can artificial intelligence be used to predict likelihood of tree failure, and, and try to get out ahead of these outage events? There has been research into the efficacy of drive by windshield surveys and walk by surveys, and they're effective at identifying trees with an elevated likelihood of failure. So there's a potential for AI to process images to predict likelihood of failure in a similar way, which would promote efficiency as a potential to reduce costs, especially of large scale tree risk assessment projects and to increase data reliability through standardization of the process.

Hm. So we took the approach to automate tree risk classification using a convolutional neural network that's the type of artificial intelligence we use today. And Nasko will give us some more background on the specifics of that in a little bit. CNNs have been used in risk management before. [00:06:00] Specifically it's been applied to earthquakes and analyzing building structural health.

However, its use in utilize vegetation management is relatively new and complex. And that's because of the required degree of accuracy yeah. As we all know there's millions of trees along distribution rights of way throughout the country. And to assess each one you needed to be pretty accurate to mitigate that risk. There's also a lot of variation source here in terms of different species appearance seasonal trends, like, you know, deciduous trees losing their leaves in the winter. And growth habit.

It has specifically been used in UVM for a couple of different applications, including estimating distance to the conductor from a tree part, identifying density of tree stands along rights of way and correlating higher density stands with incident levels of risk. But there remains a gap in deploying AI for predicting the use of or, excuse me, predicting the risk presented by individual trees. And as we brought up earlier it's really these individual tree failures that are causing a lot of these outages. So it is a really interesting opportunity to, to see if AI can be effective, here.

So now I'll get into our data and description of that. I'm sure everybody has probably seen these two BMPs on the left, here. So our data was collected in Eversource Distribution Rights-of-Way across the State of Massachusetts from, you know, all the way in the West, in the Berkshires, all the way out and onto the Cape 'cause we had quite a few different study sites, here done at a level one risk assessment procedure. And it was a likelihood [00:08:00] failure assessment and a photograph. It was level one because, you know, photograph is, by definition, a, a limited visual inspection of, of the tree in question.

So likelihood of failure you know, as we all know, it's four levels. So we had total of 665 images in this training data set, 459 of which were deemed improbable, 130 of them were, had a likelihood of failure of possible, 76 had a likelihood of failure of probable and imminent was excluded from our model due to the rarity. I think we only had two trees that had an imminent likelihood of failure. And just to give us kind of an idea of what these trees looked like in the field I have some representative photos, here. So this tree on the left this Littleleaf Linden, you know, there's no visible defects from this perspective. There's no dieback in the crown. So that was assessed at an improbable likelihood of failure.

Similarly this large London plane on the right has, you know, strong branch unions, no included bark and this was also given a improbable likelihood of failure. Some Possible trees, here this is a blue spruce, if you follow the trunk up, about three quarters of the way, there's a co-dominant union, there. And there is some dieback in the extremities of the crown. So that union was given a likelihood of failure of possible. And a similar situation, here with the silver maple on the right, there's a low co-dominant union with some included bark and there's also some dieback in the crown extremities. So this one was assessed as possible.

And probable this one is an easy [00:10:00] identification, here, just a stone dead sugar maple, and an American elm, here. On the left, you can see above that crossarm, there's significant dead wood in the crown. All those leaves are actually IV so there's a potential for decay. And these were assessed as probable. And now, Nasko will get into the methodology.

Nasko Apostolov: Thanks, Ryan. To give everyone an idea of the procedure that we used for developing our network like to share our framework with you all. So we used the raw images Ryan mentioned, 665 in total. And we performed some quick processing procedures known as down sampling and horizontal flipping. So this data augmentation was performed in order to increase the size of our data set which is, a requirement for training our model.

Once we obtained the preprocessed images we did, created the classification scenarios where we combined different likelihood of failures, different categories together. So we had a total of four scenarios where the three likelihoods were either independent or grouped, to again compare the performance of the model. And based on the different convolutional architectures that we, that we had, we created some visualizations to see how the what the performance was. And more importantly we also tried and this is part of our ongoing efforts, here, to actually [00:12:00] understand how the model learns and how those predictions are, are created.

To give everyone an idea of what a convolutional neural network is or CNN, for short it is actually based on the image recognition process of our own brains. So it essentially, it maps image data to an output variable in our case it's a it's a class, the probable, possible, or improbable. So having these raw, reprocessed images as input we perform our preprocessing procedure and then we, perform the model training which, essentially is how the model learns to numerically map the image data to the image label that was given.

And the final result is the image class prediction which tells us whether the model correctly predicted an image, which was given a certain label by, by an arborist. So our CNN implementation classifies these images accordingly. On the left, here you can see what the sample, raw input image looks like. Each image has a resolution of 4,030 kilobyte by 3,024 pixels, which is by far quite large we had to actually down sample it for a daily enhanced model performance. So on the [00:14:00] down right, you can see the down sample version.

What, what we also did in an effort to increase the size of the of, of the training set, we also performed the horizontal flipping. So from the total 665 images, we down sampled them to lower resolutions to speed up the computation, process of the model. We used three resolution sets. And then we randomly selected 20% which is equivalent to 133 images. Those were reserved purely for testing purposes so the model is not biased during the training process the stage of the of the training process.

So the model was actually trained on 532 original input images. The horizontal flipping is, this actually creates these mirror images but it doubles the, the training set to over 1,000, which is really what, what we need in order to train our model. Now a little bit about the classification scenarios that we used. So there were four in total. And three of them were binary scenarios, where the first one was the probable improbable case where we completely excluded the possible.

So this wasn't a simplification of the original research question but as you can see this leaves out one of the categories. So we also [00:16:00] decided to do some groupings. We grouped the two high risk categories, the probable and possible together. So they essentially constitute one, one class. And then we're training our model based on that and the improbable class which is left separate.

Another scenario is grouping the possible and the improbable classes together. And then comparing it against the, the probable class. And as you can see, this is just another one of the strategies that we have for comparing the performance and, and really seeing what optimizes, our convolutional neural network. And finally, we have the three categories all separate. So they're, not grouped in any way. It's essentially a harder problem for the network to solve since it has three outputs to label as opposed to just do in, various forms.

The goal of our of training our models is to maximize performance based on a specified metric and also minimize loss. This is not any just unique to this problem, but through artificial intelligence in, in general, it's an optimization problem. So what our model and what we keep track of is the true positives, the correct predictions, that belong to a certain class. The incorrect predictions as well as the, ... so essentially is they're the true positives, false positives false negatives, [00:18:00] which were the incorrect predictions, which do not belong to a particular, to the particular class as w- as well as the, the true negatives.

So the measure of performance that we chose was F1, the F1 score. Which is a metric that keeps track of the sensitivities of those, underlying factors, the true positives and false positives, false negatives. And a score, is in the zero, one range, where a higher score indicates better performance and higher sensitivity. And to give you an idea of what the split is of our total number of images, we have the 1,064 images after performing data augmentation. And 133 unique images, unseen by the model previously before we actually test its, predictive capability.

Now we'll dive into the results. And, and show you how our model performed overall. This graph tells us what the average F1 score is compared to across all four scenarios. We can see that the probable-possible improbable case, where we grouped probable and possible together performed, best across the, across the four. And similarly, the loss, which is the percent, the percent error of the model, so what, what percentage of the classification is actually incorrect and that's, that's what it tracks. [00:20:00] So the best one the best one of the four scenarios was PR/PO_IM. And it resulted in the highest overall F1 scores.

We also included confusion matrices which specific to keep track of the probable, possible, and the improbable predictions. So what essentially, the way we read them, is what, what is the, did the, the true, or the observed class, as opposed to the predicted class by the, by the model? And it tells us what is the, the highest percentage. So, in the case of PR/PO_IM on top right, we can see that the probable possible class was correctly predicted in the highest form compared to the other three cases.

One of the efforts as mentioned previously in answering this question was to actually interpret the results, and not treat the convolutional neural network as a black box, but more of a interpretable tool. So there are many visualization techniques out there that let us study the learning process what does the model deem relevant? What contributes to the final prediction? So in this slide you can see the different visualization techniques that we are currently exploring. And the most relevant one in answering, [00:22:00] a research question is the Grad-CAM, which stands for Gradient-weighted Class Activation Mapping. It's the second one from the left.

And it essentially creates an overlay on the original process input. And it highlights the regions of the image which were most relevant for the final prediction. As you can see, this was a correctly predicted image with a probability of 100%. And it's i- it's a very relevant tool for the visual representation of the network's, interpretability. This is an ongoing effort. So we continue to make progress on this.

Ryan Suttle: And I would like to just kinda point out one thing on this Grad-CAM image, here, is that there is a, a limb on the right-hand side with a very sharp bend which can be, you know, a stress raiser and a weak point under, under a significant load. And this Grad-CAM is picking up that, that area. So it's, it's looking like it's assessing the tree in a similar way to I was in the field, which was really, really exciting to see.

So I'll take over the discussion, here just some kinda summary points, some key takeaways as well as mentioning some future work that we hope to do. And then at the end hopefully we'll have some time for questions. So in summary this pilot study demonstrates that AI can be effectively used to automate tree risk assessment of least in likelihood of failure from photographs. It was able to correctly predict the field assessed likelihood of failure with a high degree of accuracy. [00:24:00] And again, this, you know, in, for, in future stages has the potential to decrease budget demands of large risk assessment projects by highlighting these, these higher or more probable likelihood of failure trees, and potentially identifying them for a future, for a more detailed assessment.

Now, one limitation is a variability among assessors, especially between the probable and possible categories of likelihood of failure. There have been previous studies done showing that improbable and imminent trees most arborists can agree on what those look like with a possible and probable, the two intermediary categories are, are kinda where professional judgment comes in, and, and experience so delineating between those two classes is a limitation that, you know, running more images through this through the CNN would, would help address.

In future work validation with multiple arborists to try to address that, that variability among excesses would be an awesome next step, here. As well as con- quantifying the interpretability of the network's classification process like Nasko described, trying to unpack that black box and see how the CNN is, is reaching its decisions, to see, you know, are, are the areas of interest the same as what the field assessor is seeing in the tree?

And finally we'd just like to, to thank everyone who made this study possible including Eversource Energy, Networks for Accessibility, Resilience, and Sustainability Laboratory and the Department of Civil Environmental Engineering, and the Department of Environmental Conservation [00:26:00] at UMass Amherst. Thank you.

Nasko Apostolov: Thank you.