THE OHIO STATE UNIVERSITY

Confidence-Driven Hierarchical Classification of Cultivated Plant Stresses

Logan Frank[1], Christopher Wiegman[2], Jim Davis[1], Scott Shearer[2]

[1] Department of Computer Science & Engineering
[2] Department of Food, Agricultural, and Biological Engineering
Ohio State University

CVL Computer Vision Lab

Motivation

Plant Stress

Plant Stress

Plant Stress

Billions of bushels of yield loss → Billions of dollars in lost revenue

Plant Stress

Billions of bushels of yield loss → Billions of dollars in lost revenue

How are plant stresses detected and monitored?

Plant Stress

Current monitoring methods involve unsustainable activities:

- Manual labor
- Expensive tests
- Long waiting periods

Plant Stress

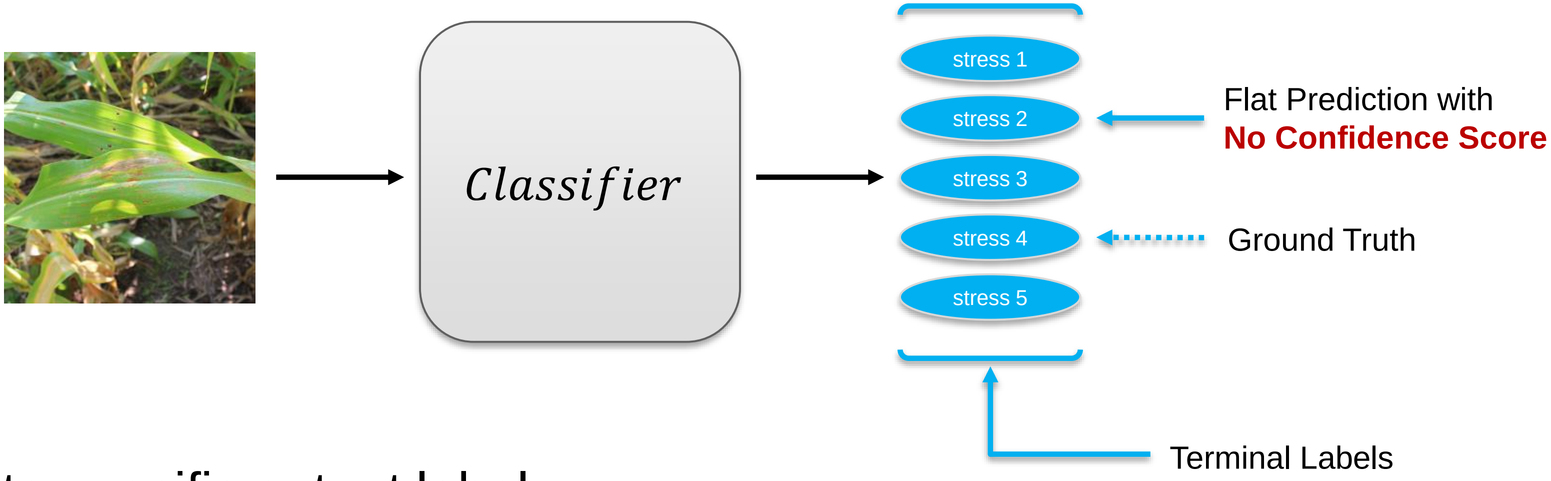
Current monitoring methods involve unsustainable activities:

- Manual labor
- Expensive tests
- Long waiting periods

Could machine learning be used for automatic image classification?

Traditional Approach

Argmax selection on softmax / logit scores



Issues:

- Limited to specific output labels
- No confidence association with predictions
- Lack of domain knowledge

Addressing Shortcomings

- Hierarchies exist naturally within agriculture that define relationships between specific plant stresses and broader plant stress categories
  - Specific stresses can be grouped into more general stress categories in accordance with traditional management strategies
  - We construct semantic trees using the domain knowledge of our agricultural engineering collaborators to model the hierarchy for different datasets

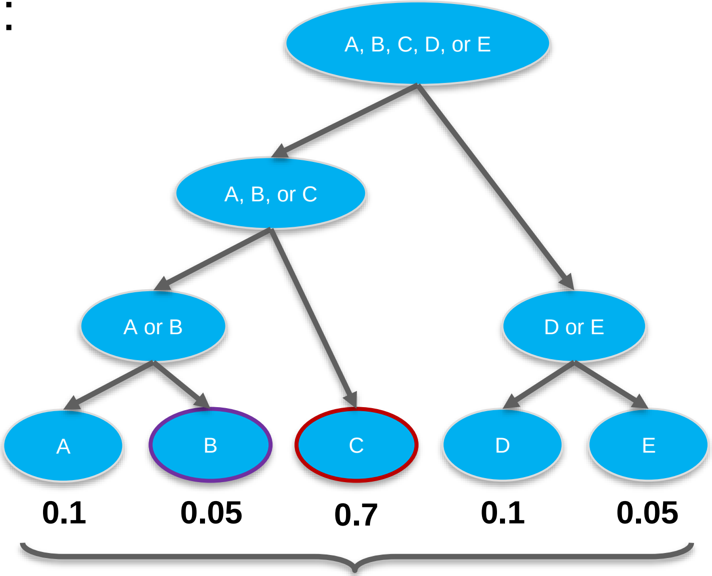
Addressing Shortcomings

- Hierarchies exist naturally within agriculture that define relationships between specific plant stresses and broader plant stress categories
  - Specific stresses can be grouped into more general stress categories in accordance with traditional management strategies
  - We construct semantic trees using the domain knowledge of our agricultural engineering collaborators to model the hierarchy for different datasets
- “Hierarchical Semantic Labeling with Adaptive Confidence”, Davis et al., ISVC 2019:
  - Builds upon a pretrained base classifier
  - Post-processing inference procedure to perform hierarchical reasoning

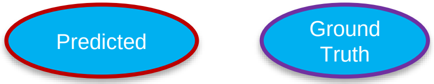
Hierarchical Approach

Hierarchical Approach: Estimation

- For every node $\boldsymbol{l}$ in the hierarchy, compute:
  - *Positive* and *negative* likelihood distributions:
    - $P(s_l \mid l)$
    - $P(s_l \mid \neg l)$
  - Priors $P(l)$ and $P(\neg l)$
  - Posterior probability distributions using Bayes' Rule
    - $P(l \mid s_l) = \frac{P(s_l \mid l)P(l)}{P(s_l \mid l)P(l)+P(s_l \mid \neg l)P(\neg l)}$

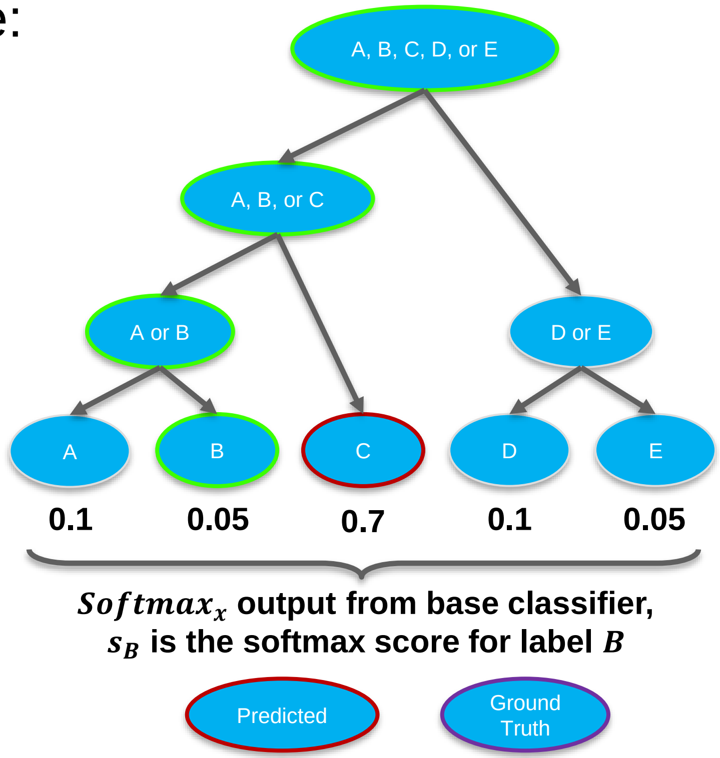


$\boldsymbol{Softmax_x}$ **output from base classifier,**
$\boldsymbol{s_B}$ **is the softmax score for label** $\boldsymbol{B}$



Hierarchical Approach: Estimation

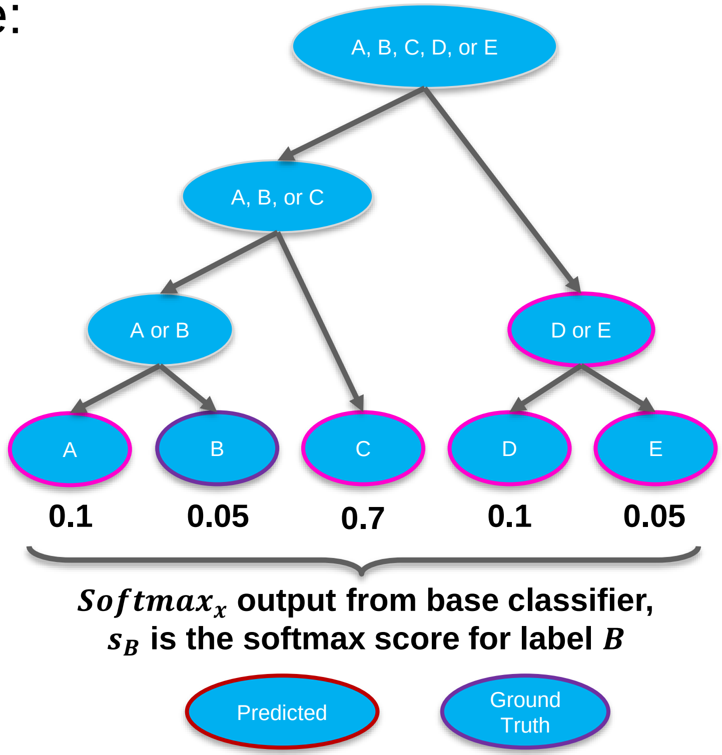
- For every node $\boldsymbol{l}$ in the hierarchy, compute:
  - *Positive* and *negative* likelihood distributions:
    - $P(s_l \mid l)$
    - $P(s_l \mid \neg l)$
  - Priors $P(l)$ and $P(\neg l)$
  - Posterior probability distributions using Bayes' Rule
    - $P(l \mid s_l) = \frac{P(s_l \mid l)P(l)}{P(s_l \mid l)P(l)+P(s_l \mid \neg l)P(\neg l)}$



Davis et al. "Hierarchical Semantic Labeling with Adaptive Confidence" ISVC 2019

Hierarchical Approach: Estimation

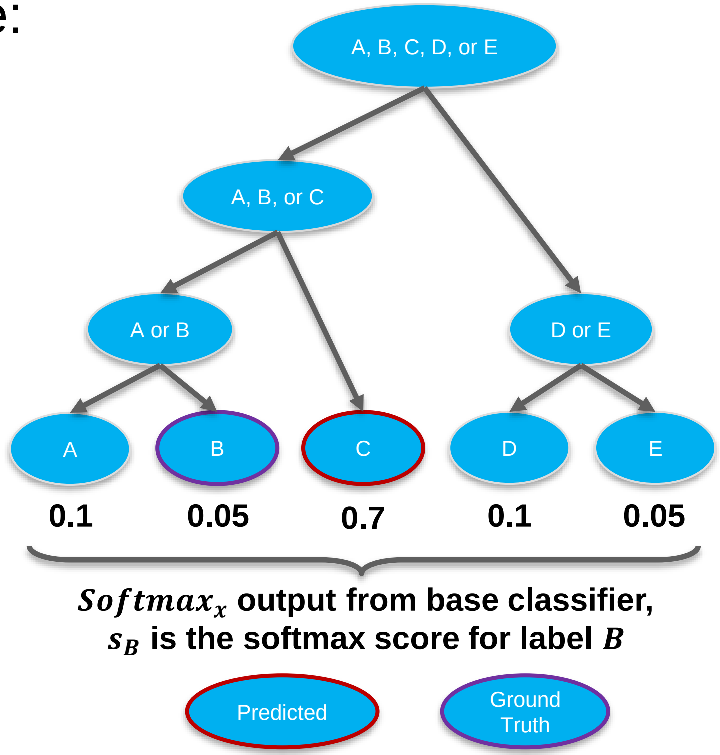
- For every node $\boldsymbol{l}$ in the hierarchy, compute:
  - *Positive* and *negative* likelihood distributions:
    - $P(s_l \mid l)$
    - $P(s_l \mid \neg l)$
  - Priors $P(l)$ and $P(\neg l)$
  - Posterior probability distributions using Bayes' Rule
    - $P(l \mid s_l) = \dfrac{P(s_l \mid l)P(l)}{P(s_l \mid l)P(l)+P(s_l \mid \neg l)P(\neg l)}$



Davis et al. "Hierarchical Semantic Labeling with Adaptive Confidence" ISVC 2019

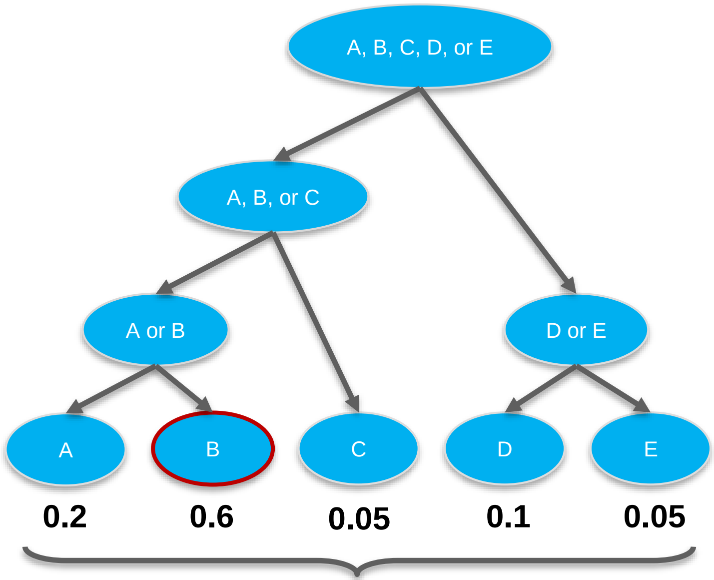
Hierarchical Approach: Estimation

- For every node $\boldsymbol{l}$ in the hierarchy, compute:
  - *Positive* and *negative* likelihood distributions:
    - $P(s_l \mid l)$
    - $P(s_l \mid \neg l)$
  - Priors $P(l)$ and $P(\neg l)$
  - Posterior probability distributions using Bayes' Rule
    - $P(l \mid s_l) = \dfrac{P(s_l \mid l)P(l)}{P(s_l \mid l)P(l) + P(s_l \mid \neg l)P(\neg l)}$



Hierarchical Approach: Inference

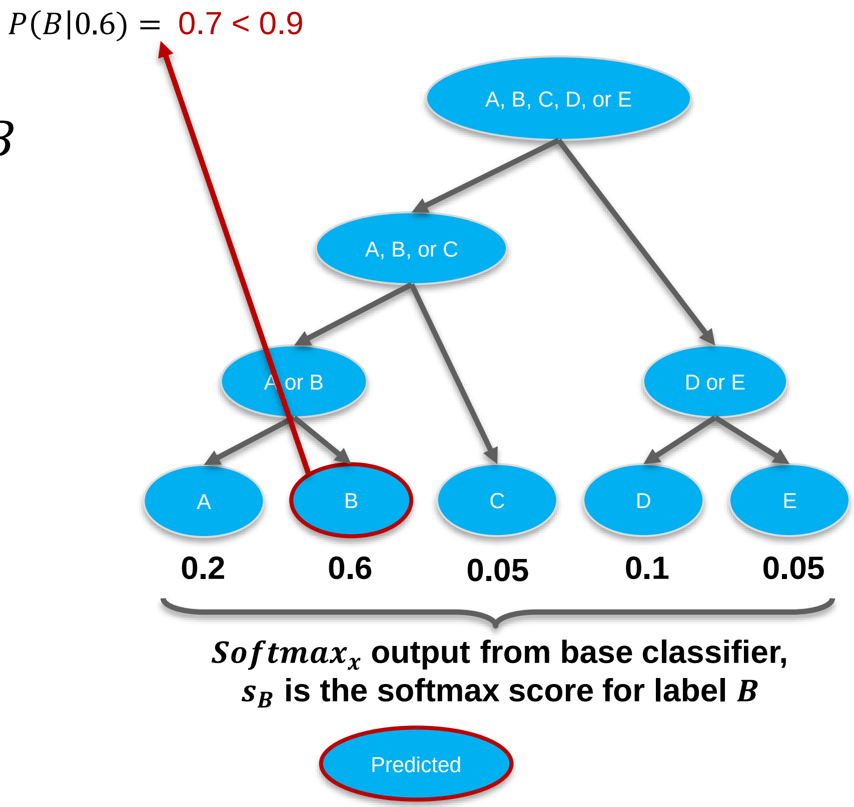
- Given confidence threshold = 0.9



$Softmax_x$ output from base classifier, s_B is the softmax score for label B

Hierarchical Approach: Inference

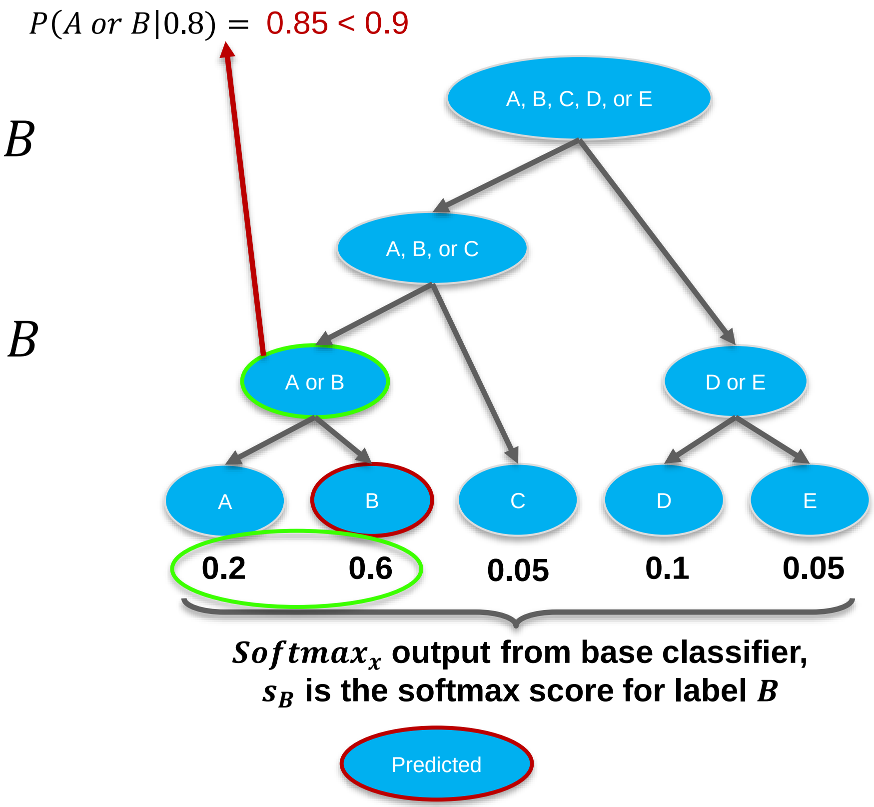
- Given confidence threshold = 0.9
- Analyze the argmax selected label: B
  - $s_B = 0.6$
  - $P(B \mid s_B) = 0.7$



Davis et al. "Hierarchical Semantic Labeling with Adaptive Confidence" ISVC 2019

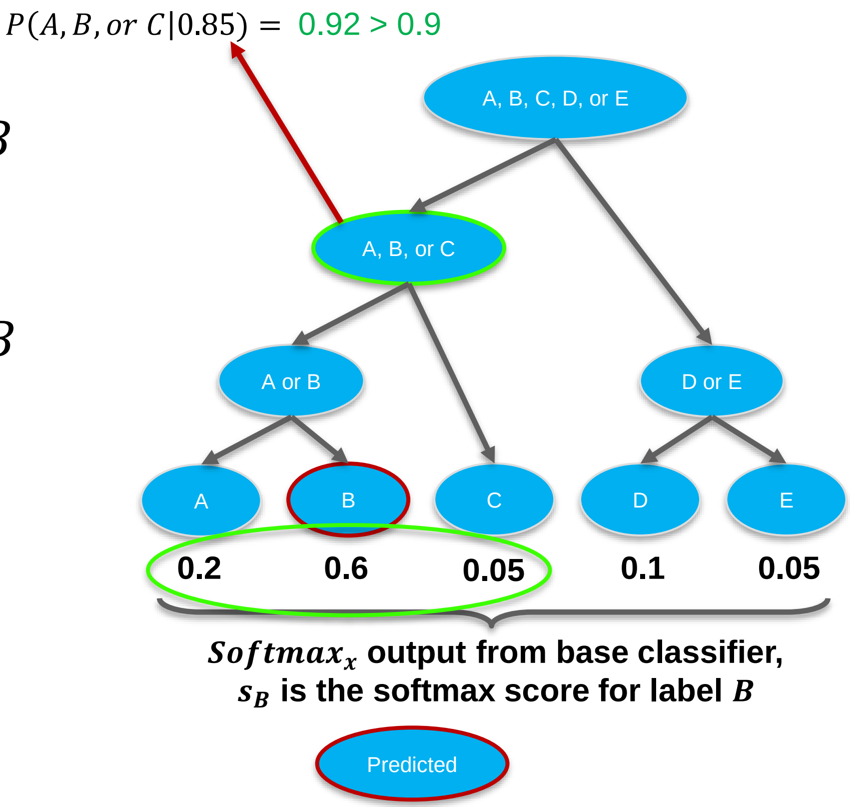
Hierarchical Approach: Inference

- Given confidence threshold = 0.9
- Analyze the argmax selected label: B
  - $s_B = 0.6$
  - $P(B \mid s_B) = 0.7$
- Analyze the immediate parent: A or B
  - $s_{A\ or\ B} = s_A + s_B = 0.8$
  - $P(A\ or\ B \mid s_{A\ or\ B}) = 0.85$



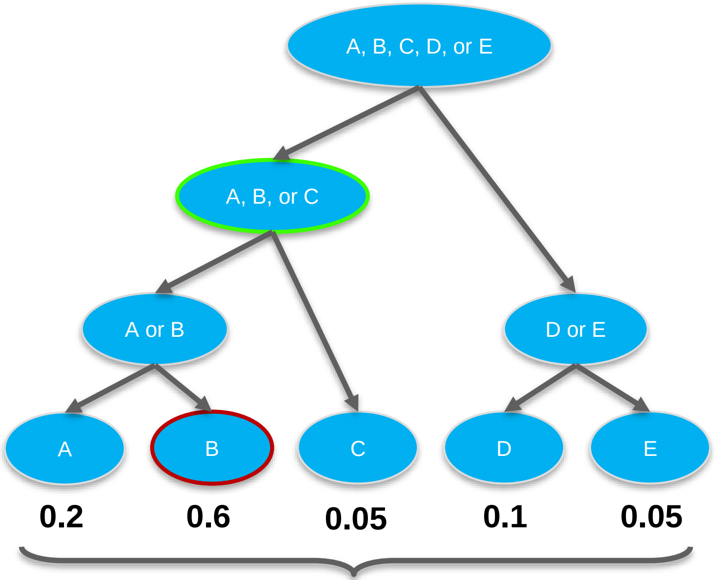
Hierarchical Approach: Inference

- Given confidence threshold = 0.9
- Analyze the argmax selected label: B
  - $s_B = 0.6$
  - $P(B \mid s_B) = 0.7$
- Analyze the immediate parent: A or B
  - $s_{A\ or\ B} = s_A + s_B = 0.8$
  - $P(A\ or\ B \mid s_{A\ or\ B}) = 0.85$
- Analyze the grandparent: $A, B,$ or C
  - $s_{A,\ B, or\ C} = s_A + s_B + s_C = 0.85$
  - $P(A, B, or\ C \mid s_{A,\ B,\ or\ C}) = 0.92$



Hierarchical Approach: Inference

- Given confidence threshold = 0.9
- Analyze the argmax selected label: B
  - $s_B = 0.6$
  - $P(B \mid s_B) = 0.7$
- Analyze the immediate parent: A or B
  - $s_{A\ or\ B} = s_A + s_B = 0.8$
  - $P(A\ or\ B \mid s_{A\ or\ B}) = 0.85$
- Analyze the grandparent: $A, B,$ or C
  - $s_{A,\ B, or\ C} = s_A + s_B + s_C = 0.85$
  - $P(A, B, or\ C \mid s_{A,\ B,\ or\ C}) = 0.92$
- Our final prediction label is: $A, B,$ or C



$Softmax_x$ output from base classifier,
s_B is the softmax score for label B

Applied Experiments

Datasets

- Tomato subset of PlantVillage
  - Widely used in agricultural community
  - 9 stress classes and healthy
  - 256x256 RGB images
- OSU Corn
  - 10 stress classes and healthy
  - 4K resolution RGB images
- OSU Soybean
  - 5 stress classes and healthy
  - 4K resolution RGB images

Plant Stress Relational Trees

Tomato:

<table>
<tr><td colspan="10">Unknown</td></tr>
<tr><td rowspan="5">healthy</td><td colspan="9">Stressed</td></tr>
<tr><td rowspan="3">spider mites</td><td rowspan="3">bacterial spot</td><td colspan="2">Virus</td><td colspan="5">Fungal / Oomycete</td></tr>
<tr><td rowspan="2">mosaic virus</td><td rowspan="2">yellow leaf curl virus</td><td colspan="3">Hemi-Biotroph</td><td colspan="2">Necrotroph</td></tr>
<tr><td>late blight</td><td>septoria leaf spot</td><td>leaf mold</td><td>target spot</td><td>early blight</td></tr>
</table>

Corn:

<table>
<tr><td colspan="11">Unknown</td></tr>
<tr><td rowspan="6">healthy</td><td colspan="10">Stressed</td></tr>
<tr><td colspan="5">Biotic</td><td colspan="5">Abiotic</td></tr>
<tr><td rowspan="3">holcus spot</td><td rowspan="3">corn borer</td><td colspan="3">Fungal</td><td rowspan="3">herbicide sensitivity</td><td colspan="4">Nutrient Stress</td></tr>
<tr><td rowspan="2">common rust</td><td colspan="2">Necrotrophic</td><td rowspan="2">nitrogen burn</td><td colspan="3">Nutrient Deficiency</td></tr>
<tr><td>grey leaf spot</td><td>northern corn leaf blight</td><td>phosphorus deficiency</td><td>nitrogen deficiency</td><td>magnesium / potassium deficiency</td></tr>
</table>

Soybean:

<table>
<tr><td colspan="6">Unknown</td></tr>
<tr><td rowspan="5">healthy</td><td colspan="5">Stressed</td></tr>
<tr><td rowspan="3">dicamba damage</td><td colspan="4">Biotic</td></tr>
<tr><td rowspan="2">bacterial blight / phyllosticta</td><td rowspan="2">insect damage</td><td colspan="2">Fungal</td></tr>
<tr><td>sudden death syndrome</td><td>frogeye leaf spot</td></tr>
</table>

Experiments: Evaluation

	Tomato					
	Base	50%	80%	85%	90%	95%
C-Persist	1.0	.91	.71	.66	.60	.44
C-Withdrawn	-	.00	.01	.02	.02	.03
C-Soften	-	.09	.28	.32	.38	.53
IC-Remain	1.0	.81	.29	.23	.22	.11
IC-Withdrawn	-	.00	.05	.05	.05	.08
IC-Reform	-	.19	.67	.71	.73	.80
avg-sIG	-	.78	.65	.61	.58	.49
% Valid (¬root)	100	99.8	98.6	97.6	97.5	96.6
Accuracy	82.1	86.0	94.6	95.8	96.1	98.2

	Corn					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.89	.70	.49	.46	.41
C-Withdrawn	-	.01	.06	.13	.13	.13
C-Soften	-	.10	.24	.37	.41	.46
IC-Remain	1.00	.77	.39	.27	.27	.13
IC-Withdrawn	-	.02	.21	.26	.26	.26
IC-Reform	-	.21	.40	.48	.48	.61
avg-sIG	-	.65	.56	.45	.44	.40
% Valid (¬root)	100	98.6	87.9	80.9	80.9	80.9
Accuracy	68.8	74.7	87.7	92.4	92.4	95.8

	Soybean					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.92	.87	.87	.76	.33
C-Withdrawn	-	.04	.05	.07	.18	.18
C-Soften	-	.05	.07	.06	.06	.49
IC-Remain	1.00	.59	.38	.34	.28	.13
IC-Withdrawn	-	.09	.20	.23	.29	.29
IC-Reform	-	.32	.43	.43	.43	.58
avg-sIG	-	.76	.72	.72	.63	.40
% Valid (¬root)	100	96.3	93.0	91.9	82.1	82.1
Accuracy	80.0	81.5	84.5	85.1	85.5	98.7

Experiments: Evaluation

	Tomato					
	Base	50%	80%	85%	90%	95%
C-Persist	1.0	.91	.71	.66	.60	.44
C-Withdrawn	-	.00	.01	.02	.02	.03
C-Soften	-	.09	.28	.32	.38	.53
IC-Remain	1.0	.81	.29	.23	.22	.11
IC-Withdrawn	-	.00	.05	.05	.05	.08
IC-Reform	-	.19	.67	.71	.73	.80
avg-sIG	-	.78	.65	.61	.58	.49
% Valid (¬root)	100	99.8	98.6	97.6	97.5	96.6
Accuracy	82.1	86.0	94.6	95.8	96.1	98.2

	Corn					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.89	.70	.49	.46	.41
C-Withdrawn	-	.01	.06	.13	.13	.13
C-Soften	-	.10	.24	.37	.41	.46
IC-Remain	1.00	.77	.39	.27	.27	.13
IC-Withdrawn	-	.02	.21	.26	.26	.26
IC-Reform	-	.21	.40	.48	.48	.61
avg-sIG	-	.65	.56	.45	.44	.40
% Valid (¬root)	100	98.6	87.9	80.9	80.9	80.9
Accuracy	68.8	74.7	87.7	92.4	92.4	95.8

	Soybean					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.92	.87	.87	.76	.33
C-Withdrawn	-	.04	.05	.07	.18	.18
C-Soften	-	.05	.07	.06	.06	.49
IC-Remain	1.00	.59	.38	.34	.28	.13
IC-Withdrawn	-	.09	.20	.23	.29	.29
IC-Reform	-	.32	.43	.43	.43	.58
avg-sIG	-	.76	.72	.72	.63	.40
% Valid (¬root)	100	96.3	93.0	91.9	82.1	82.1
Accuracy	80.0	81.5	84.5	85.1	85.5	98.7

Experiments: Evaluation

	Tomato					
	Base	50%	80%	85%	90%	95%
C-Persist	1.0	.91	.71	.66	.60	.44
C-Withdrawn	-	.00	.01	.02	.02	.03
C-Soften	-	.09	.28	.32	.38	.53
IC-Remain	1.0	.81	.29	.23	.22	.11
IC-Withdrawn	-	.00	.05	.05	.05	.08
IC-Reform	-	.19	.67	.71	.73	.80
avg-sIG	-	.78	.65	.61	.58	.49
% Valid (¬root)	100	99.8	98.6	97.6	97.5	96.6
Accuracy	82.1	86.0	94.6	95.8	96.1	98.2

	Corn					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.89	.70	.49	.46	.41
C-Withdrawn	-	.01	.06	.13	.13	.13
C-Soften	-	.10	.24	.37	.41	.46
IC-Remain	1.00	.77	.39	.27	.27	.13
IC-Withdrawn	-	.02	.21	.26	.26	.26
IC-Reform	-	.21	.40	.48	.48	.61
avg-sIG	-	.65	.56	.45	.44	.40
% Valid (¬root)	100	98.6	87.9	80.9	80.9	80.9
Accuracy	68.8	74.7	87.7	92.4	92.4	95.8

	Soybean					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.92	.87	.87	.76	.33
C-Withdrawn	-	.04	.05	.07	.18	.18
C-Soften	-	.05	.07	.06	.06	.49
IC-Remain	1.00	.59	.38	.34	.28	.13
IC-Withdrawn	-	.09	.20	.23	.29	.29
IC-Reform	-	.32	.43	.43	.43	.58
avg-sIG	-	.76	.72	.72	.63	.40
% Valid (¬root)	100	96.3	93.0	91.9	82.1	82.1
Accuracy	80.0	81.5	84.5	85.1	85.5	98.7

Experiments: Evaluation

	Tomato					
	Base	50%	80%	85%	90%	95%
C-Persist	1.0	.91	.71	.66	.60	.44
C-Withdrawn	-	.00	.01	.02	.02	.03
C-Soften	-	.09	.28	.32	.38	.53
IC-Remain	1.0	.81	.29	.23	.22	.11
IC-Withdrawn	-	.00	.05	.05	.05	.08
IC-Reform	-	.19	.67	.71	.73	.80
avg-sIG	-	.78	.65	.61	.58	.49
% Valid (¬root)	100	99.8	98.6	97.6	97.5	96.6
Accuracy	82.1	86.0	94.6	95.8	96.1	98.2

	Corn					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.89	.70	.49	.46	.41
C-Withdrawn	-	.01	.06	.13	.13	.13
C-Soften	-	.10	.24	.37	.41	.46
IC-Remain	1.00	.77	.39	.27	.27	.13
IC-Withdrawn	-	.02	.21	.26	.26	.26
IC-Reform	-	.21	.40	.48	.48	.61
avg-sIG	-	.65	.56	.45	.44	.40
% Valid (¬root)	100	98.6	87.9	80.9	80.9	80.9
Accuracy	68.8	74.7	87.7	92.4	92.4	95.8

	Soybean					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.92	.87	.87	.76	.33
C-Withdrawn	-	.04	.05	.07	.18	.18
C-Soften	-	.05	.07	.06	.06	.49
IC-Remain	1.00	.59	.38	.34	.28	.13
IC-Withdrawn	-	.09	.20	.23	.29	.29
IC-Reform	-	.32	.43	.43	.43	.58
avg-sIG	-	.76	.72	.72	.63	.40
% Valid (¬root)	100	96.3	93.0	91.9	82.1	82.1
Accuracy	80.0	81.5	84.5	85.1	85.5	98.7

- As confidence increases:

Experiments: Evaluation

	Tomato					
	Base	50%	80%	85%	90%	95%
C-Persist	1.0	.91	.71	.66	.60	.44
C-Withdrawn	-	.00	.01	.02	.02	.03
C-Soften	-	.09	.28	.32	.38	.53
IC-Remain	1.0	.81	.29	.23	.22	.11
IC-Withdrawn	-	.00	.05	.05	.05	.08
IC-Reform	-	.19	.67	.71	.73	.80
avg-sIG	-	.78	.65	.61	.58	.49
% Valid (¬root)	100	99.8	98.6	97.6	97.5	96.6
Accuracy	82.1	86.0	94.6	95.8	96.1	98.2

	Corn					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.89	.70	.49	.46	.41
C-Withdrawn	-	.01	.06	.13	.13	.13
C-Soften	-	.10	.24	.37	.41	.46
IC-Remain	1.00	.77	.39	.27	.27	.13
IC-Withdrawn	-	.02	.21	.26	.26	.26
IC-Reform	-	.21	.40	.48	.48	.61
avg-sIG	-	.65	.56	.45	.44	.40
% Valid (¬root)	100	98.6	87.9	80.9	80.9	80.9
Accuracy	68.8	74.7	87.7	92.4	92.4	95.8

	Soybean					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.92	.87	.87	.76	.33
C-Withdrawn	-	.04	.05	.07	.18	.18
C-Soften	-	.05	.07	.06	.06	.49
IC-Remain	1.00	.59	.38	.34	.28	.13
IC-Withdrawn	-	.09	.20	.23	.29	.29
IC-Reform	-	.32	.43	.43	.43	.58
avg-sIG	-	.76	.72	.72	.63	.40
% Valid (¬root)	100	96.3	93.0	91.9	82.1	82.1
Accuracy	80.0	81.5	84.5	85.1	85.5	98.7

- As confidence increases:
  - Many originally correct predictions are kept at the terminal level

Experiments: Evaluation

	Tomato					
	Base	50%	80%	85%	90%	95%
C-Persist	1.0	.91	.71	.66	.60	.44
C-Withdrawn	-	.00	.01	.02	.02	.03
C-Soften	-	.09	.28	.32	.38	.53
IC-Remain	1.0	.81	.29	.23	.22	.11
IC-Withdrawn	-	.00	.05	.05	.05	.08
IC-Reform	-	.19	.67	.71	.73	.80
avg-sIG	-	.78	.65	.61	.58	.49
% Valid (¬root)	100	99.8	98.6	97.6	97.5	96.6
Accuracy	82.1	86.0	94.6	95.8	96.1	98.2

	Corn					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.89	.70	.49	.46	.41
C-Withdrawn	-	.01	.06	.13	.13	.13
C-Soften	-	.10	.24	.37	.41	.46
IC-Remain	1.00	.77	.39	.27	.27	.13
IC-Withdrawn	-	.02	.21	.26	.26	.26
IC-Reform	-	.21	.40	.48	.48	.61
avg-sIG	-	.65	.56	.45	.44	.40
% Valid (¬root)	100	98.6	87.9	80.9	80.9	80.9
Accuracy	68.8	74.7	87.7	92.4	92.4	95.8

	Soybean					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.92	.87	.87	.76	.33
C-Withdrawn	-	.04	.05	.07	.18	.18
C-Soften	-	.05	.07	.06	.06	.49
IC-Remain	1.00	.59	.38	.34	.28	.13
IC-Withdrawn	-	.09	.20	.23	.29	.29
IC-Reform	-	.32	.43	.43	.43	.58
avg-sIG	-	.76	.72	.72	.63	.40
% Valid (¬root)	100	96.3	93.0	91.9	82.1	82.1
Accuracy	80.0	81.5	84.5	85.1	85.5	98.7

- As confidence increases:
  - Many originally correct predictions are kept at the terminal level
  - Varying levels of softening across the datasets

Experiments: Evaluation

	Tomato					
	Base	50%	80%	85%	90%	95%
C-Persist	1.0	.91	.71	.66	.60	.44
C-Withdrawn	-	.00	.01	.02	.02	.03
C-Soften	-	.09	.28	.32	.38	.53
IC-Remain	1.0	.81	.29	.23	.22	.11
IC-Withdrawn	-	.00	.05	.05	.05	.08
IC-Reform	-	.19	.67	.71	.73	.80
avg-sIG	-	.78	.65	.61	.58	.49
% Valid (¬root)	100	99.8	98.6	97.6	97.5	96.6
Accuracy	82.1	86.0	94.6	95.8	96.1	98.2

	Corn					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.89	.70	.49	.46	.41
C-Withdrawn	-	.01	.06	.13	.13	.13
C-Soften	-	.10	.24	.37	.41	.46
IC-Remain	1.00	.77	.39	.27	.27	.13
IC-Withdrawn	-	.02	.21	.26	.26	.26
IC-Reform	-	.21	.40	.48	.48	.61
avg-sIG	-	.65	.56	.45	.44	.40
% Valid (¬root)	100	98.6	87.9	80.9	80.9	80.9
Accuracy	68.8	74.7	87.7	92.4	92.4	95.8

	Soybean					
	Base	50%	80%	85%	90%	95%
C-Persist	1.00	.92	.87	.87	.76	.33
C-Withdrawn	-	.04	.05	.07	.18	.18
C-Soften	-	.05	.07	.06	.06	.49
IC-Remain	1.00	.59	.38	.34	.28	.13
IC-Withdrawn	-	.09	.20	.23	.29	.29
IC-Reform	-	.32	.43	.43	.43	.58
avg-sIG	-	.76	.72	.72	.63	.40
% Valid (¬root)	100	96.3	93.0	91.9	82.1	82.1
Accuracy	80.0	81.5	84.5	85.1	85.5	98.7

- As confidence increases:
  - Many originally correct predictions are kept at the terminal level
  - Varying levels of softening across the datasets
  - Several originally incorrect predictions are reformed to a correct generalized label

Examples at 90% Confidence

Incorrect & Reformed

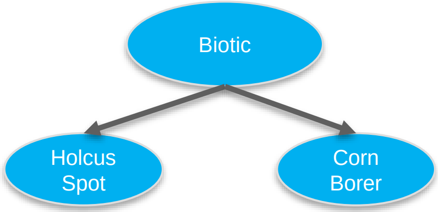
Ground Truth: Holcus Spot

Base Classifier: Corn Borer

Final Label: Biotic

Implication

- We now have correct information on how to proceed with treating a stress, maintaining user trust



Examples at 90% Confidence

Correct & Withdrawn

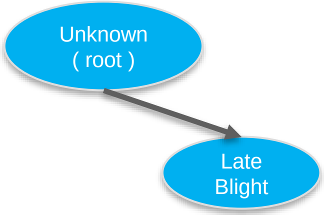
Ground Truth: Late Blight

Base Classifier: Late Blight

Final Label: Unknown (root)

Implication

- Further analysis is required to make an accurate statement regarding the stress
- Prefer to withdraw than to make an incorrect prediction



*Intermediate ancestral nodes omitted

Summary

- Hierarchical classification approach for plant stress identification that addresses many shortcomings of previous works
  - Ability to output generalized labels
  - Each prediction has a confidence guarantee
  - Methods for incorporating domain knowledge
- CNN platform with improved potentially for widespread adoption in the agricultural community
- Future Work
  - Experimenting with different tree structures (e.g., phylogenetic, etc.)
  - Implementing on drones for real time surveillance of crop fields

Code available: https://www.github.com/loganfrank/agriculture

Thank You

Questions? Please come to my Q&A session!

Code available: https://www.github.com/loganfrank/agriculture