Fusion of attribution maps for automated canola blackleg severity scoring

Weihao (Patrick) Li, Emi Tanaka
Australian National University

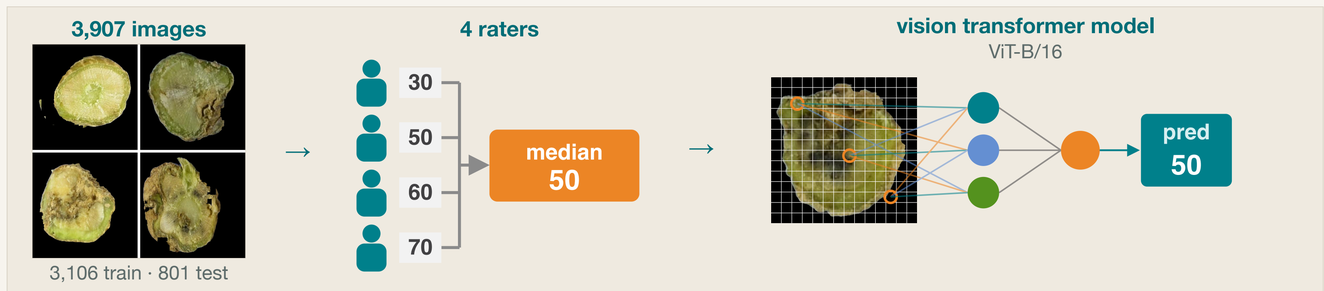
AAGI Analytics for the Australian Grains Industry

Australian National University

Attribution maps for automated canola blackleg severity prediction can *disagree* substantially. We develop metrics to quantify their properties and use them to guide a fusion function that combines the maps into a single summary.

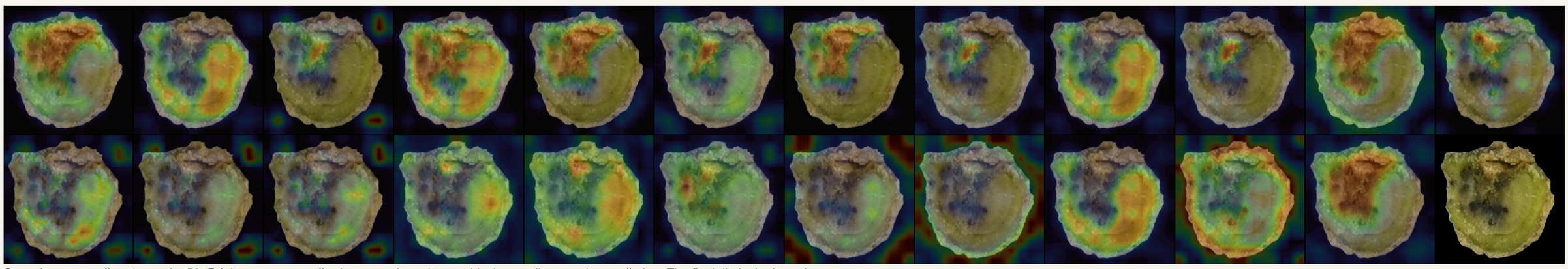
The existing CSIRO pipeline.

A vision transformer (ViT) model predicts canola blackleg severity from segmented stem images scored by four raters. Our goal: reliable attribution maps that explain predictions.



One prediction, many different explanations.

Fifteen attribution methods produce 23 maps for the same prediction, yet disagree on which regions contribute to it.



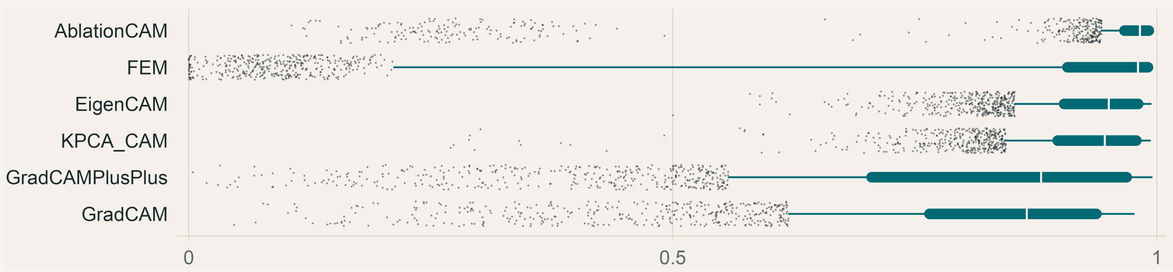
Same image, predicted severity 50. Brighter = more attribution on regions that positively contribute to the prediction. The final tile is the input image.

Define what “good” means. Then optimise it.

Metrics tailored to this task expose how unstable individual attribution maps are. That instability is why we fuse rather than select.

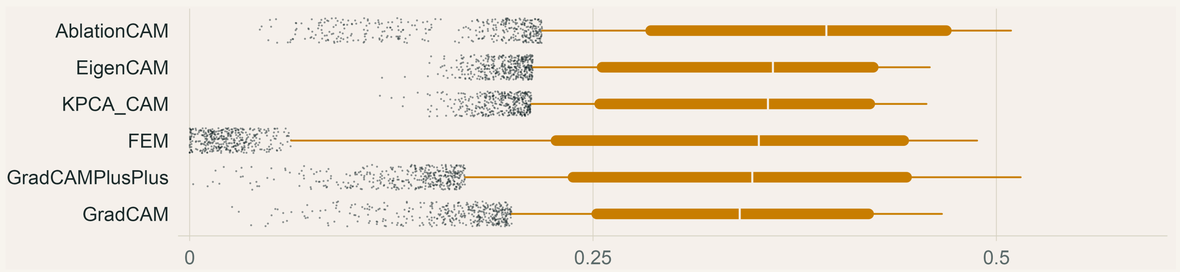
Foreground mass $M_{\text{fg}} = \sum_{i,j} L'_{ij} F_{ij}$

Does attribution stay on the stem?



Darkness-weighted mass $M_{\text{dark}} = \sum_{i,j} L'_{ij} F_{ij} D_{ij}$

Does it favour dark lesion pixels?



L' = normalised attribution · F = stem mask · D = pixel darkness. Six leading methods, 3,907 images each, ordered by median. Box: interquartile range. Whiskers: 10th to 90th percentile. Dots: images below the 10th percentile.

LEARN THE WEIGHTS

$$L(\alpha) = -\bar{M}(\alpha) + \lambda R(\alpha)$$

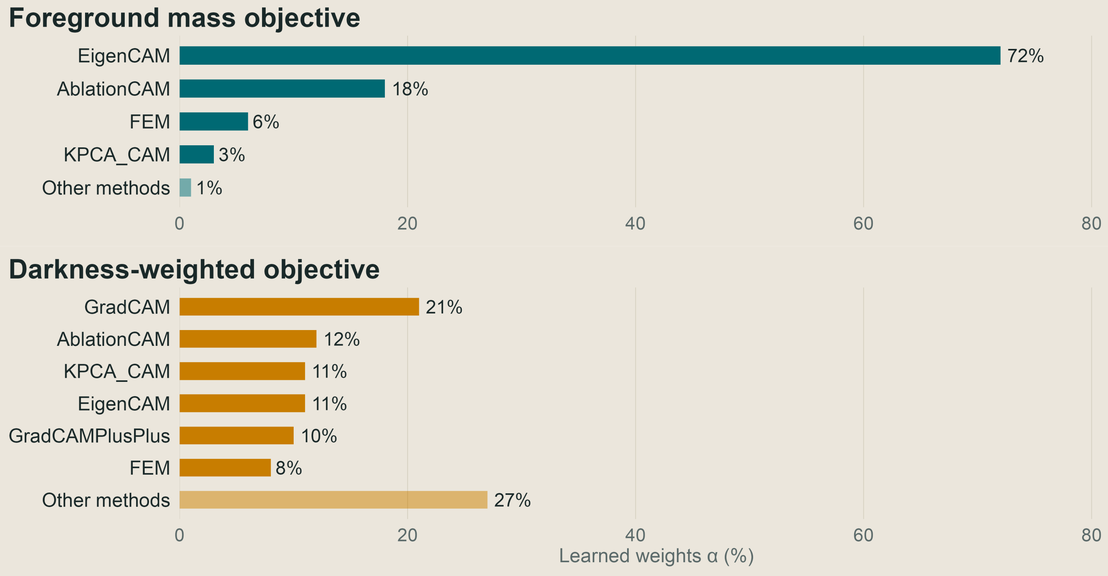
$\bar{M}$ = mean quality · R = SD across images · weights fitted on training data

→

FUSED MAP

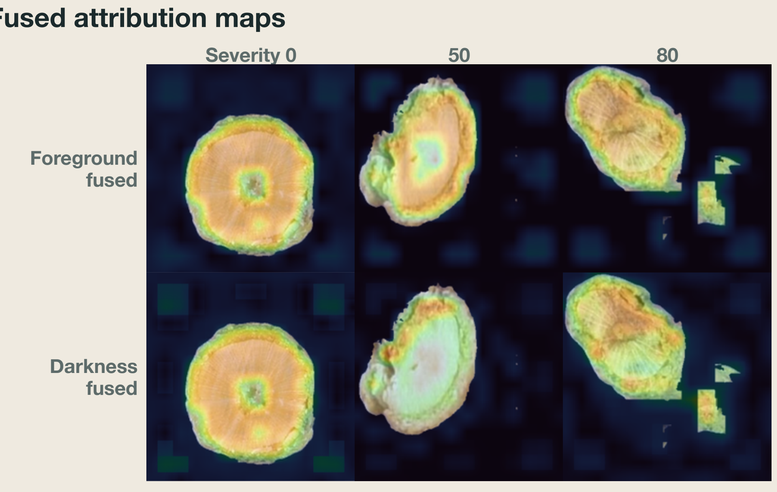
$$L_f = \sum_k \alpha_k L_k$$

Foreground mass objective



Darkness-weighted objective

Fused attribution maps



At severity 0, with no lesions present, attribution spreads across the foreground. At high severity, foreground weighting keeps attribution within the stem, while **darkness weighting highlights dark lesion pixels**.

Thanks to **Luke Barrett** and **Saba Rabab** (CSIRO) for the data, trained ViT model and field metadata.

GRDC Grains Research & Development Corporation · Adelaide University · Curtin University · The University of Queensland Australia

The support of the Grains Research and Development Corporation (GRDC) provided through the Analytics for the Australian Grains Industry (AAGI) project is acknowledged.