Using R Based Image Analysis to Quantify Rust on Perennial Ryegrass

UNIVERSITY OF MINNESOTA

Garett Heineck[1], Eric Watkins[1], Jacob Jungers[2] and Ian McNish[2]

(1) Horticultural Science, University of Minnesota, St. Paul, MN, (2) Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN

Introduction

- Crown and stem rust caused by *Puccinia coronata* f. sp. *lolii* and *Puccinia graminis* subsp. *graminicola* are major diseases of perennial ryegrass (*Lolium perenne* L.) when grown for turfgrass, forage, and seed.
- Plant breeders and pathologists often quantify rust severity in the field using the modified Cobb scale, but this method is subjective, labor intensive, and dependent on the skill and experience of the scorer.
- Our objective was to develop a novel, open-source system that couples both ImageJ and R to quantify rust severity on simple RGB images.
- We validated the system's performance with count data and visual ratings.

Materials and Methods

Plant material and image collection

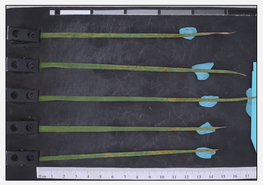
Three image series (n = 50 each) were collected from a diverse collection of 50 perennial ryegrass that varied greatly for rust susceptibility.

Whole plant (WP) images were taken in the field under ambient conditions using monopod 80 cm above the ground during peak rust incidence in Aug 2017.

Five leaf images (FL) were taken in cross pattern from each plant and imaged in the field immediately after placement on an imaging board in a light box lit by 150 cm of high color rendering index LED light strip.

Single leaf (SL) images were randomly selected out of the 5L images.

Data collection efficiency



Efficiency of data collection for crown and stem rust severity		
Visual ratings, (sec plant $^{-1}$)	*WP imaging, (sec plant^{-1})*	*FL sample imaging, (sec plate^{-1})*
6 seconds	**4 seconds**	**140 seconds**

Data collection efficiency was slightly higher for WP imaging compared to visual rating in the field (this does not include image processing time).

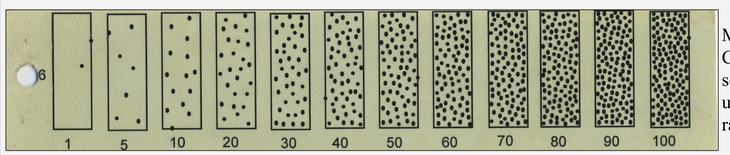
Five leaf images were very time consuming compared to visual or WP data collection.

Visual scoring of images for rust incidence

Ability to visually score rust across the three image series was assessed by a panel of raters (n=9) varying in disease rating experience.

Before rating images for rust severity, three calibration images were used to train raters *a priori* on rust rating using the modified Cobb scale.

Inter-rater reliability was assessed using intraclass correlation (ICC) estimation. Inter-rater reliability determined consistency of rater's scores.

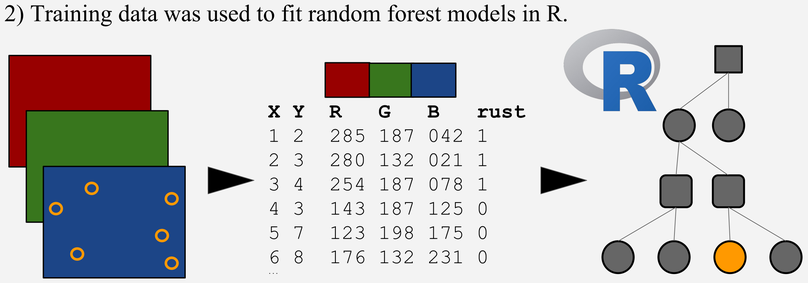


Modified Cobb scale scorecard used by the rater panel.

Image analysis pipeline

1) ImageJ was used to generate training data from selected pixels (non-plant, plant, rust).

2) Training data was used to fit random forest models in R.



3) Background pixels were filtered using the random forest model.

Original image was loaded into R and vectorized into three columns for RGB prediction.

Each pixel was assigned a probability of being biological or not based on votes from the forest of decision trees.

The coordinates of highly probable pixels were saved and the original RGB values are overlaid leaving only the useful foreground pixels.

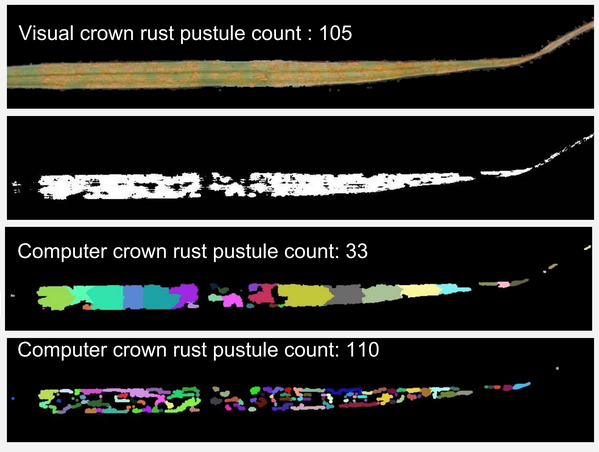
4) Rust severity was quantified using R packages random forest and EBImage.

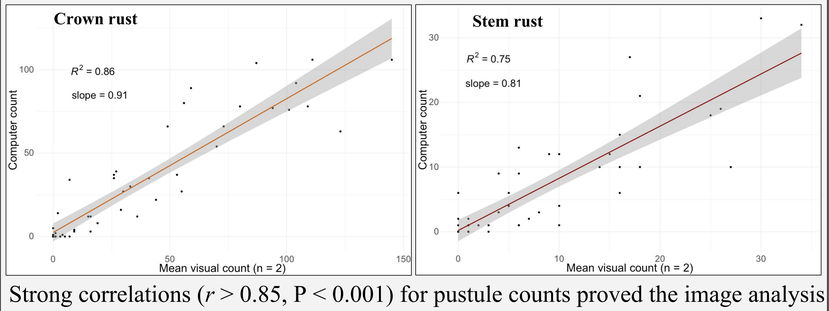
Non-biologically related pixels were filtered with RGB overlay from original image.

Random forest model trained on crown rust with a global threshold applied to each pixel.

EBImage was applied with watershed operation. Low accuracy.

Morphological operation finding circles before water shedding operation. High accuracy.





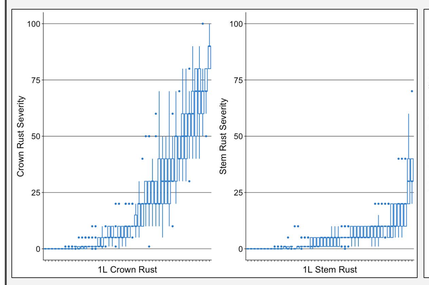
Strong correlations ($r > 0.85$, $P < 0.001$) for pustule counts proved the image analysis system discriminated stem and crown rust on the same leaves.

Results and Discussion

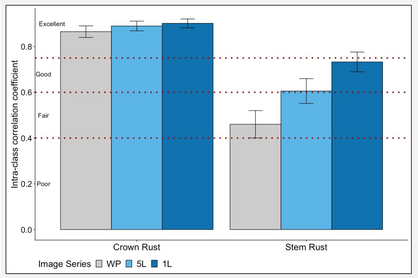
Assessment of rater variability

Agreement on severity for SL images was low, as scores approached 50%, the distribution of the rater scores was maximized.

Crown rust scores across all image series all had excellent ICC values, meaning raters had very consistent perceptions of severity. Rater scores for stem rust for showed fair consistency for WP while FL and SL showed good consistency.



Box plot of crown and stem rust ratings across the 50 SL images showing differences in agreement.



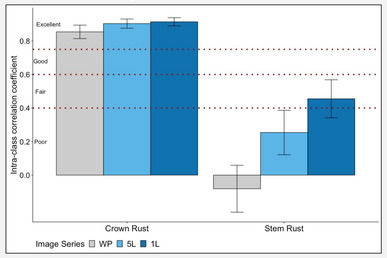
ICC values for each image series and rust showing relative rank of rater scores (1= perfect consistency).

Correlations between image series

	Rater Crown Rust (*r*)		
	WP	FL	SL
WP	-	***	***
FL	0.89	-	***
SL	0.86	0.93	-
	Rater Stem Rust (*r*)		
WP	-	*	***
FL	0.36	-	***
SL	0.51	0.70	-

	Computer Crown Rust (*r*)		
	WP	FL	SL
WP	-	***	***
FL	0.84	-	***
SL	0.80	0.94	-
	Computer Stem Rust (*r*)		
WP	-	NS	NS
FL	0.20	-	***
SL	0.09	0.85	-

Consistency and correlation between rater and computer scores



	Computer severity score	
Rater severity score	*Crown rust (r)*	*Stem rust (r)*
WP	0.94 ***	NS
5L	0.94 ***	0.65 ***
1L	0.93 ***	0.79 ***

*, **, *** significant at the 0.05, 0.01, and 0.001 probability levels

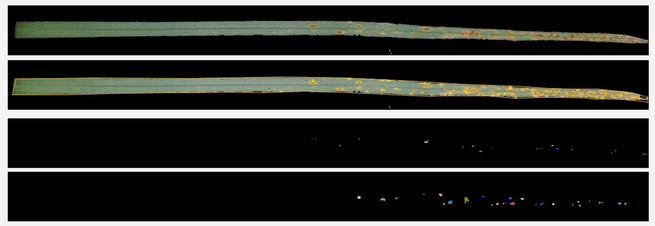
Computer and rater scores were very consistent for crown rust, but not stem rust. This is likely due to the inability to see stem rust on WP.

Correlations between raters and the computer were excellent for crown rust and SL stem rust images. Stem rust was not significantly correlated for WP.

Median rater severity
Total severity: *11.1%*

Hand calculated
Total severity: *3.5%*

Computer severity:
Stem rust (top): *0.4%*
Crown (bottom): *3.0%*
Total severity: *3.4%*



Raters overestimated actual percent infestation of disease. Stem rust was more challenging than crown rust in general.

Conclusions

- An open-source, automated image analysis system was developed and accurately predicted crown rust severity on WP, FL and SL images.
- This system does not accurately quantify stem rust on a whole plant scale, but will work for stem or leaf samples.
- Future research should focus on implementing multispectral imaging that would add additional explanatory power to the random forest model.

References:

Law, A., and M. Wiener. 2002. Classification and Regression by randomForest. R News 2(3).

Pau, G., F. Fuchs, O. Sklyar, M. Boutros, and W. Huber. 2010. EBImage—an R package for image processing with applications to cellular phenotypes. Bioinformatics 26(7): 979–981.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.