Machine Learning; The Future of High Throughput Plant Phenotyping.

- **Isha Joshi**

Machine learning uses established and annotated resources as an illustration to retrieve and accumulate information. It predicts an unseen data and classify it in a defined category. Furthermore, giving us a multidisciplinary statistical approach, which can be looked upon as a combination of human intelligence with computer accuracy and memory. Self-experience is utilised to learn and draw patterns in the submitted data. Additionally, with increase in experience; its scrutiny of decision making also increases. Recent developments in the field of plant phenotyping termed as high throughput phenotyping has led breeders and researcher to look at a specific trait more broadly at individual and combinatorial level. Plant architecture can vary in different conditions and can indicate the state of agronomical important traits and the physiological function associated with them.

Machine learning architectures in combination with public datasets enlisted in table 1 are used for Image based phenotyping. An important application of machine learning is high throughput phenotyping where automated platforms such as UAV, ground and aircraft robot equipped with multiple sensors inducing remote sensing, Spectro radiometry, Light Detection and Ranging (LIDAR) , visible to far-infrared hyperspectral , thermal , fluorescence , 3D laser scanning , trichromatic (RGB), UV illumination and halogen excitation are used to capture and store data in the form of images which are further processed as training or testing datasets by machine learning supervised or unsupervised algorithms. The valuation is normalised by an inclusive range of absorbance index and Leaf Reflectance parameters recorded by the sensors which can trace non-photosynthetically active constituents, water, nitrogen and chlorophyll content [4-5]. Various studies which have used image based high throughput phenotyping are enlisted in table 1.

The available plethora of insightful information through research projects is not easily accessible by the farming community. It is certainly difficult for them to collect the samples and take them to a specific laboratory to get them tested for any bizarre phenotype. Smartphone assisted diagnostic approaches can bridge this gap by providing an easily available, economical and approachable way especially for smallholder farmers and researchers with limited funds. The use of broad-spectrum pesticide and fertilizer can be ineffective to treat a specific condition and instead increase the problem in certain cases when overused.

Automated agricultural drones mounted with artificial intelligence technologies can be used for the selective spray of water, specific nutrients, insecticide and herbicide wherever needed. When the above is equipped with a high-resolution HD camera, It can capture images detailing about the plant phenotype, soil quality, potential pest invasions more quickly and precisely. This can lead to the development of less expensive and higher productivity farming approach ensuring food security and higher investment returns. Algorithms which can differentiate more than one parameter simultaneously and concurrently are the need of the hour.

The high throughput phenotyping which has been widely used in plant-disease profiling should be parallelly looked out for other physiological traits in varied and diverse experimental replicates. Exploration of various absorption spectra and secondary metabolite image-based mapping can lead to the use of identified and unidentified medicinal plants in low research cost. Artificial intelligence studies like disease forecasting in accordance with the weather conditions are highly translational and should be encouraged [32]. Machine learning is the future of the plant science research but only if the limitation in the form of an available narrow database range is well dealt with.

TABLE.1- Machine learning in image based high throughput phenotyping						
S. no	**USAGE**	**ALGORITHM/ CLASSIFIERS**	**SENSOR**	**ASSOCIATED TRAIT**	**ORGANISM OF STUDY**	**REFERENCE**
1.	Biotic Stress Identification And Classification	NN	Spectral/ Hyperspectral Images, Thermography	*Puccinia striiformis f.sp.Tritici infection*	Wheat	[6]
			Thermal Images	*Alternaria infection*		
		SVM	Multicolour Fluorescence, Thermography	*Oidium neolycopersici infection*	Tomato	[7]
		LR SVM NN	Portable Field Spectrometer, Spectroradiometer, Infrared Thermometer	Pathogenic bacteria	Melon	[8]
		RF	Fluorescence Imaging	*Septoria tritici blotch*	Wheat Lemon	[9]
				myrtle rust	Myrtle Tress	[10]
		SVM KNN QDA LDA	Spectroscopy, UAV- And Aircraft-Based Sensors, Spectroradiometer	*Huanglongbing (HLB) infection*	Citrus	[11-13]
		SAM	Remote sensing	*Heterodera schachtii, Rhizoctonia solani infection*	Sugar Beet	[14]
		Bayes factor DAR	Hyperspectral images	*Rust, net blotch, and powdery mildew disease*	Barley	[15]
		ANN variant	RGB images	*Bacterial soft rot, Phythopthora black rot, Bacterial brown spot*	Orchid	[16]
2.	Abiotic Stress (Drought)	SVM GPC	Visible /Thermal Images	Water stress	Spinach	[17]
		SVIM			Tomato	[18]
		DAR	Hyperspectral		Barley	[19]
3.	Presence of toxins	HBBE MLPNN LDA	CCD images	Aflatoxins	Chili Pepper	[20]
4.	Image-based plant phenotyping	CNN	Images taken from Canon 650D and	QTL analysis by root and shoot mapping	Wheat	[21]

			Nikon D5100 DSLR camera			
5.	3D phenotyping of entire plant	Segmentation Algorithm	RGB-depth camera	Shoot and segmented leaf architecture	Rosebush	[22]
6.	Pollution	LDA k means	RGB images Clover	Ozone	Trifolium Subterraneu m L.	[23]
7.	Herbicide tolerance	Obia	UAV-based RGB images and multispectral image	Weed identification	Sunflower	[24]
8.	Nutrient deficiency	SVM Variant	Scanned images	Nitrogen, phosphorus, and potassium (NPK) stress.	Rice	[25]
		DCNN	RGB images	Iron and potassium	Soyabean	[26]
9.	Canopies and soil measurement s	NN	Hyperspectral images	Phytophthora infestans	Tomato	[27]
10 .	Preplanning risk prediction	MR, NN, RF	Self-generated dataset	Stagonospora nodorum blotch	Wheat	[28]
11 .	Plant physiological stress	NDVI	Optical remote sensor	leaf chlorophyll content estimation Nitrogen sensing	Spring Maize	[29]
12 .	Digitized natural history collection analysis	CNN	RGB Images	Mercury-stained specimens	Herbarium	[30]
13 .	Plant identification	Pl@ntNet	Images via social network	Multiorgan Identification	2200 Plant Species	[31]
Public Image-based datasets/databas e		plant village, Cifar-10, Real field, Wheat Disease Database 2017, Flavia, Foliage, LeafSnap3, Real wheat field, Real environment, CASC-IFW, Bisque, Swedish leaf, Oxford Flower 17, Oxford Flower 1o2, TRY, Imagenet.				
Absorption Spectral		CAI—Cellulose Absorption Index, LCA—Lignin-Cellulose Absorption Index, NTDI—Normalized Difference Tillage Index, LWVI-1 – Normalized Difference, Leaf water VI 2, DLAI—Difference Leaf Area Index, PWI—Plant Water Index, NLI—Nonlinear Vegetation Index, DWSI—Disease Water Stress Index, NDVI—Normalized Difference, Vegetation Index, MCARI—Modified Chlorophyll Absorption Ratio Index, GI—Greenness Index, CAR—Chlorophyll Absorption Ratio, GNDVI—Green Normalized Difference Vegetation Index, OSAVI—Optimized Soil Adjusted Vegetation Index, CI r—Coloration Index red, CI g—Coloration Index green				

References-

1. Singh, A., Ganapathysubramanian, B., Singh, A.K. and Sarkar, S., 2016. Machine learning for high-throughput stress phenotyping in plants. *Trends in plant science*, *21*(2), pp.110-124.
2. Sun, S., Wang, C., Ding, H. and Zou, Q., 2020. Machine learning and its applications in plant molecular studies. *Briefings in Functional Genomics*, *19*(1), pp.40-48.
3. Ma, C., Zhang, H.H. and Wang, X., 2014. Machine learning for Big Data analytics in plants. *Trends in plant science*, *19*(12), pp.798-808.
4. Randive, P.U., Deshmukh, R.R., Janse, P.V. and Gupta, R.S., 2018, December. Discrimination Between Healthy and Diseased Cotton Plant by Using Hyperspectral Reflectance Data. In *International Conference on Recent Trends in Image Processing and Pattern Recognition* (pp. 342-351). Springer, Singapore.
5. Perez-Sanz, F., Navarro, P.J. and Egea-Cortines, M., 2017. Plant phenomics: an overview of image acquisition technologies and image data analysis algorithms. *GigaScience*, *6*(11), p.gix092.
6. Moshou, D., Bravo, C., West, J., Wahlen, S., McCartney, A. and Ramon, H., 2004. Automatic detection of 'yellow rust'in wheat using reflectance measurements and neural networks. *Computers and electronics in agriculture*, *44*(3), pp.173-188.
7. Prince, G., Clarkson, J.P. and Rajpoot, N.M., 2015. Automatic detection of diseased tomato plants using thermal and stereo visible light images. *PloS one*, *10*(4), p.e0123262.
8. Pineda, M., Pérez-Bueno, M.L. and Barón, M., 2018. Detection of bacterial infection in melon plants by classification methods based on imaging data. *Frontiers in plant science*, *9*, p.164.
9. Odilbekov, F., Armonienė, R., Henriksson, T. and Chawade, A., 2018. Proximal phenotyping and machine learning methods to identify Septoria tritici blotch disease symptoms in wheat. *Frontiers in plant science*, *9*, p.685.
10. Heim, R.H.J., Wright, I.J., Chang, H.C., Carnegie, A.J., Pegg, G.S., Lancaster, E.K., Falster, D.S. and Oldeland, J., 2018. Detecting myrtle rust (Austropuccinia psidii) on lemon myrtle trees using spectral signatures and machine learning. *Plant Pathology*, *67*(5), pp.1114-1121.
11. Wetterich, C.B., Kumar, R., Sankaran, S., Junior, J.B., Ehsani, R. and Marcassa, L.G., 2013, October. A comparative study on application of computer vision and fluorescence imaging spectroscopy for detection of citrus huanglongbing disease in USA and Brazil. In *Laser Science* (pp. JW3A-26). Optical Society of America.
12. Garcia-Ruiz, F., Sankaran, S., Maja, J.M., Lee, W.S., Rasmussen, J. and Ehsani, R., 2013. Comparison of two aerial imaging platforms for identification of Huanglongbing-infected citrus trees. *Computers and Electronics in Agriculture*, *91*, pp.106-115.
13. Sankaran, S., Mishra, A., Maja, J.M. and Ehsani, R., 2011. Visible-near infrared spectroscopy for detection of Huanglongbing in citrus orchards. *Computers and electronics in agriculture*, *77*(2), pp.127-134.

14. Hillnhütter, C., Mahlein, A.K., Sikora, R.A. and Oerke, E.C., 2011. Remote sensing to detect plant stress induced by Heterodera schachtii and Rhizoctonia solani in sugar beet fields. *Field Crops Research*, *122*(1), pp.70-77.
15. Wahabzada, M., Mahlein, A.K., Bauckhage, C., Steiner, U., Oerke, E.C. and Kersting, K., 2015. Metro maps of plant disease dynamics—automated mining of differences using hyperspectral images. *Plos one*, *10*(1), p.e0116902.
16. Huang, K.Y., 2007. Application of artificial neural network for detecting Phalaenopsis seedling diseases using color and texture features. *Computers and Electronics in agriculture*, *57*(1), pp.3-11.
17. Smith, H.K., Clarkson, G.J., Taylor, G., Thompson, A.J., Clarkson, J. and Rajpoot, N.M., 2014. Automatic detection of regions in spinach canopies responding to soil moisture deficit using combined visible and thermal imagery. *PloS one*, *9*(6), p.e97612.
18. Römer, C., Wahabzada, M., Ballvora, A., Pinto, F., Rossini, M., Panigada, C., Behmann, J., Léon, J., Thurau, C., Bauckhage, C. and Kersting, K., 2012. Early drought stress detection in cereals: simplex volume maximisation for hyperspectral image analysis. *Functional Plant Biology*, *39*(11), pp.878-890.
19. Kersting, K., Xu, Z., Wahabzada, M., Bauckhage, C., Thurau, C., Roemer, C., Ballvora, A., Rascher, U., Leon, J. and Pluemer, L., 2012, July. Pre-Symptomatic Prediction of Plant Drought Stress Using Dirichlet-Aggregation Regression on Hyperspectral Images. In *AAAI*.
20. Ataş, M., Yardimci, Y. and Temizel, A., 2012. A new approach to aflatoxin detection in chili pepper by machine vision. *Computers and electronics in agriculture*, *87*, pp.129-141.
21. Pound, M.P., Atkinson, J.A., Townsend, A.J., Wilson, M.H., Griffiths, M., Jackson, A.S., Bulat, A., Tzimiropoulos, G., Wells, D.M., Murchie, E.H. and Pridmore, T.P., 2017. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. *Gigascience*, *6*(10), p.gix083.
22. Chéné, Y., Rousseau, D., Lucidarme, P., Bertheloot, J., Caffier, V., Morel, P., Belin, É. and Chapeau-Blondeau, F., 2012. On the use of depth camera for 3D phenotyping of entire plants. *Computers and Electronics in Agriculture*, *82*, pp.122-127.
23. Kruse, O.M.O., Prats-Montalbán, J.M., Indahl, U.G., Kvaal, K., Ferrer, A. and Futsaether, C.M., 2014. Pixel classification methods for identifying and quantifying leaf surface injury from digital images. *Computers and electronics in Agriculture*, *108*, pp.155-165.
24. Peña, J.M., Torres-Sánchez, J., Serrano-Pérez, A., De Castro, A.I. and López-Granados, F., 2015. Quantifying efficacy and limits of unmanned aerial vehicle (UAV) technology for weed seedling detection as affected by sensor resolution. *Sensors*, *15*(3), pp.5609-5626.
25. Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T. and Klukas, C., 2014. Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *The Plant Cell*, *26*(12), pp.4636-4655.
26. Ghosal, S., Blystone, D., Singh, A.K., Ganapathysubramanian, B., Singh, A. and Sarkar, S., 2018. An explainable deep machine vision framework for plant stress phenotyping. *Proceedings of the National Academy of Sciences*, *115*(18), pp.4613-4618.

27. Wang, X., Zhang, M., Zhu, J. and Geng, S., 2008. Spectral prediction of Phytophthora infestans infection on tomatoes using artificial neural network (ANN). *International Journal of Remote Sensing*, *29*(6), pp.1693-1706.
28. Mehra, L.K., Cowger, C., Gross, K. and Ojiambo, P.S., 2016. Predicting pre-planting risk of Stagonospora nodorum blotch in winter wheat using machine learning models. *Frontiers in plant science*, *7*, p.390.
29. Yanli, L., Youlu, B. and Liping, Y., 2008. Diagnosis on nitrogen status using GreenSeeker in spring maize. *Journal of Maize Sciences*.
30. Schuettpelz, E., Frandsen, P.B., Dikow, R.B., Brown, A., Orli, S., Peters, M., Metallo, A., Funk, V.A. and Dorr, L.J., 2017. Applications of deep convolutional neural networks to digitized natural history collections. *Biodiversity Data Journal*, (5).
31. Joly, A., Goëau, H., Bonnet, P., Bakić, V., Barbe, J., Selmi, S., Yahiaoui, I., Carré, J., Mouysset, E., Molino, J.F. and Boujemaa, N., 2014. Interactive plant identification based on social image data. *Ecological Informatics*, *23*, pp.22-34.
32. Kaundal, R., Kapoor, A.S. and Raghava, G.P., 2006. Machine learning techniques in disease forecasting: a case study on rice blast prediction. *BMC bioinformatics*, *7*(1), p.485.