Extraction of Phenotypic Traits for Drought Stress Study Using Hyperspectral Images

Swati Bhugra$^{(\boxtimes)}$ ⓘ, Nitish Agarwal ⓘ, Shubham Yadav ⓘ, Soham Banerjee ⓘ, Santanu Chaudhury ⓘ, and Brejesh Lall ⓘ

Department of Electrical Engineering, Indian Institute of Technology, Delhi, India
{eez138301,ee1130477,ee1130499,ee1130502,santanuc,brejesh}@ee.iitd.ac.in

Abstract. High-throughput identification of digital traits encapsulating the changes in plant's internal structure under drought stress, based on hyperspectral imaging (HSI) is a challenging task. This is due to the high spectral and spatial resolution of HSI data and lack of labelled data. Therefore, this work proposes a novel framework for phenotypic discovery based on autoencoders, which is trained using Simple Linear Iterative Clustering (SLIC) superpixels. The distinctive archetypes from the learnt digital traits are selected using simplex volume maximisation (SiVM). Their accumulation maps are employed to reveal differential drought responses of wheat cultivars based on t-distributed stochastic neighbour embedding (t-SNE) and the separability is quantified using cluster silhouette index. Unlike prior methods using raw pixels or feature vectors computed by fusing predefined indices as phenotypic traits, our proposed framework shows potential by separating the plant responses into three classes with a finer granularity. This capability shows the potential of our framework for the discovery of data-driven phenotypes to quantify drought stress responses.

Keywords: Autoencoders · SLIC superpixels · Hyperspectral · Drought stress · t-SNE · Deep learning · SiVM

1 Introduction

Drought stress is a major limiting factor for crop productivity. To select high yielding cultivars under drought conditions, current strategies depend on both (a) genotypic data of the cultivar and (b) quantification of its physiological and structural characteristics (phenotypic data) [1]. In this context, digital traits based on hyperspectral imaging (HSI) data have the potential to reveal changes in the plant's internal structure non-destructively [1]. Previous studies are based on extraction of single vegetation indices computed using two or three spectral bands. However, using single vegetation indices only quantifies specific changes to detect drought stress [2]. Behmann et al. [3] formulated a feature based on fusion of all the vegetation indices to characterise different stages of leaf senescence, whereas Römer et al. [4] used the entire spectra of the pixels in the HSI data. Since the stress labels for the pixels are not available, extracting digital traits to

study drought is an unsupervised task. Behmann et al. employed a framework combining k-means to extract drought clusters and based on these cluster labels, used Support Vector Machine (SVM) for drought stress classification. Here, the number of cluster centres and annotation of each cluster centre to its corresponding drought level was performed manually by an expert. Similarly, Römer et al. employed simplex volume maximisation (SiVM) to extract archetypes from the raw spectra of the pixels. To avoid the selection of noisy spectra as an archetype, an expert manually extracted only the spectrum belonging to plant pixels based on the domain knowledge.

Recent establishment of phenotyping platforms provides automated imaging of large number of experimental plants for drought stress study [1]. Since, the aforementioned approaches rely on experts, these will scale badly to this growing amount of data and are also prone to human bias. Thus, RGB imaging modules available in these platforms have been frequently exploited in contrast to HSI for drought analysis [1]. But, HSI data can capture intricate phenotypic information as compared to the corresponding RGB representations of the plant canopy [1]. Therefore, we propose a framework based on single-layer and multi-layer stacked autoencoders (AE) [5] to learn shallow and deep features respectively from HSI data in an unsupervised manner. The compact representations obtained from these networks were utilized to select distinctive archetypes using SiVM [6]. To identify different clusters that represent different degrees of drought stress level using the agglomerative representation computed from these archetypes for each HSI data, t-SNE [7] was employed. The separability of these clusters was quantified using Silhouette coefficient [8]. Although many methods have been proposed for the purpose of drought stress identification [9,10], but to the best of our knowledge, this is the first work that utilizes deep networks on HSI data to learn an implicit representation of features for drought stress characterization. To show the eligibility of our proposed approach based on the separation of different responses effectively, we empirically compared the silhouette coefficient with the (a) classical approach of using raw pixel spectra [4] and (b) feature comprising of different indices [3].

The rest of the paper is organised as follows: In Sect. 2 the dataset is described, Sect. 3 explains the methodology and in Sect. 4 the results are discussed.

2   Dataset

The drought experiment was conducted on wheat pots at the Plant Phenomics Facility, Indian Agricultural Research Institute (IARI), Pusa, New-Delhi during Rabi season of 2016. To examine the differential responses of drought, two genotypes of wheat crop: C-306 (drought tolerant) and HD-2967 (drought sensitive) were investigated for a period of 5 continuous days. Both genotypes were divided into three groups (six replicates each) in terms of water intensity i.e. well-watered, reduced watered and unwatered. HSI data was captured in the spectral range of 400 nm to 1000 nm at equal wavelength intervals resulting in
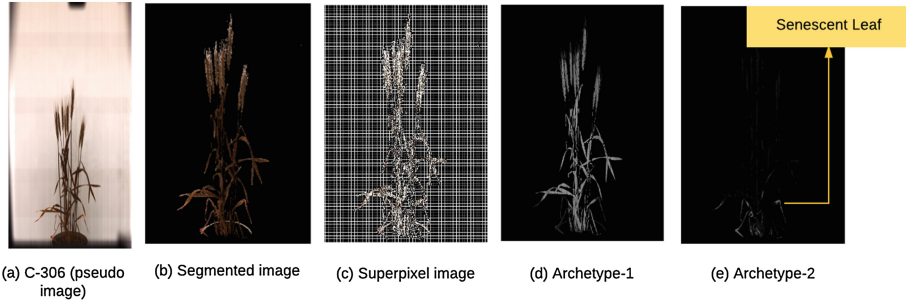
(a) C-306 (pseudo image)   (b) Segmented image   (c) Superpixel image   (d) Archetype-1   (e) Archetype-2

Fig. 1. Selected archetypes representation on C-306 (control)

108 bands and a corresponding pseudo colour image was also collected along the side view.

3   Methodology

The steps of the proposed framework to study the temporal dynamics of drought stress are explained in the following subsections.

3.1   Pre-processing Step

The HSI data contains wheat canopy and non-canopy elements such as soil, water and background. The segmentation of pseudo image is obtained graph-cut algorithm [11]. In addition to using the color features for each pixel as an input to the graph-cut, texture features [12] are also computed. The texture response is given by: $f(I; \beta, r, \sigma_H, \sigma_L) = \exp(\beta/(H_r * I_i + (G_{\sigma_H} * I_j - G_{\sigma_L} * I_j)))$ where, I is the pseudo image, G_{σ_H} and G_{σ_L} are Gaussian filters with σ_H and σ_L respectively, difference of the Gaussian kernels highlights high texture regions, H_r is a uniform circular filter with radius r that highlights the smooth regions in the image and β is the fall-off rate. The segmented pseudo image is used as a mask to extract the plant pixels from all the hyperspectral bands. Due to the highly correlated spectra of the neighbouring pixels in the segmented HSI data, superpixels based on Simple Linear Iterative Clustering (SLIC) [13] are extracted. The homogeneous regions obtained using SLIC computes a better representative spectra with less noise than the use of raw pixel spectra for subsequent analysis.

3.2   Feature Learning

The superpixels extracted from the HSI data were used to train the autoencoder and stacked autoencoder to learn shallow and deep features respectively in an unsupervised manner. The autoencoder (AE) [5] comprises of an encoder and a decoder. The encoder obtains the latent representation of dimension $H < M$ from

the input features $x \in \mathbb{R}^M$ given by: $h = \phi(\mathbf{P}x+b)$, where $\mathbf{P} \in \mathbb{R}^{H \times M}$ is a matrix of learned weights, $b \in \mathbb{R}^H$ is bias vector and ϕ is an activation function. We used logistic sigmoid function as the activation function. A decoder reconstructs the input features $\hat{x}$ using the latent representation given by: $\hat{x} = \epsilon(\mathbf{P}'h + b')$, where ϵ is a logistic sigmoid activation function, $\mathbf{P}'$ is the weight matrix and b' the bias vector. Beginning with random initialisation of $\{\mathbf{P}, \mathbf{P}', b, b'\}$, the training process is formulated as an unsupervised optimisation of a cost function which measures the error between the input and its reconstruction given by: $\theta = argmin_\theta \mathbf{L}(x, \hat{x})$ with respect to the parameters $\theta = \{\mathbf{P}, \mathbf{P}', b, b'\}$, where $\mathbf{L}$ is defined as the squared difference between x and $\hat{x}$. The weights are updated with stochastic gradient descent which can be efficiently implemented using the Back-Propagation algorithm. In order to avoid over-fitting, a standard $L2$ norm weight regularisation [14] is employed for the elements of $\mathbf{P}$. For H hidden nodes, N training examples and M features (spectral bands) in the training data, this is given by:

$$\eta_w = \sum_h^H \sum_j^N \sum_i^M (w_{ji}{}^h)^2 \qquad (1)$$

Using the Kullback-Leibler (KL) divergence, sparsity regularisation [15] is included with sparsity ($\rho = .5$) and is computed as shown below:

$$\eta_s = \sum_i^H \rho \, log\Big(\frac{\rho}{\hat{\rho}_i}\Big) + (1 - \rho) \, log\Big(\frac{1 - \rho}{1 - \hat{\rho}_i}\Big) \qquad (2)$$

The cost function for the unsupervised optimisation problem is given by:

$$\mathbb{E} = \frac{1}{N} \sum_j^N \eta_j + \lambda\eta_w + \beta\eta_s \qquad (3)$$

where, η_j is the squared error for the jth training data, λ and β is set to .01 and .5 respectively.

Stacked autoencoders (SAE) [5] are constructed using greedy layer wise strategy on the learnt AE. The learnt representation h of the input feature x is used as an input to another autoencoder which learns a latent representation v and so on. A deep network is obtained by stacking the layer wise trained autoencoders. The latent representation obtained from both the aforementioned architectures is then employed to quantify drought stress response, as explained in the next subsection.

3.3   Drought Stress Quantification

At the pixel level, well defined labels to denote different drought stress responses (DSR) are not available. Thus, to obtain distinctive archetypes characterising drought stages from the learnt features, SiVM is employed. SiVM extracts the archetypes by fitting a simplex with the maximum volume to the data. It selects

the archetypes from the data matrix $\mathbf{V}$. Thus, the matrix $\mathbf{W}$ is defined as $\mathbf{W} = \mathbf{VG}$ with n data points and k selected archetypes where, $\mathbf{G} \in \mathbb{R}^{n \times k}$ and is restricted to unary column vectors and $\mathbf{H}$ is restricted to convexity. The number of extreme archetypes is chosen based on the increase in the accumulated maps of the selected archetypes [4]. If $\mathbf{W}_m$ is the matrix of selected m archetypes, h_{dm} is the normalised aggregation (belongingness histogram) of $\mathbf{H}_{dm}$ of dth HSI data where $d = 1, 2, \cdots, N$ images, then $(m+1)^{th}$ archetype is added to $\mathbf{W}_m$, if $\mathbb{E}[h_{d(m+1)} - h_{dm}] \geq 0.1$. Thus, only those archetypes are selected which results in significant accumulation of the plant pixels in all HSI data. Belongingness histogram are represented in a 2D space using t-SNE [7] to study the clusters or patterns in terms of drought responses. t-SNE maps a high dimensional data into a lower dimension space by computing pair-wise similarity matrix and simultaneously preserving local and global structure of the data, whereas other classical approaches such as principal component analysis (PCA) [7], multidimensional scaling (MDS) [7] may not capture the non-linear relationship in high dimensional data. The stress responses obtained from t-SNE are evaluated using silhouette coefficient [8]. It is used to quantify the separability of different responses.
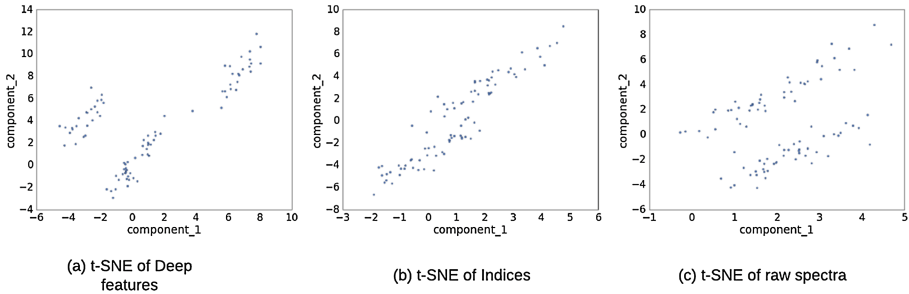

(a) t-SNE of Deep features

(b) t-SNE of Indices

(c) t-SNE of raw spectra

Fig. 2. Drought stress characterisation

4   Results

The segmented pseudo image, with the corresponding SLIC superpixels is shown in Fig. 1(b) and (c) respectively. $160,000$ superpixels obtained from six images belonging to different irrigation treatments were used to train the autoencoder. The autoencoder was trained in an unsupervised manner and no supervised fine tuning was applied. To improve the convergence of the training algorithms, data was normalised, i.e. the spectral band with zero mean and the spectral values normalised to unit variance. The learning rate was kept low i.e. 10^{-4}. In the first AE, the 108 spectral bands were mapped to a latent representation in a subspace with a hidden layer of 80 nodes. For a deep representation of the input spectra, the latent representation obtained was used to train the second AE. The hidden layer of this second AE was estimated to be 30, based on the small reconstruction

error. The trained model was validated based on mean square error of the test data, which was of the order of 10^{-2} and lower as compared to the first AE. Thus for the subsequent drought stress analysis, stacked AE model was employed. The representation obtained for the SLIC pixels, was used as an input to SiVM and four archetypes were selected automatically based on the aforementioned criteria in Sect. 3.3. The visualisation of the accumulation map of C-306 cultivar (control day 1) corresponding to the 1st and 2nd archetype is shown in Fig. 1(d) and (e) respectively. The pixels belonging to high intensity shows high similarity to the archetype. The 1st latent representation selected encapsulates the spectra of the healthy leaves and the 2nd representation captures the leaves with senescence. This illustrates the effectiveness of the autoencoder to separate the different drought responses within the plant canopy based on unsupervised dimensionality reduction. Although training the weight matrix requires a large amount of computation time, but once the trained model is obtained the encoding of the data to extract the compact features is fast.

Since drought symptoms do not manifest over local regions but the entire plant (as shown in Fig. 1), aggregated belongingness for all the pixels in each HSI data represents the overall stress response of the plant. To uncover the different set of drought responses captured using these learnt features, t-SNE (with perplexity $= 5$) was employed and silhouette coefficient was used to validate the separability. The t-SNE computed for the learnt feature is shown in Fig. 2(a). The significance of the groups was obtained by linking image to the genotype and drought/control day. It was discovered that in Fig. 2(a), one cluster consisted of the plants belonging to the control group, the second cluster belonged to the C-306 replicates at drought days $3, 4, 5$ and the third clusters comprised of the HD-2967 replicates at the same drought days (silhouette coefficient .811). This shows a statistical difference between the drought response of both the varieties on the same day of drought stress, which was successfully captured using the learnt features. On the other hand, based on the archetypes selected from (a) raw spectra [4] and (b) the feature vector comprising of indices [3], only two distinct clusters: control and drought (with silhouette coefficient 0.711 and 0.69 respectively) were identified. This finer granularity in classification shows the potential of our framework to reveal data-driven phenotypic patterns.

5   Conclusion

In this article, we presented a novel framework that provides phenotypic expression of drought stress based on deep learning and based on these expressions, captured the phenotypic difference between a drought susceptible cultivar (HD-2967) and a drought tolerant cultivar (C-306). The findings contribute to the ongoing studies to predict drought stress based on phenotypic traits. The application of deep networks for drought study based on HSI data is largely unexplored, thus more investigations of the learnt features and its relation to different physiological responses in plants is essential for an accurate and high throughput drought characterization.

References

1. Humplík, J.F., Lazár, D., Husiĉková, A., Spíchal, L.: Automated phenotyping of plant shoots using imaging methods for analysis of plant stress responses–a review. Plant Meth. **11**(1), 29 (2015)
2. Thenkabail, P.S., Smith, R.B., De Pauw, E.: Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. Remote Sens. Environ. **71**(2), 158–182 (2000)
3. Behmann, J., Steinrücken, J., Plümer, L.: Detection of early plant stress responses in hyperspectral images. ISPRS J. Photogrammetry Remote Sens. **93**, 98–111 (2014)
4. Römer, C., Wahabzada, M., Ballvora, A., Pinto, F., Rossini, M., Panigada, C., Behmann, J., Léon, J., Thurau, C., Bauckhage, C., Kersting, K.: Early drought stress detection in cereals: simplex volume maximisation for hyperspectral image analysis. Funct. Plant Biol. **39**(11), 878–890 (2012)
5. Palm, R.B.: Prediction as a candidate for learning deep hierarchical models of data, vol. 5. Technical University of Denmark (2012)
6. Thurau, C., et al.: Yes we can: simplex volume maximization for descriptive web-scale matrix factorization, pp. 1785–1788. ACM (2010)
7. van der Maaten, L., Hinton, G.: Visualizing data using t-SNE. J. Mach. Learn. Res. **9**, 2579–2605 (2008)
8. Bolshakova, N., Azuaje, F.: Cluster validation techniques for genome expression data. Signal Process. **83**(4), 825–833 (2003)
9. Singh, A., Ganapathysubramanian, B., Singh, A.K., Sarkar, S.: Machine learning for high-throughput stress phenotyping in plants. Trends Plant Sci. **21**(2), 110–124 (2016)
10. Bhugra, S., Anupama, A., Chaudhury, S., Lall, B., Chugh, A.: Phenotyping of xylem vessels for drought stress analysis in rice. In: Fifteenth IAPR International Conference on Machine Vision Applications (MVA), pp. 428–431 (2017)
11. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Mach. Intell. **23**(11), 1222–1239 (2001)
12. Minervini, M., Abdelsamea, M.M., Tsaftaris, S.A.: Image-based plant phenotyping with incremental learning and active contours. Ecol. Inf. **23**, 35–48 (2014)
13. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, P.: SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. **34**(11), 2274–2282 (2012)
14. Møller, M.F.: A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw. **6**(4), 525–533 (1993)
15. Olshausen, B.A., Field, D.J.: Sparse coding with an overcomplete basis set: a strategy employed by V1. Vis. Res. **37**(23), 3311–3325 (1997)