UNSUPERVISED DETECTION OF ASH DIEBACK DISEASE (*HYMENOSCYPHUS FRAXINEUS*) USING DIFFUSION-BASED HYPERSPECTRAL IMAGE CLUSTERING

Sam L. Polk [1], *Aland H. Y. Chan* [2], *Kangning Cui* [3],
Robert J. Plemmons [4], *David A. Coomes* [2], *and James M. Murphy* [1*]

[1] Department of Mathematics, Tufts University
[2] Department of Plant Sciences, University of Cambridge
[3] Department of Mathematics, City University of Hong Kong
[4] Departments of Mathematics and Computer Science, Wake Forest University

ABSTRACT

Ash dieback (*Hymenoscyphus fraxineus*) is an introduced fungal disease that is causing the widespread death of ash trees across Europe. Remote sensing hyperspectral images encode rich structure that has been exploited for the detection of dieback disease in ash trees using supervised machine learning techniques. However, to understand the state of forest health at landscape-scale, accurate unsupervised approaches are needed. This article investigates the use of the unsupervised Diffusion and VCA-Assisted Image Segmentation (D-VIS) clustering algorithm for the detection of ash dieback disease in a forest site near Cambridge, United Kingdom. The unsupervised clustering presented in this work has high overlap with the supervised classification of previous work on this scene (overall accuracy = 71%). Thus, unsupervised learning may be used for the remote detection of ash dieback disease without the need for expert labeling.

Index Terms— Clustering, Diffusion, Forestry, Graphs, Hyperspectral Images, Unsupervised Machine Learning.

1. INTRODUCTION

Ash dieback disease, caused by the *Hymenoscyphus fraxineus* ascomycete, poses a major threat to the health of European forests and ash-dependent biota [1]. To aid epidemiologists and forest managers in the modeling and mitigation of ash dieback disease, healthy and infected trees must be identified at landscape-scale [2]. Recent years have brought significant advances in methods for passive remote sensing of forest pathogens [3, 4]. For example, remote sensing hyperspectral images (HSIs) are high-dimensional images that encode rich structure that can be exploited by machine learning algorithms for the detection of diseases in forests [2, 3, 4]. While HSIs may be used for disease detection, the large volume of

HSI data generated over forests makes manual labeling (generally required for supervised learning) infeasible. Thus, unsupervised machine learning algorithms are needed to produce accurate disease mappings of forests using HSIs.

This article implements the unsupervised Diffusion and VCA-Assisted Image Segmentation (D-VIS) on an HSI generated over a forest in Madingley Village, near Cambridge, United Kingdom [2]. D-VIS is closely related to Diffusion and Volume maximization-based Image Clustering (D-VIC), which has been shown to successfully recover ground truth labels from benchmark HSIs [5]. We compare an unsupervised D-VIS clustering to a disease mapping generated by a supervised Random Forest (RF) [2]. High levels of overlap were observed between unsupervised and supervised partitions, indicating that unsupervised learning—and D-VIS, in particular—may be used to detect ash dieback disease in European forests even when no ground truth labels are available.

The rest of this article is structured as follows. In Section 2, background material is provided on HSI segmentation, diffusion geometry, spectral unmixing, and the D-VIS clustering algorithm. In Section 3, the Madingley HSI is described, and disease mappings obtained by D-VIS are presented. We conclude and suggest directions for future work in Section 4.

2. BACKGROUND

2.1. Hyperspectral Image Segmentation

Let $X = \{x_i\}_{i=1}^n \subset \mathbb{R}^D$ be the set of HSI pixel spectra, where n and D indicate the number of pixels and bands, respectively. HSI segmentation algorithms partition an HSI into *clusters* of pixels $\{X_k\}_{k=1}^K$ [6, 7]. Ideally, any two pixels in the same cluster will be similar, but two pixels from different clusters will be dissimilar [6]. *Supervised* HSI segmentation algorithms rely on some fraction of the image's ground truth labels to obtain a partition of X. Conversely, *unsupervised* HSI segmentation algorithms (also called HSI clustering algorithms) require no ground truth labels to partition X [6].

2287

2.2. Diffusion Geometry

Graph-based HSI clustering algorithms treat each pixel as a node in an undirected graph [8]. The edges between pixels can be encoded in an adjacency matrix $\mathbf{W} \in \mathbb{R}^{n \times n}$, where $\mathbf{W}_{ij} = 1$ if x_j is one of the N ℓ^2-nearest neighbors of x_i and $\mathbf{W}_{ij} = 0$ otherwise. Define $\mathbf{P} = \mathbf{D}^{-1}\mathbf{W}$, where $\mathbf{D} \in \mathbb{R}^{n \times n}$ is the diagonal *degree matrix* defined by $\mathbf{D}_{ii} = \sum_{j=1}^{n} \mathbf{W}_{ij}$. We identify $\mathbf{P}$ as the transition matrix for a Markov diffusion process on HSI pixels. Assuming $\mathbf{P}$ is irreducible and aperiodic, there is a unique $\pi \in \mathbb{R}^{1 \times n}$ satisfying $\pi \mathbf{P} = \pi$ [8].

Diffusion distances enable the comparison of HSI pixels in the context of the diffusion process encoded in $\mathbf{P}$ [8, 9]. Define $D_t(x_i, x_j) = \sqrt{\sum_{k=1}^{n}((\mathbf{P}^t)_{ik} - (\mathbf{P}^t)_{jk})^2 / \pi_k}$ to be the diffusion distance between x_i and x_j at time $t \geq 0$ [8]. For datasets with highly coherent and well-separated latent cluster structure, the maximum within-cluster diffusion distance is bounded away from the minimum between-cluster diffusion distance across a large range of t [9, 10]. As such, diffusion distances can be used to learn latent structure in HSIs.

2.3. Spectral Unmixing

HSIs are usually generated at a relatively coarse spatial resolution (often with a spatial resolution as high as 10 m) [7]. As such, in many scenarios (e.g., disease mapping in forests), a single pixel may correspond to a spatial region containing multiple materials. If m is the number of materials present in the scene, the goal of a linear *spectral unmixing* algorithm is to find two nonnegative matrices, $\mathbf{A} \in \mathbb{R}^{n \times m}$ and $\mathbf{U} = (u_1, u_2, \ldots, u_m)^\top \in \mathbb{R}^{m \times D}$ such that $x_i \approx \sum_{j=1}^{m} \mathbf{A}_{ij} u_j$ for each $x_i \in X$ [11, 12]. Ideally, the rows of $\mathbf{U}$ will encode the intrinsic spectral signatures associated with materials in the scene, while the rows of $\mathbf{A}$ encode the relative frequency that those materials appear in a given pixel. Information about material abundance can be summarized using *pixel purity*, defined for each $x_i \in X$ by $\eta(x_i) = \frac{\max_{1 \leq j \leq m} \mathbf{A}_{ij}}{\sum_{j=1}^{m} \mathbf{A}_{ij}}$. The function $\eta(x)$ will be nearly 1 if the spatial region corresponding to the pixel x contains predominantly one material, and $\eta(x)$ will be small for mixed pixels containing many materials [5, 13].

2.4. Diffusion and VCA-Assisted Image Segmentation

In its first stage, D-VIS (Algorithm 1) locates K pixels that exemplify underlying structure in the HSI. To locate these *cluster modes*, D-VIS first calculates $\eta(x)$ using HySime [14] to learn m and VCA [12, 15] to learn $\mathbf{A}$ and $\mathbf{U}$. This differs slightly from D-VIC, which relies on Alternating Volume Maximization to learn $\mathbf{U}$ [5, 11]. Next, D-VIS calculates density: $p(x) = \sum_{y \in NN(x,N)} \exp(-\|x - y\|_2^2 / \sigma_0^2)$, where $NN(x, N)$ is the set of N ℓ^2-nearest neighbors of x in X and $\sigma_0 > 0$ is a *density scale* that controls interactions between

---

Algorithm 1: Diffusion and VCA-Assisted Image Segmentation (D-VIS)

Input: X (HSI), K (# clusters), N (# neighbors), σ_0 (density scale), t (diffusion time)
Output: $\hat{\mathcal{C}} = \{X_k\}_{k=1}^{K}$ (clustering)

1 Calculate $\eta(x)$ by implementing HySime [14] to estimate m and VCA [12, 15] to estimate $\mathbf{U}$ and $\mathbf{A}$;
2 Calculate $\zeta(x)$ using density $p(x)$ (see Section 2.4);
3 Label cluster modes $\hat{\mathcal{C}}(x_{m_k}) = k$ for $1 \leq k \leq K$, where $\{x_{m_k}\}_{k=1}^{K}$ are the K maximizers of $\mathcal{D}_t(x) = \zeta(x)d_t(x)$; $d_t(x)$ is as in Section 2.4;
4 In order of non-increasing $\zeta(x)$, assign unlabeled points $x \in X$ the label $\hat{\mathcal{C}}(x) = \hat{\mathcal{C}}(x^*)$, where $x^* = \underset{y \in X}{\operatorname{argmin}} \{D_t(x, y) | \zeta(y) \geq \zeta(x) \wedge \hat{\mathcal{C}}(y) > 0\}$;

---

points. The functions $p(x)$ and $\eta(x)$ are combined into a single measure $\zeta(x) = \frac{2\bar{p}(x)\bar{\eta}(x)}{\bar{p}(x) + \bar{\eta}(x)}$, where $\bar{p}(x) = \frac{p(x)}{\max_{y \in X} p(y)}$ and $\bar{\eta}(x) = \frac{\eta(x)}{\max_{y \in X} \eta(y)}$. As the harmonic mean of pixel purity and density, $\zeta(x)$ will be large if and only if x is both modal (i.e., close to its N ℓ^2-nearest neighbors) and representative of a single material in the scene (i.e., high-purity).

Cluster modes are identified in D-VIS as the maximizers of $\mathcal{D}_t(x) = \zeta(x)d_t(x)$, where

$$d_t(x) = \begin{cases} \underset{y \in X}{\min}\{D_t(x, y) | \zeta(y) \geq \zeta(x)\} & x = \underset{y \in X}{\operatorname{argmin}} \zeta(y), \\ \underset{y \in X}{\max} D_t(x, y) & \text{otherwise.} \end{cases}$$

The function $d_t(x)$ outputs the diffusion distance at time t between x and its diffusion distance-nearest neighbor of higher ζ-value. Thus, $\mathcal{D}_t$-maximizers are high-density, high-purity pixels that are far in diffusion distance at time t from other high-density, high-purity pixels. As such, these K points are suitable exemplars for the materials in the scene. Cluster modes are given unique labels, and each non-modal $x \in X$ is, in order of non-increasing $\zeta(x)$, given the label of its $\mathcal{D}_t$-nearest neighbor of higher ζ-value that is already labeled.

3. UNSUPERVISED ASH DIEBACK DETECTION

3.1. Madingley Hyperspectral Image

This article presents the implementation of D-VIS on hyperspectral data collected by a human-crewed aircraft over a 512m × 356m region of temperate deciduous forest near Madingley, on the outskirts of Cambridge, United Kingdom, in August 2018 [2]. Spectral reflectance was recorded using a Norsk Elektro Optikk hyperspectral camera (Hyspex VNIR 1800) at a spectral resolution of 3.26 nm across wavelengths 410-1001 nm and at a high spatial resolution of 0.32 m. Thus, reflectance at a total of $D = 186$ spectral bands was recorded

for 1816835 pixels in a scene with spatial dimensions of 1601×1113. QUick Atmospheric Correction [7] was implemented to remove atmospheric effects on pixel spectra, and spectral signatures were normalized to prevent differences in illumination from corrupting classification results [2].

3.2. Supervised Classification of Ash Dieback Disease

To enable the detection of ash dieback disease in the Madingley scene, a Partial Least Squares Discriminant Analysis (PLSDA) classifier was trained to locate pixels corresponding to ash trees in the Madingley scene [2]. Ground truth labels were collected for 166 tree crowns in the Madingley region as well as 256 tree crowns from three other forest sites near Cambridge [2]. Manually delineated and labeled tree crowns from these four forest sites were separated into training (70%) and testing (30%) sets. The resulting species mapping was highly accurate, with an overall accuracy of 85.3% on the testing dataset [2]. The ground truth labels and PLSDA species mapping are visualized in Fig. 1.

Ash dieback disease was identified using a supervised RF classifier [2]. The RF was trained to detect three disease classes—infected, severely infected, and healthy—among trees labeled ash in the PLSDA species mapping. Models were trained and tested using the average pixel spectra of tree crowns, where the testing set included 16 trees from each disease class. The resulting RF classifier demonstrated strong recovery of ash dieback disease in the Madingley scene, with an overall accuracy of 77.1% [2].

3.3. Unsupervised Classification of Ash Dieback Disease

This section presents the results of unsupervised clustering of the Madingley HSI using the D-VIS clustering algorithm. Because ash trees affected by dieback often have a mosaic of healthy and dead branches, the pixels corresponding to visibly infected trees at the original 0.32 m spatial resolution may resemble the pixels of healthy trees [2]. Therefore, to provide a more holistic assessment of tree health (as opposed to that

of the health of individual branches), the HSI was downsampled to a spatial resolution of 1.28 m using bicubic interpolation [16] so that each pixel covered multiple branches.

Pixels corresponding to species other than ash in the PLSDA species mapping were discarded, leaving $n = 72775$ pixels across a 401×279 scene. D-VIS relied on a sparse K-nearest neighbors graph with $N = 150$ edges per pixel, a density scale $\sigma_0 = 3.89 \times 10^{-4}$, and $t = 2^5$. We set $K = 2$ so that classes corresponded to healthy and dieback-infected trees. After cluster analysis, majority voting was implemented among pixels in each delineated tree crown, yielding an unsupervised tree crown-level disease mapping.

For validation, the unsupervised D-VIS clustering was compared against the supervised RF classification of ash dieback disease obtained in prior work [2] after aligning labels using the Hungarian algorithm. The two dieback classes in the RF disease mapping (infected and severely infected) were combined, yielding a single "dieback" class. A matching matrix summarizing the overlap between the supervised RF and unsupervised D-VIS labelings is provided in Table 1, while disease mappings are visualized in Fig. 2. D-VIS and RF disease mappings exhibited a high level of overlap, with D-VIS achieving an overall accuracy of 71.0% and average accuracy of 71.3%. Thus, unsupervised clustering algorithms such as D-VIS may be used for the detection of ash dieback disease using hyperspectral data, even when no ground truth labels are available.

4. CONCLUSIONS

We conclude that the unsupervised D-VIS clustering algorithm can successfully identify ash dieback disease from remote sensing HSIs. Future work includes implementing the pipeline developed in this article on additional forest regions as further validation, as well as scaling up our approach to be implemented on larger forests. Moreover, we expect the unsupervised learning procedure outlined in this article to be useful for the remote detection of other damaging agents in forests and hope to consider this problem in future work. Finally, the performance of D-VIS is likely to improve upon further modification (e.g., modifying for active learning [13, 17, 18] or adding spatial regularization [18, 19, 20]).

Ground Truth Labels   PLSDA Species Mapping
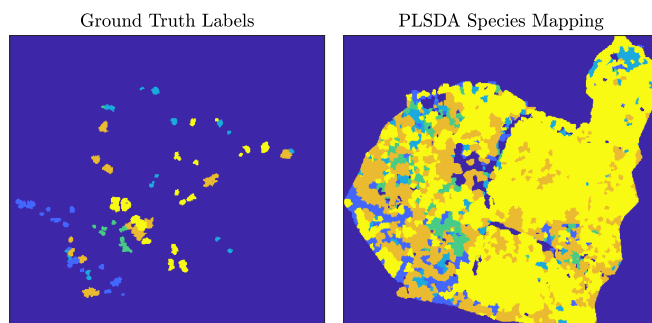
Fig. 1: Ground truth labels and PLSDA species mapping for the Madingley HSI [2]. Colors indicate different tree species. The class marked in yellow corresponds to ash trees.

	Healthy	**Dieback**	**Producer's Acc.**
Healthy	**27460**	12895	68.0%
Dieback	8238	**24182**	74.6 %
User's Acc.	76.9%	65.2%	

Table 1: Matching matrix showing overlap between the unsupervised D-VIS and supervised RF ash dieback mappings [2]. Rows summarize how pixels labeled by the RF in a fixed class were classified by D-VIS. "Acc." indicates Accuracy.

Supervised RF Disease Mapping   Unsupervised D-VIS Disease Mapping
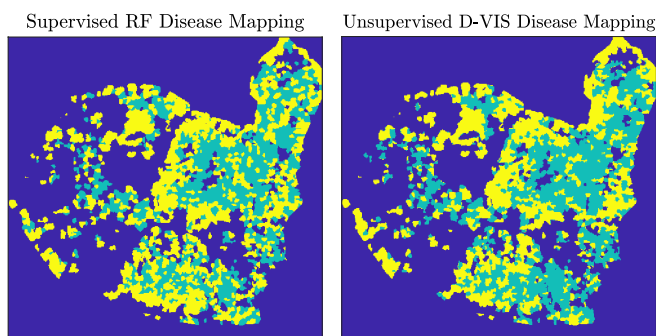
Fig. 2: Supervised RF [2] and unsupervised D-VIS disease mappings of the Madingley HSI. Significant overlap exists between the two labelings, indicating that unsupervised methods such as D-VIS may be used for ash dieback disease mapping when no ground truth labels are available.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] L. V. McKinney et al., "The ash dieback crisis: genetic variation in resistance can prove a long-term solution," *Plant Pathol*, vol. 63, no. 3, pp. 485–499, 2014.

[2] A. H. Y. Chan, C. Barnes, T. Swinfield, and D. A. Coomes, "Monitoring ash dieback (*Hymenoscyphus fraxineus*) in British forests using hyperspectral remote sensing," *Remote Sens Ecol Conserv*, vol. 7, no. 2, pp. 306–320, 2021.

[3] C. Stone and C. Mohammed, "Application of remote sensing technologies for assessing planted forests damaged by insect pests and fungal pathogens: a review," *Curr For Rep*, vol. 3, no. 2, pp. 75–92, 2017.

[4] L. T. Waser, M. Küchler, K. Jütte, and T. Stampfer, "Evaluating the potential of Worldview-2 data to classify tree species and different levels of ash mortality," *Remote Sens*, vol. 6, no. 5, pp. 4515–4545, 2014.

[5] S. L. Polk, K. Cui, R. J. Plemmons, and J. M. Murphy, "Diffusion and volume maximization-based clustering of highly mixed hyperspectral images," *arXiv preprint arXiv:2203.09992*, 2022.

[6] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, Springer, 2001.

[7] M. T. Eismann, *Hyperspectral remote sensing*, SPIE, 2012.

[8] R. R. Coifman and S. Lafon, "Diffusion maps," *Appl Comput Harm Anal*, vol. 21, no. 1, pp. 5–30, 2006.

[9] M. Maggioni and J. M. Murphy, "Learning by unsupervised nonlinear diffusion," *J Mach Learn Res*, vol. 20, no. 160, pp. 1–56, 2019.

[10] J. M. Murphy and S. L. Polk, "A multiscale environment for learning by diffusion," *Appl Comput Harm Anal*, vol. 57, pp. 58–100, 2022.

[11] T. Chan, W. Ma, A. Ambikapathi, and C. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Trans Geosci Remote Sens*, vol. 49, no. 11, pp. 4177–4193, 2011.

[12] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans Geosci Remote Sens*, vol. 43, no. 4, pp. 898–910, 2005.

[13] S. L. Polk, K. Cui, R. J. Plemmons, and J. M. Murphy, "Active diffusion and VCA-assisted image segmentation of hyperspectral images," *arXiv preprint arXiv:2204.06298*, 2022.

[14] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans Geosci Remote Sens*, vol. 46, no. 8, pp. 2435–2445, 2008.

[15] R. Bro and S. De Jong, "A fast non-negativity-constrained least squares algorithm," *J Chemom*, vol. 11, no. 5, pp. 393–401, 1997.

[16] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans Signal Process*, vol. 29, no. 6, pp. 1153–1160, 1981.

[17] M. Maggioni and J. M. Murphy, "Learning by active nonlinear diffusion," *Found Data Sci*, vol. 1, no. 3, pp. 271, 2019.

[18] J. M. Murphy, "Spatially regularized active diffusion learning for high-dimensional images," *Pattern Recognit Lett*, vol. 135, pp. 213–220, 2020.

[19] J. M. Murphy and M. Maggioni, "Spectral–spatial diffusion geometry for hyperspectral image clustering," *IEEE Geosci Remote Sens Lett*, vol. 17, no. 7, pp. 1243–1247, 2019.

[20] S. L. Polk and J. M. Murphy, "Multiscale clustering of hyperspectral images through spectral-spatial diffusion geometry," in *Proc IEEE Geosci Remote Sens Symp*, 2021, pp. 4688–4691.