PROBABILISTIC DEEP DISCRIMINANT ANALYSIS FOR WIND BLADE SEGMENTATION

Raül Pérez-Gonzalo[1,2,*], *Andreas Espersen*[2] *and Antonio Agudo*[1]

[1]Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Barcelona, Spain
[2]Wind Power LAB, Copenhagen, Denmark

ABSTRACT

Linear discriminant analysis improves class separability but struggles with non-linearly separable data. To overcome this, we introduce Deep Discriminant Analysis (DDA), which directly optimizes the Fisher criterion utilizing deep networks. To ensure stable training and avoid computational instabilities, we incorporate signed between-class variance, bound outputs with a sigmoid function, and convert multiplicative relationships into additive ones. We present two stable DDA loss functions and augment them with a probability loss, resulting in Probabilistic DDA (PDDA). PDDA effectively minimizes class overlap in output distributions, producing highly confident predictions with reduced within-class variance. When applied to wind blade segmentation, PDDA showcases notable advances in performance and consistency, critical for wind energy maintenance. To our knowledge, this is the first application of DDA to image segmentation.

***Index Terms*—** Class Separability, Fisher Criterion, Deep Discriminant Analysis, Probability Loss, Blade Segmentation

1. INTRODUCTION

Linear Discriminant Analysis (LDA) is a popular linear classification technique that projects data onto a lower-dimensional space to enhance class separability. This projection is guided by the Fisher criterion, defined as the ratio of between-class to within-class variance [1]. Apart from being computationally efficient and versatile for the multi-class setting, LDA provides robust and reliable results when handling multicollinearity and normality [2].

Despite its strengths, LDA struggles with non-linearly separable data [3]. To address this, various extensions have been proposed, including uncertainty modeling for scatter matrices [4], regularization techniques like Tikhonov and subspace regularization [5], and advanced discriminant models with greater number of discriminants such as orthogonal [6], local [7], and heteroscedastic LDA [8]. LDA norm-based variants [9], kernel methods [10] and, more recently, the integration with deep learning, termed as DDA [11, 12], have shown promising advances in addressing LDA's limitations.

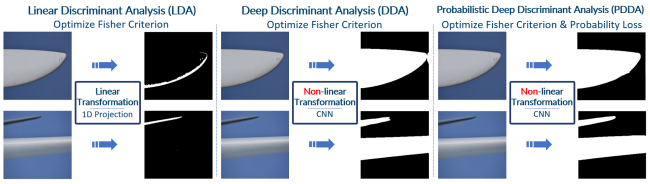

Fig. 1: **Enhancing segmentation via non-linear class separability and probability loss**. Probabilistic DDA corresponds specifically to our proposed PDDA framework.

DDA handles complex signals like images by mapping input data into a deep subspace with better class separability. It typically optimizes a Fisher-like criterion [13, 14, 15] or an equivalent least-squares problem [16]. However, these approaches often rely on reformulations of the Fisher criterion or require eigenvalue and scatter matrix computations, becoming computationally costly and unstable, respectively.

Conversely, modern image segmentation relies on encoder-decoder architectures [17, 18], with recent developments incorporating attention mechanisms to refine feature extraction [19, 20, 21]. Lightweight models enable efficiency in constrained environments [22, 23], while zero-shot models like SAM [24] generalize to domain-agnostic tasks [25, 26].

This work proposes optimizing the Fisher criterion through a Convolutional Neural Network (CNN) by adapting it to a stable loss. Novel DDA optimization is achieved by enforcing the positive class to be projected to higher values, preventing extreme gradients through the sigmoid function. This adaptation enables bounded outputs, additive relationships, and smoother gradients, facilitating robust training. We introduce two novel DDA loss functions: a log-scaled Fisher criterion and a linear-based decomposition, and combine them with a probability-based loss [27] to form Probabilistic DDA (PDDA) (see Fig. 1), providing stronger supervision for low-contrast boundaries, where standard losses often struggle.

We validate PDDA in image segmentation, becoming, to our knowledge, the first application of DDA to this task. Tailored for lightweight models with strong separation objectives, PDDA achieves accurate wind blade segmentation with limited labeled data [28], capturing subtle textures and boundaries under complex conditions. This yields robust, cost-efficient solutions for the wind industry settings [29].

This work has been partially supported by the project GRAVATAR PID2023-151184OB-I00 funded by MCIU/AEI/10.13039/501100011033 and by ERDF, UE; and by the project GreenVAR of the Fundación Ramón Areces.

2. DEEP DISCRIMINANT ANALYSIS

LDA [1] projects data into a lower-dimensional space where class separability is maximized. By integrating class labels into its objective, LDA ensures that feature vectors $\mathbf{x}$ from different classes remain well separated after projection. For two classes, LDA computes a linear projection $y = \mathbf{w}^\top \mathbf{x}$ that best separates the background (C_0) and blade (C_1) classes, where $\mathbf{x}$ represents an RGB pixel and $\mathbf{w}$ is a learned weight vector. This projection is optimized by minimizing the inverse Fisher criterion:

$$\mathcal{L}_{LDA} = \frac{s_{C_0}^2 + s_{C_1}^2}{(\mu_{C_0} - \mu_{C_1})^2} \,, \tag{1}$$

where the within-class variance is measured as the sum of the projected class variances $s_{C_i}^2$, and the between-class one as the difference of the projected class means μ_{C_i}, for $i \in \{0, 1\}$. This leads to a generalized eigenvalue problem over $\mathbf{w}$.

Non-linear transformations enhance class separability by mapping data into a higher-dimensional space but cannot be analytically optimized as linear models. This section extends LDA to non-linear cases using a novel iterative gradient-based optimization approach that prevents gradient explosion.

2.1. Towards Gradient Iterative Optimization

We propose adopting a non-linear transformation $y = f(\mathbf{x})$ such as a neural network to separate classes. As finding the optimal non-linear solution analytically is no longer possible, we need to leverage gradient-based optimization techniques. However, directly optimizing the Fisher criterion from Eq. (1) leads to training instability, as the gradient of the between-class variance $\mathcal{G} = \nabla(\mu_{C_0} - \mu_{C_1})^2$ can point in opposite directions depending on the sampled instance:

$$\mathcal{G} = 2(\mu_{C_0} - \mu_{C_1})(\nabla\mu_{C_0} - \nabla\mu_{C_1}) \begin{cases} < 0 \text{ if } \mu_{C_1} > \mu_{C_0} \\ > 0 \text{ if } \mu_{C_0} > \mu_{C_1} \end{cases} . \tag{2}$$

To resolve this, we remove the squaring operation and instead minimize the signed mean difference $(\mu_{C_0} - \mu_{C_1})$, while replacing the variance ratio with a product. This allows the model to learn a directional relationship where class C_1 is projected to higher values than C_0, i.e., $f(\mathbf{x}_{C_1}) > f(\mathbf{x}_{C_0})$. This reformulation enables stable gradient optimization while maintaining a discriminative objective:

$$\mathcal{L} = (\mu_{C_0} - \mu_{C_1})(s_{C_0}^2 + s_{C_1}^2) \,. \tag{3}$$

2.2. Preventing Exploding Gradients

In the linear setting, $\mathcal{L}_{LDA}$ (Eq. (1)) is scale-invariant, as rescaling the input $\mathbf{x}$ preserves the ratio of within-class to between-class variance. However, this invariance does not hold for non-linear transformations, where scaling the projected features can artificially minimize $\mathcal{L}_{LDA}$ without improving class separability. As a result, the network may increase feature magnitudes excessively, leading to exploding gradients. To counteract this, we constrain the output of the non-linear projection $f(\mathbf{x})$ using a sigmoid activation σ, bounding the output to $[0, 1]$ and stabilizing gradient updates.

Beyond numerical stability, bounding the feature space also simplifies estimation of the discriminant threshold $\mathcal{T}$ for classifying pixels as background (C_0) or blade (C_1). A projected pixel $y = f(\mathbf{x})$ is assigned to C_1 if $\sigma(y) \geq \mathcal{T}$, and to C_0 otherwise. While in binary cross entropy the threshold is typically set near 0.5, the Fisher criterion does not originate from a probability derivation, so the optimal $\mathcal{T}$ must be found from the projected space, which is a priori unknown.

By restricting the output space to $[0, 1]$, we enable an efficient class-threshold search to evaluate segmentation performance across candidate thresholds and select the one that maximizes class separability.

2.3. Direct Deep Discriminant Analysis

Building on Sec. 2.1 and Sec. 2.2 over Eq. (1), using the non-squared class mean difference for gradient stability and bounding the projected outputs via a sigmoid, we define two loss functions to enable DDA. These formulations are designed to facilitate gradient-based optimization while maintaining interpretability with respect to class separability. We introduce a parameter $\lambda_F \in \mathbb{R}$ to balance the contribution of class variance terms.

Logarithmic DDA Loss. To balance the contribution of both class separability terms, we transform the division relationship in Eq. (3) into an additive one by applying logarithms. Since the signed between-class variance $\mu_{C_0} - \mu_{C_1}$ lies in $[-1, 1]$, we add a constant $\epsilon = 1 + 10^{-8}$ to ensure the logarithm remains within its domain and apply it to both variance terms, ensuring that both contribute equally, writing:

$$\mathcal{L}_{DDA}^{(\ln)} = \ln\left(\epsilon + \mu_{C_0} - \mu_{C_1}\right) + \lambda_F \ln\left(\epsilon + s_{C_0}^2 + s_{C_1}^2\right) \,. \tag{4}$$

Delta DDA Loss. An alternative separability criterion is to minimize the delta difference between class variances $(s_{C_0}^2 - s_{C_1}^2) - (\mu_{C_0} - \mu_{C_1})^2$ [3]. Following a similar rationale, we can rearrange this criterion to adopt gradient optimization by adopting the signed between-class variance, and changing the delta difference to a sum, aiming that the blade pixels are projected to higher values than the background ones:

$$\mathcal{L}_{DDA}^{(\Delta)} = (\mu_{C_0} - \mu_{C_1}) + \lambda_F \left(s_{C_0}^2 + s_{C_1}^2\right) \,. \tag{5}$$

Given m the binary ground-truth class, the conventional class-wise means and variances are redefined for each class $i \in \{0, 1\}$ using differential formulations over the dataset $\mathcal{D}$.

$$\mu_{C_i} = \frac{1}{|C_i|} \sum_{y \in \mathcal{D}} m^i (1 - m)^{(1-i)} \sigma(y) \,, \tag{6}$$

$$s_{C_i}^2 = \frac{1}{|C_i| - 1} \sum_{\mathbf{y} \in \mathcal{D}} m^i (1 - m)^{(1-i)} (\sigma(y) - \mu_{C_i})^2 \,, \tag{7}$$

$$C_i = \sum_{y \in \mathcal{D}} m^i (1 - m)^{(1-i)} \,. \tag{8}$$

2.4. Probabilistic Deep Discriminant Analysis

While class separability losses guide $f(\mathbf{x})$ to learn discriminative projections, they lack probabilistic interpretation. In contrast, binary cross-entropy reflects the negative log-likelihood of a Bernoulli distribution, enabling outputs to be treated as class probabilities. To bridge this gap, we combine class separability with a probability-based loss [27], which emphasizes learning from low-confidence examples.

Focal Loss. This loss enhances the segmentation mapping by penalizing low-confidence predictions with γ parameter. To prevent bias towards the majority class (background, in this context), class-specific weights α set the class balance:

$$\mathcal{L}_P = -\sum_{y\in\mathcal{D}} \left[\alpha\,(1-\sigma(y))^{\gamma}\, m \ln \sigma(y) \right. \\ \left. +(1-\alpha)\sigma(y)^{\gamma}(1-m)\ln(1-\sigma(y))\right] . \quad (9)$$

Probabilistic DDA Loss. We define the final loss as a linear combination of the DDA loss (with the superscript indicating the chosen DDA) and Focal Loss, weighted by λ_P:

$$\mathcal{L}_{PDDA} = \mathcal{L}_P + \lambda_P \mathcal{L}_{DDA} . \quad (10)$$

2.5. Implementation Details

Input images are resized to 256×256, min-max normalized, and augmented with flipping and cropping. We adopt U-Net [30] as the backbone architecture. We use Adam optimizer with an initial learning rate of 10^{-4} and batch size 8. A custom scheduler reduces the learning rate after three validation plateaus. The variance balance parameters are $\lambda_F = 0.4$ for $\mathcal{L}_{DDA}^{(\Delta)}$ and $\lambda_F = 0.9$ for $\mathcal{L}_{DDA}^{(\ln)}$. Focal loss parameter are fine-tuned over the validation set to $\gamma = 2$ and $\alpha = 0.25$.

3. EXPERIMENTAL RESULTS

This section compares LDA and DDA performance and, then, presents both quantitative and qualitative evaluations to showcase the effectiveness of PDDA, highlighting the methods' robustness by comparing results across different windfarms. The wind blade dataset employed is taken from [28].

3.1. Direct Deep Discriminant Analysis

We first explore the benefits of using a non-linear transformation as a deep network for optimizing discriminant analysis. As shown in Tab. 1, DDA significantly outperforms LDA. Both optimization objectives, $L_{DDA}^{(\Delta)}$ in Eq. (5) and $L_{DDA}^{(\ln)}$ in Eq. (4), yield comparable gains over LDA, with accuracy improving by about 20% and recall doubling. A dummy classifier predicting only the majority class (background) already achieves 69% accuracy, close to LDA's 72.59%, confirming that the blade segmentation task is non-linearly separable. From a qualitative perspective, Fig. 2 illustrates that LDA often fails in complex backgrounds or only partially segments the blade, whereas DDA correctly identifies the blade region in most cases (last two rows). Remaining challenges (top two rows) are further addressed by PDDA in Sec. 3.4.

Method	λ_P	Accuracy [%]	F1 [%]	mIoU [%]	$\mu_{C_1} - \mu_{C_0}$ ↑	$s^2_{C_0}$ ↓	$s^2_{C_1}$ ↓
LDA	-	72.59	58.46	53.67	-	-	-
DDA$^{(\Delta)}$	-	94.06	91.52	87.11	.81	.033	.14
DDA$^{(\ln)}$	-	95.50	91.59	88.70	.83	.023	.13
$\mathcal{L}_P$	-	96.28	94.05	90.91	.69	.007	.05
PDDA$^{(\Delta)}$	.0001	94.27	85.12	86.34	.47	.007	.02
PDDA$^{(\Delta)}$	.001	95.15	93.23	90.12	.57	.008	.04
PDDA$^{(\Delta)}$	.01	96.82	94.57	91.87	.66	.010	.07
PDDA$^{(\Delta)}$	.1	97.10	95.42	93.05	.77	.010	.10
PDDA$^{(\Delta)}$	1	**97.62**	**95.53**	**93.29**	.77	.007	.10
PDDA$^{(\Delta)}$	2	95.82	91.47	90.07	.82	.016	.09
PDDA$^{(\ln)}$	.0001	96.51	93.04	90.85	.49	.010	.05
PDDA$^{(\ln)}$	.001	96.56	93.48	91.57	.54	.009	.05
PDDA$^{(\ln)}$	.01	97.19	94.80	92.32	.64	.014	.06
PDDA$^{(\ln)}$	.1	**97.97**	**96.28**	**94.34**	.76	.012	.11
PDDA$^{(\ln)}$	1	96.77	94.82	92.23	.74	.009	.17
PDDA$^{(\ln)}$	2	96.00	90.70	89.73	.80	.016	.13

Table 1: **Quantitative comparison** of $\mathcal{L}_P$, DDA and PDDA.

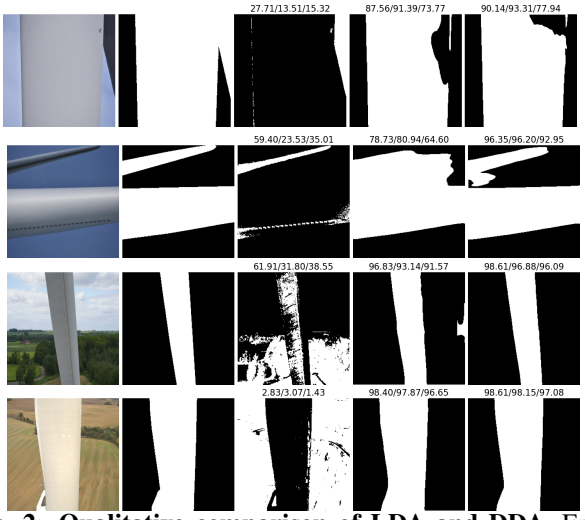

Fig. 2: **Qualitative comparison of LDA and DDA.** From left-to-right: input image, ground-truth mask, LDA, DDA$^{(\Delta)}$ and DDA$^{(\ln)}$ estimations with accuracy, F1, and mIoU above.

3.2. Probabilistic Deep Discriminant Analysis

Detailed insights into the adoption of PDDA are provided in Tab. 1, including performance metrics and class variance breakdowns. Both $\mathcal{L}_P$ and DDA achieve remarkable segmentation performance, but $\mathcal{L}_P$ overperforms DDA, because it follows the state-of-the-art BU-Net [28] designed for blade wind segmentation. Although the between-class variance $(\mu_{C_1} - \mu_{C_0})$ is not as pronounced as with DDA, optimizing $\mathcal{L}_P$ results in significantly lower within-class variances ($s^2_{C_i}$ for $i \in \{0, 1\}$), even though it does not directly minimize them. By balancing the probability and DDA terms with $\mathcal{L}_P$, PDDA retains the high between-class variance from DDA and the low within-class variances from $\mathcal{L}_P$, leading to top-performing segmentation metrics. Note that PDDA$^{(\Delta)}$ is balanced with $\lambda_P = 1$, whereas PDDA$^{(\ln)}$ requires $\lambda_P = 0.1$ due to $10 \cdot \mathcal{L}_{DDA}^{(\Delta)} \approx \mathcal{L}_{DDA}^{(\ln)}$ over the range $[0, 1]$.

3.3. Quantitative Evaluation

Table 2 provides a quantitative comparison of several state-of-the-art segmentation methods, trained from scratch following

Method	Accuracy [%]	Precision [%]	Recall [%]	F1 [%]	mIoU [%]	IoU_{C_0} [%]	IoU_{C_1} [%]
U-Net [30]	86.24	95.51	68.93	77.95	75.94	80.31	71.57
DeepLabv3+ [17]	94.14	96.36	87.38	89.03	87.47	90.31	84.62
SW [18]	93.48	93.57	91.71	91.37	87.44	88.64	86.23
ResNeSt [20]	94.23	96.84	91.47	92.77	89.63	90.40	88.86
SAM [24]	94.36	97.29	91.22	92.60	91.66	92.31	91.01
CLIPSeg [25]	82.70	77.02	75.52	74.29	75.09	80.16	70.02
DiffSeg [26]	96.37	82.08	89.74	85.73	86.40	91.66	81.13
EfficientFormer [23]	96.42	95.47	93.63	94.55	93.51	94.02	92.99
MobileViT [22]	96.14	95.44	93.33	94.38	93.47	94.06	92.88
Mask2Former [19]	96.68	95.63	93.89	94.76	93.72	94.51	92.93
BU-Net [28]	97.39	**99.42**	93.35	95.73	93.80	94.70	92.90
$\text{PDDA}^{(\Delta)}$	97.62	94.29	97.79	95.53	93.29	94.36	92.21
$\text{PDDA}^{(\text{ln})}$	**97.97**	94.66	**98.71**	**96.28**	**94.34**	**95.28**	**93.40**

Table 2: **Quantitative comparison on blade segmentation** with respect to state-of-the-art methods.

the data pipeline and augmentation from Sec. 2.5. The U-Net [30], trained with binary cross-entropy, is used as the baseline, given its role as our backbone architecture. While it achieves respectable accuracy, its recall rate of 68.93% reveals a notable shortfall in capturing the blade region.

More advanced architectures, such as Mask2Former [19], markedly improve on U-Net, with stronger results across multiple metrics, particularly in recall. Zero-shot models like SAM [24] also achieve impressive quantitative scores. BU-Net [28] surpasses prior state-of-the-art methods as it is tailored to our specific dataset. Still, $\text{PDDA}^{(\Delta)}$ and $\text{PDDA}^{(\text{ln})}$ stand out by delivering top-tier performance compared to these segmentation algorithms. Notably, $\text{PDDA}^{(ln)}$ achieves the highest scores across all metrics except precision, where BU-Net [28] leads. This underscores $\text{PDDA}^{(ln)}$'s superiority, with 97.97% accuracy, 96.28% F1-score and 94.34% mIoU.

Moreover, PDDA's inference computational cost is comparable to the well-studied U-Net [30], offering both an efficient and highly effective segmentation solution.

3.4. Qualitative Evaluation

Despite the high performance metrics, Fig. 3 exemplifies some images where probabilistic and DDA are necessary in conjunction to accomplish high-quality segmentation. This complexity arises from limited data and complex blade shapes, high color transitions within regions, and shadows partially obscuring blade areas. Combining DDA with probability-based loss improves accuracy, overperforming models trained solely with DDA or probability losses. In particular, we can observe that PDDA excels other models in capturing complex boundaries in the hub or additional wind turbine blades behind the primary salient blade.

3.5. Generalization Across Windfarms

To assess PDDA's generalizability and robustness, the test set comprises 20 randomly selected images from 10 distinct windfarms and inspection campaigns [28]. Figure 4 presents the performance distribution per windfarm. The results indicate consistently high performance across all windfarms, confirming the robustness and generalization capabilities of our approach in effectively segmenting unseen test images.

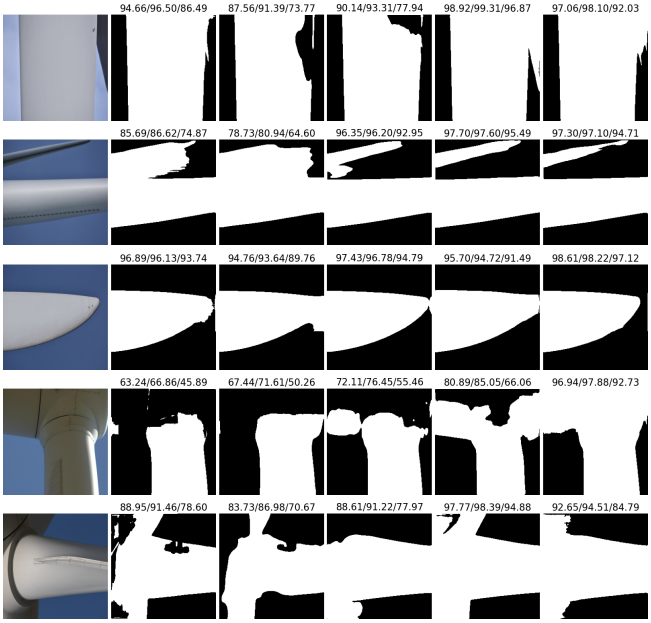

Fig. 3: **Qualitative comparison of $\mathcal{L}_P$, DDA and PDDA.** From left to right columns: input image, $\mathcal{L}_P$, $\text{DDA}^{(\Delta)}$, $\text{DDA}^{(\text{ln})}$, $\text{PDDA}^{(\Delta)}$ and $\text{PDDA}^{(\text{ln})}$ estimations with accuracy, F1-score and mIoU above.

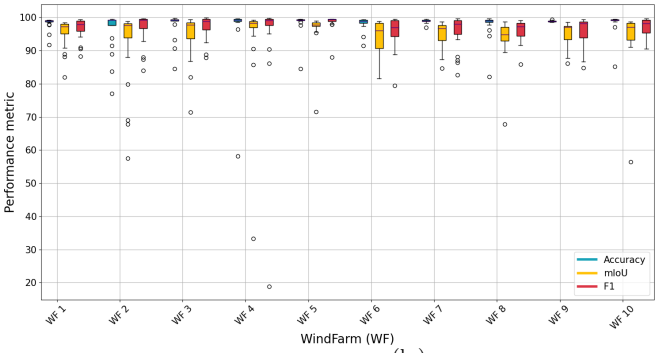

Fig. 4: **Boxplot results of $\text{PDDA}^{(\text{ln})}$ over the test set for each windfarm** to study the robustness of our algorithm.

4. CONCLUSION

LDA is a foundational linear classifier, but its reliance on linear separability limits broader applicability. We presented DDA, which leverages neural networks to directly optimize the Fisher criterion in a stable manner without costly eigen-decomposition. Through novel loss functions, DDA improves class separability and segmentation quality and provides stronger supervision for low-contrast boundaries common in wind blade imagery. We also introduced PDDA, which combines discriminant learning with probability-based loss to reduce class uncertainty and within-class variance. Experiments on wind blade segmentation, a domain with scarce labels and fine-grained structures, demonstrate PDDA's effectiveness and practical impact. By unifying discriminant analysis with probabilistic learning in lightweight models, PDDA provides robust, cost-efficient segmentation and opens new directions for broader industrial applications.

5. REFERENCES

[1] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Ann. Eugenics*, vol. 7, no. 2, pp. 179–188, 1936.

[2] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*, vol. 4, Springer, 2006.

[3] K. Fukunaga, *Introduction to statistical pattern recognition*, Elsevier, 2013.

[4] R. Saeidi, R. F. Astudillo, and D. Kolossa, "Uncertain LDA: Including observation uncertainties in discriminative transforms," *PAMI*, vol. 38, no. 7, pp. 1479–1488, 2015.

[5] X. Zhao, J. Guo, F. Nie, L. Chen, Z. Li, and H. Zhang, "Joint principal component and discriminant analysis for dimensionality reduction," *PAMI*, vol. 31, no. 2, pp. 433–444, 2019.

[6] P.-F. Hsieh, D.-S. Wang, and C.-W. Hsu, "A linear feature extraction for multiclass classification problems based on class mean and covariance discriminant information," *PAMI*, vol. 28, no. 2, pp. 223–235, 2005.

[7] M. Sugiyama, "Dimensionality reduction of multimodal labeled data by local fisher discriminant analysis," *JMLR*, vol. 8, no. 5, 2007.

[8] W. Zheng, C. Lu, Z. Lin, T. Zhang, Z. Cui, and W. Yang, "ℓ_1-norm heteroscedastic discriminant analysis under mixture of gaussian distributions," *NNLR*, vol. 30, no. 10, pp. 2898–2915, 2019.

[9] R. Martín-Clemente and V. Zarzoso, "LDA via L1-PCA of whitened data," *TSP*, vol. 68, pp. 225–240, 2020.

[10] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *NC*, vol. 12, no. 10, pp. 2385–2404, 2000.

[11] J. Mao and A. K. Jain, "Discriminant analysis neural networks," in *ICNN*, 1993, vol. 1, pp. 300–305.

[12] Q. Wang, Z. Qin, F. Nie, and Y. Yuan, "Convolutional 2D LDA for nonlinear dimensionality reduction," in *IJCAI*, 2017, pp. 2929–2935.

[13] D. Díaz-Vico, A. Omari, A. Torres-Barrán, and J. R. Dorronsoro, "Deep fisher discriminant analysis," in *ACI*, 2017, pp. 501–512.

[14] L. Li, M. Doroslovački, and M. H. Loew, "Discriminant analysis deep neural networks," in *CISS*, 2019.

[15] G. Zhong, Y. Zheng, X.-Y. Zhang, H. Wei, and X. Ling, "Convolutional discriminant analysis," in *ICPR*, 2018.

[16] D. Díaz-Vico and J. R. Dorronsoro, "Deep least squares fisher discriminant analysis," *NNLR*, vol. 31, no. 8, pp. 2752–2763, 2020.

[17] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *ECCV*, 2018.

[18] X. Pan, X. Zhan, J. Shi, X. Tang, and P. Luo, "Switchable whitening for deep representation learning," in *ICCV*, 2019.

[19] B. Cheng, I. Misra, A. G. Schwing, A. Kirillov, and R. Girdhar, "Masked-attention mask transformer for universal image segmentation," in *CVPR*, 2022.

[20] H. Zhang, C. Wu, Z. Zhang, Y. Zhu, H. Lin, Z. Zhang, Y. Sun, T. He, J. Mueller, R. Manmatha, M. Li, and A. Smola, "ResNeSt: Split-attention networks," in *CVPRW*, 2022.

[21] S. Singhal, R. Pérez-Gonzalo, A. Espersen, and A. Agudo, "Dual-space augmented intrinsic-lora for wind turbine segmentation," in *ICASSP*, 2025.

[22] S. Mehta and M. Rastegari, "MobileViT: Lightweight, general-purpose, and mobile-friendly vision transformer," in *ICLR*, 2022.

[23] Y. Li et al., "Rethinking vision transformers for mobilenet size and speed," in *ICCV*, 2023.

[24] A. Kirillov et al., "Segment anything," in *ICCV*, 2023, pp. 4015–4026.

[25] T. Lüddecke and A. Ecker, "Image segmentation using text and image prompts," in *CVPR*, 2022.

[26] J. Tian, L. Aggarwal, A. Colaco, Z. Kira, and M. Gonzalez-Franco, "Diffuse attend and segment: Unsupervised zero-shot segmentation using stable diffusion," in *CVPR*, 2024.

[27] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *ICCV*, 2017.

[28] R. Pérez-Gonzalo, A. Espersen, and A. Agudo, "Robust wind turbine blade segmentation from RGB images in the wild," in *ICIP*, 2023.

[29] R. Pérez-Gonzalo, A. Espersen, and A. Agudo, "Generalized nested latent variable models for lossy coding applied to wind turbine scenarios," in *ICIP*, 2024.

[30] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015.