



Addressing Class Imbalance in Oil Palm Disease and Micronutrient Deficiency Detection Using Meta-Learned Transfer Metric Learning

Hartono¹, Erianto Ongko²

¹Department of Informatics, Faculty of Engineering, Universitas Medan Area, Medan, Indonesia

²Department of Computer Science, Institut Modern Arsitektur dan Teknologi, Medan, Indonesia

¹hartono@staff.uma.ac.id, ²eriantoongko9@gmail.com

Abstract

Class imbalance is a major challenge in oil palm disease and nutrient deficiency detection, where healthy samples dominate while diseased or deficient cases are underrepresented, often leading to biased models with high false-negative rates. To address this issue, this study proposes MetaTMLDA (Meta-Learned Transfer Metric Learning with Distribution Alignment), a hybrid framework that combines Transfer Metric Learning (TML) with MW-FixMatch. TML learns discriminative and domain-invariant features, while MW-FixMatch employs a meta-learned weighting mechanism to adaptively reweight samples, improving sensitivity to minority classes and enhancing robustness against pseudo-label noise. Experiments on four public datasets—Ganoderma Disease Detection, Palm Oil Leaf Disease, and Leaf Nutrient Detection for Boron and Magnesium—demonstrated that the proposed method consistently outperforms TML-DA, MW-FixMatch, SMOTE, Random Undersampling, and Biased SVM. On the smaller datasets (Ganoderma and Palm Oil Leaf Disease), MetaTMLDA achieved accuracy of 0.976, precision 0.951, recall 0.915, Cohen's Kappa 0.912, and macro F1-score 0.933 for Ganoderma, and accuracy of 0.980, precision 0.972, recall 0.957, Kappa 0.911, and macro F1-score 0.964 for Palm Oil Leaf Disease. On the larger datasets (Boron and Magnesium), the model reached near-perfect accuracy of 0.995, with precision up to 0.967, recall up to 0.973, Kappa above 0.919, and macro F1-scores up to 0.969, highlighting its robustness and balanced predictive performance. These findings confirm that MetaTMLDA effectively addresses both class imbalance and domain shift, providing a scalable solution for precision agriculture through earlier and more reliable detection of oil palm health issues.

Keywords: class imbalance; MW-FixMatch; MetaTMLDA; oil palm; transfer metric learning;

How to Cite: H. Hartono and E. Ongko, "Addressing Class Imbalance in Oil Palm Disease and Micronutrient Deficiency Detection Using Meta-Learned Transfer Metric Learning", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 9, no. 5, pp. 1273 - 1282, Oct. 2025.

Permalink/DOI: <https://doi.org/10.29207/resti.v9i5.6857>

Received: June 28, 2025

Accepted: October 18, 2025

Available Online: October 28, 2025

This is an open-access article under the CC BY 4.0 License

Published by Ikatan Ahli Informatika Indonesia

1. Introduction

Class imbalance remains a persistent challenge in supervised and semi-supervised machine learning[1], particularly in classification tasks where one class is significantly underrepresented[2]. In binary settings, this typically manifests as a skewed distribution of labels, where the majority class dominates the dataset [3], and the minority class, often representing critical or anomalous cases, is underrepresented or sparsely sampled[4]. This issue is especially pronounced in agricultural applications, such as disease detection[5] and micronutrient deficiency detection in oil palm plantations[6].

Here, healthy plants are overwhelmingly common in the data, while instances of specific diseases or nutrient

deficiencies are rare but agronomically vital to identify early[7]. In such cases, machine learning models trained on imbalanced datasets are prone to bias toward the majority class[8], resulting in high false-negative rates [9]. This is detrimental to plantation management, where failure to detect early signs of Ganoderma infection[10] or boron deficiency, for instance, can lead to significant yield loss and economic impact[11].

Moreover, collecting and annotating data for the minority class in agricultural environments presents unique challenges, including seasonal availability, dependency on experts, and image variability across different locations or devices[12]. These factors introduce domain shift—a discrepancy between labeled source data (e.g., from one plantation or sensor type) and unlabeled target data (e.g., from a different region

or camera)[13]. Traditional supervised or semi-supervised learning techniques often fail in such scenarios, as they assume balanced data and distributional homogeneity between training and deployment contexts[14].

To address the dual challenges of class imbalance and domain shift, recent research has explored advanced learning paradigms that combine meta-learning and meta-training[15]. Moniz and Cerqueira proposed ATOMIC to predict the best workflows based on past data, utilizing meta-features extracted from datasets and a meta-model trained with XGBoost to efficiently estimate workflow performance[16]. Chamlal et al. proposed the Hybrid Multi-Criteria Meta-Learner (HML), a meta-learning approach designed to handle imbalanced data by integrating ensemble learning, multi-objective optimization, and clustering[17]. Guo et al. proposed a meta-learning-based approach called MAEL to address class imbalance in medical data. MAEL adaptively learns sampling strategies using a meta-learner and reinforcement learning to focus on instances of the minority class [18].

The use of meta-learning, often combined with meta-training and transfer learning, has proven effective in addressing data bias and enhancing transfer networks. For example, Jindal and Singh[19] proposed CIDSTL-Net, which adopts a two-stage training approach: first, developing domain-specific models, and then fine-tuning them on target medical datasets. Isnani et al[20]. Applied transfer learning using deep learning to classify 29 types of Indonesian crops with limited data. By fine-tuning the final layer of pre-trained CNN models- EfficientNet-B0, ResNet-18, VGG-19, and AlexNet -they achieved efficient training and improved accuracy.

Another method that combines meta-learning and transfer learning is Transfer Metric Learning-Based Distribution Alignment (TML-DA). This approach learns a discriminative and transferable metric space to align the source and target domains effectively[21]. TML-DA exploits within-class and between-class relationships, derived from both source samples and pseudo-labeled target samples, to compute metric losses that preserve the local class structure. Additionally, a distribution alignment loss is incorporated to encourage the learning of domain-invariant representations.

Compared to meta-learning frameworks like MLBE-ICF and MAEL, which focus on classifier adaptability and sample balancing, TML-DA stands out by learning transferable and discriminative features without relying on extensive labeled data or complex ensembles. While meta-learning often struggles with limited generalizability due to hyperparameter sensitivity, TML-DA efficiently aligns both marginal and conditional distributions in homogeneous, transductive settings. Unlike hybrid methods such as HML that require computationally intensive ensemble generation and clustering, TML-DA offers a more straightforward yet robust approach through iterative distribution

alignment and pseudo-labeling based on inter-domain similarity.

Although TML-DA outperforms meta-learning and transfer learning methods in terms of overall performance, it still has notable limitations. Specifically, it relies on pseudo-labels generated through nearest neighbor selection, which may introduce confirmation bias if early predictions are inaccurate[22]. Moreover, it treats all pseudo-labeled samples equally, lacking an adaptive weighting mechanism to account for varying learning difficulties and class imbalances[23]. These limitations can be addressed by MW-FixMatch, which employs a meta-learned weight network to adaptively reweight both supervised and unsupervised losses based on a class-balanced meta-dataset, resulting in more robust learning under conditions of class imbalance and uncertain labeling.

MW-FixMatch, a class-imbalance-aware semi-supervised learning based on the FixMatch framework, introduces a meta-learned weight network to address this problem[24]. It reweights each sample's contribution to the loss function during training, based on its reliability, using a meta-dataset that reflects balanced class distributions. This adaptive reweighting helps mitigate the dominance of the majority class and guides the model to focus on informative minority-class samples, even in the presence of pseudo-label noise. Despite its robustness to imbalance, MW-FixMatch does not explicitly address domain shift, which limits its applicability in cross-plantation or cross-device scenarios commonly encountered in oil palm monitoring.

To overcome the limitations of each method, we propose an integrated approach that combines TML-DA and MW-FixMatch, tailored for the detection of oil palm diseases and micronutrient deficiencies. The combination of TML-DA and MW-FixMatch offers a robust solution for handling class imbalance and domain distribution mismatch in semi-supervised learning. While MW-FixMatch addresses class imbalance through a re-weighting mechanism that adjusts the importance of labeled and unlabeled data, TML-DA focuses on aligning marginal and conditional distributions between source and target domains using transfer metric learning. When integrated, these methods complement each other: TML-DA enhances domain generalisation by learning domain-invariant features, while MW-FixMatch reduces confirmation bias and improves pseudo-label reliability. Compared to conventional transfer learning approaches, such as standard pre-trained models tested on Indonesian crop classification, the combined method provides stronger performance, particularly in scenarios with limited labels and skewed class distributions.

A balanced meta-dataset guides the update of this network to ensure fair representation and learning across all classes. In the context of oil palm plantation management, this approach facilitates the more accurate

and early identification of diseases and deficiencies, thereby contributing to more innovative and sustainable agricultural practices. By addressing the critical interplay between class imbalance and domain shift, this study introduces a novel framework for meta-learned transfer metric learning, offering both technical advancement and practical relevance in the domain of precision agriculture.

2. Methods

The methodology of this study consists of several key components, starting with Transfer Metric Learning, MW-FixMatch, and their integration in MetaTMLDA, followed by complexity analysis, dataset preparation, and evaluation metrics.

2.1 Transfer Metric Learning-based Distribution Alignment (TML-DA)

The Transfer Metric Learning-Based Distribution Alignment (TML-DA) is designed to learn a transferable metric representation that aligns the distribution between labelled source domain data and unlabeled target domain data. TML-DA begins by initializing with labeled source data D_s , unlabeled target data D_t , the number of training epochs T , a learning rate η , a margin parameter α , the number of neural network layers M , and the number of neurons per layer q_m . For each training epoch, the process constructs within-class pairs D_c^w and between-class pairs D_c^b using Formulas 1 and 2[21].

$$D_c^w \subseteq D^c \times D^c := \{(x_i, x_j) | y_i = y_j = c\} \quad (1)$$

$$D_c^b \subseteq D^c \times D^{\tilde{c}} := \{(x_i, x_j) | y_i = c, y_j = \tilde{c}\} \quad (2)$$

c denotes the c -th class, $\tilde{c}$ denotes a class other than c .

Then, for each within-class and between-class pair, it computes the feature representations of the samples through the network's final layer $f^{(M)}(x_i), f^{(M)}(x_j), f^{(M)}(x_t)$ followed by the computation of similarity using Euclidean distances between these representations[21] as shown in Formula 3.

$$d_{f^{(M)}}(x_i, x_j) = \|f^{(M)}(x_i) - f^{(M)}(x_j)\|_2^2 \quad (3)$$

To better handle domain shift, TML-DA estimates the importance weighting parameter $\frac{Pr_t(x_i)}{Pr_s(x_i)}$ using Formula 4[21].

$$\frac{Pr_t(x)}{Pr_s(x)} \propto \frac{1}{P(S=1|x)} \quad (4)$$

which adjusts for the distribution difference between source and target domains. It also assigns pseudo-labels to target domain samples using a nearest neighbor strategy using Formula 5[21].

$$NNS(x_i^t, T_c^{ne}) = \min_{T_c} \{d_{f^{(M)}}(x_i^t, T_c)\} \quad (5)$$

Two loss functions are then computed: the metric learning loss $\mathcal{L}_{TML}$ using Formula 6[21].

$$\mathcal{L}_{TML} = \sum_{c=1}^C \left(\sum_{(x_i, x_j) \in D_c^w} \frac{Pr_t(x_i)Pr_t(x_j)}{Pr_s(x_i)Pr_s(x_j)} \mathcal{L}_{pull}^c(x_i, x_j) + \sum_{(x_i, x_j) \in D_c^b} \frac{Pr_t(x_i)Pr_t(x_j)}{Pr_s(x_i)Pr_s(x_j)} \mathcal{L}_{push}^c(x_i, x_j) \right) + \lambda \sum_{m=1}^M (\|W\|_F^2 + \|b\|_F^2) \quad (6)$$

which preserves class relationships, and the domain distribution alignment loss $\mathcal{L}_{DDA}$ using Formula 7[21].

$$\mathcal{L}_{DDA} = (\omega_s - \omega_t)^T (\omega_s - \omega_t) + \sum_{c=1}^C (V_s^c - v_t^c)^T (V_s^c - v_t^c) + \{A_1 \sum_{D_c^w} d_{kl}(x_i^t, x_i^s) - A_2 \sum_{D_c^b} d_{kl}(x_i^t, x_i^s)\} \quad (7)$$

which ensures that source and target feature distributions become more similar. For each layer in the network, the parameters (weights $W_{tml}^{(m)}$ and biases $b_{tml}^{(m)}$) are updated using the gradient descent method with respect to the combined loss function. The learning rate is decayed gradually at each epoch to stabilize training. If the change in the total loss between consecutive epochs falls below a small threshold ϵ , the training loop is terminated early, indicating convergence. TML-DA ultimately outputs the learned network parameters that effectively align and transfer knowledge between domains.

2.2 MW-FixMatch

The MW-FixMatch is a meta-learning-based semi-supervised learning method designed to improve performance on imbalanced datasets by dynamically adjusting the influence of each sample during training. The process starts by initializing the parameters of the primary model θ_0 and the meta-weighting network ϕ_0 . For each training epoch from $t = 0$ to $T - 1$, MW-FixMatch samples mini-batches from the labeled dataset D_L and the unlabeled dataset D_U . Next, it computes the supervised loss $\mathcal{L}_{U_j}$, where pseudo-labels are typically used. Using these losses, the process calculates adaptive sample weights w_i and w_j via a weighting function $f_w(\mathcal{L}_u; \phi_t)$, which is parameterized by the current state of the meta-weight network. These weights aim to reflect the importance or difficulty of each sample.

The next step, MW-FixMatch forms a weighted total loss, combining the supervised and unsupervised losses scaled by their respective weights and a balancing hyperparameter λ . This weighted loss is used to simulate an update step for the main model parameters, resulting in a temporary parameter set. To guide the training of the meta-weight network, this process evaluates a meta-loss using a separate meta-dataset D_M , the gradient of this meta-loss is then used to update the meta-parameters ϕ_{t+1} , ensuring that future sample weights are optimized to improve generalization. Using the updated meta-parameters ϕ_{t+1} . This iterative process continues for T epochs, yielding a model trained with dynamic, context-aware sample weighting that improves robustness, especially under class imbalance and noisy labels.

2.3. Meta-Learned Transfer Metric Learning with Distribution Alignment (MetaTMLDA)

This study proposed MetaTMLDA, a novel hybrid framework that integrates Transfer Metric Learning (TML) and a meta-weighted semi-supervised learning strategy, MW-FixMatch. This method outlines the MetaTMLDA, a meta-learning-based method for domain adaptation in class-imbalanced scenarios. The method begins by receiving three types of input data: labeled source domain data D_s , unlabeled target domain data D_t , and a meta-dataset D_m , along with training hyperparameters such as the number of epochs T , learning rates η and β , margin α , and number of neural network layers M . Initially, the parameters of two networks are randomly initialized: a metric network (responsible for learning feature representations) and a weight network (used for adaptive instance reweighting). The method proceeds through a series of training epochs. In each epoch, mini-batches are sampled from the source, target, and meta datasets. Using these samples, within-class and between-class pairs are constructed to enable the metric learning process. Pseudo-labels are generated for the unlabeled target data using a nearest-neighbor approach based on the learned embeddings. For each training pair, the embeddings are extracted from the final layer of the

metric network. Pairwise distances are computed to capture similarities and dissimilarities, followed by the computation of two main loss components: the task-specific metric loss $\mathcal{L}_{TML}$ and the domain distribution alignment loss $\mathcal{L}_{DDA}$. These two losses are summed to obtain a total loss per instance. Each instance loss is then reweighted using a meta-learned function f_w parameterized by ϕ , the weight network. The overall weighted loss is computed, and the metric network parameters are updated via gradient descent using this weighted loss. Following this, a meta-loss is computed on the meta-dataset using an updated model to guide the optimization of the weight network parameters ϕ . The weight network is updated accordingly, improving its ability to assign meaningful importance to instances during training. A learning rate decay is applied each epoch, and the training loop includes an early stopping condition based on the convergence of total loss changes. The final outputs are the optimized parameters of the metric network and the weight network, which together enable effective domain adaptation through task-specific metric learning and meta-learned reweighting.

A comparison of TML-DA, MW-FixMatch, and MetaTMLDA is presented in Table 1[21], [24].

Table 1. Comparison of TML-DA, MW-FixMatch, and MetaTMLDA

Feature / Step	TML-DA	MW-FixMatch	MetaTMLDA
Main Goal	Transfer metric learning with distribution alignment for domain adaptation	Semi-supervised learning with FixMatch enhanced by meta-weighting	Combination of transfer metric learning, distribution alignment, and meta-learning-based weight adjustment
Input Data	Labeled source data (Ds) and unlabeled target data (Dt)	Supervised dataset (DL), unsupervised dataset (DU), and meta-dataset (DM)	Labeled source data (Ds), unlabeled target data (Dt), and meta-dataset (Dm)
Core Components	Metric learning, distribution alignment, pseudo-labeling	FixMatch framework, meta-weight network for loss weighting	Metric learning, distribution alignment, pseudo-labeling, meta-weight network
Pseudo-label Generation	Nearest neighbor prediction on target data	FixMatch pseudo-labels from model predictions with confidence threshold	Nearest neighbor prediction on target data
Loss Functions	Metric loss (LTML) + distribution alignment loss (LDDA)	Weighted sum of supervised loss (LL) and unsupervised loss (LU)	Metric loss (LTML) + distribution alignment loss (LDDA), weighted via meta-learned network
Weighting Mechanism	No explicit weight adjustment per sample	Meta-learned weights for both supervised and unsupervised losses	Meta-learned weights for each pair in metric learning
Parameter Updates	Gradient descent on metric network parameters	Alternating updates of model parameters and weight network parameters	Alternating updates of metric network and meta-learned weight network
Learning Rate Adjustment	Gradual decay ($\eta \leftarrow 0.95\eta$)	Fixed learning rate per schedule	Gradual decay ($\eta \leftarrow 0.95\eta$)
Early Stopping Criteria	Stop when change in total loss $< \epsilon$	Not explicitly defined	Stop when change in total loss $< \epsilon$
Unique Advantage	Directly aligns distributions while preserving class structure in metric space	Combines strong semi-supervised learning (FixMatch) with meta-weighting for better label noise handling	Integrates metric learning and distribution alignment with adaptive meta-learning-based sample weighting

2.4 Computational Complexity Analysis

To further clarify the technical characteristics of each method, a brief complexity analysis is provided. For TML-DA, the main computational cost arises from constructing within-class and between-class pairs and calculating similarity distances. For a dataset with n samples and embedding dimension d , the complexity is approximately $O(n^2d)$ due to pairwise distance computations. The gradient updates scale linearly with

the number of layers M resulting in an overall complexity of $O(n^2d + Md^2)$. In the case of MW-FixMatch, the method involves supervised and unsupervised loss calculations along with updates to the meta-weight network. For each batch of size b , supervised and unsupervised losses have complexity $O(bd)$.

The meta-update step adds an extra cost proportional to the meta-dataset size m , giving an overall complexity of

approximately $O(bd + md)$ per iteration. Although MW-FixMatch introduces additional overhead from meta-learning, it avoids the quadratic pairwise operations present in TML-DA. Finally, MetaTMLDA, as a hybrid method, combines the pairwise operations of TML-DA with the meta-weighting of MW-FixMatch, leading to a complexity of about $O(n^2d + md)$. While this introduces higher overhead compared to either method individually, the integrated approach remains computationally feasible for medium-scale datasets due to the use of early stopping and learning rate scheduling mechanisms.

2.5 Dataset

Four image datasets are selected to evaluate the performance of classifiers as shown in Table 2. The study utilises four distinct datasets with varying degrees of class imbalance. The Ganoderma Disease Detection dataset comprises 82 samples, with the majority class accounting for 52 samples (63.41%) and the minority class comprising 30 samples (36.59%). The Palm Oil Leaf Disease dataset contains 50 samples, exhibiting a significant imbalance where the majority class consists of 44 samples (88%) and the minority class comprises only six samples (12%). Two larger datasets, Leaf Nutrient Detection (Boron) and Leaf Nutrient Detection (Magnesium), each contain 6,000 samples. In the Boron dataset, the majority class comprises 4,679 samples (77.98%), and the minority class consists of 1,321 samples (22.02%). Similarly, the Magnesium dataset

has 4,728 samples (78.8%) in the majority class and 1,272 samples (21.2%) in the minority class. All datasets are publicly available at the Roboflow[25].

Table 2. Dataset

Dataset	Total Samples	Number of Samples in The Majority Class	Number of Samples in The Minority Class
Ganoderma Disease Detection	82	52 (63.41%)	30 (36.59%)
Palm Oil Leaf Disease	50	44(88%)	6 (12%)
Leaf Nutrient Detection (Boron)	6000	4679 (77.98%)	1321 (22.02%)
Leaf Nutrient Detection (Magnesium)	6000	4728 (78.8%)	1272 (21.2%)

For each dataset, the samples were divided into training and testing subsets to ensure fair and consistent evaluation across different classifiers. A stratified sampling approach was employed to maintain the original class distribution in both the training and testing sets, thereby preventing bias towards the majority class. In this study, 80% of the data was used for training and 20% for testing for all four datasets. This ratio was selected to provide sufficient data for model training while retaining a representative portion for unbiased performance evaluation. Table 3 summarises the distribution of samples for training and testing across both majority and minority classes[25].

Table 3. The Distribution of Samples for Training and Testing Across Both Majority and Minority Class

Dataset	Total Samples	Train Samples (80%)	Test Samples (20%)	Train Majority	Train Minority	Test Majority	Test Minority
Ganoderma Disease Detection	82	66	16	42	24	10	6
Palm Oil Leaf Disease	50	40	10	35	5	9	1
Leaf Nutrient Detection (Boron)	6000	4800	1200	3743	1057	936	264
Leaf Nutrient Detection (Magnesium)	6000	4800	1200	3782	1018	946	254

2.6. Classifier Performance

Classifier Performance will be measured using Accuracy, Precision, Recall, and Cohen's Kappa. This classifier performance measurement is carried out based on the confusion matrix, which can be seen in Table 4[26].

Table 4. Confusion Matrix

		Predictive Positive Class	Predictive Negative Class
Actual Class	Positive	True Positive (TP)	False Negative (FN)
Actual Class	Negative	False Positive (FP)	True Negative (TN)

The Accuracy, Precision, Recall, Cohen's Kappa, and Macro F1-Score calculations can be seen in Formulas 8 to 12[27].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

$$Kappa = \frac{P_o - P_c}{1 - P_c} \quad (11)$$

$$Macro\ F1 - Score = 2 \times \left(\frac{M_{Precision} \times M_{Recall}}{M_{Precision} + M_{Recall}} \right) \quad (12)$$

The study employs several metrics to comprehensively evaluate model performance: Accuracy (Eq. 8) for overall correctness, Precision (Formula 9) for reliability of optimistic predictions, Recall (Formula 10) for the ability to detect positives, Cohen's Kappa (Formula 11) for agreement adjusted by chance, and Macro F1-Score (Formula 12) as the harmonic mean of macro-averaged precision and recall for balanced class evaluation.

3. Results and Discussions

In this section, the outcomes of the experiments are reported and analyzed to highlight the effectiveness of the proposed MetaTMLDA framework in comparison with existing methods

3.1 Results

This study compares the proposed method not only with state-of-the-art approaches such as TML-DA and MW-FixMatch, but also with three widely used techniques

for handling imbalanced data: Random Undersampling (RUS)[28], SMOTE (Synthetic Minority Oversampling Technique)[29], and Biased Support Vector Machine (BSVM)[30]. The purpose of this comparison is to highlight the ability of the proposed method to outperform conventional approaches in Ganoderma

disease detection, especially under imbalanced class distributions. The performance of each method is summarized in Table 3, using evaluation metrics such as Accuracy, Precision, Recall, Cohen's Kappa, and Macro F1-Score. The results can be seen in Tables 5 to 8.

Table 5. Performance for Ganoderma Disease Detection Dataset

	Proposed Method	TML-DA	MW-FixMatch	Random Undersampling	SMOTE	Biased Support Vector Machine
Accuracy	0.976	0.968	0.971	0.939	0.963	0.951
Precision	0.951	0.946	0.901	0.921	0.929	0.943
Recall	0.915	0.891	0.907	0.801	0.798	0.912
Cohen's Kappa	0.912	0.876	0.899	0.898	0.901	0.887
Macro F1-Score	0.933	0.918	0.904	0.857	0.859	0.927

The experimental results on the Ganoderma Disease Detection Dataset (Table 5 and Figure 1) have demonstrated that the Proposed Method outperforms all comparative techniques across most evaluation metrics. It achieves the highest accuracy (0.976), precision (0.951), recall (0.915), Cohen's Kappa (0.912), and macro F1-score (0.933), indicating strong performance in both balanced and imbalanced class scenarios. While MW-FixMatch and Random Undersampling perform competitively in terms of accuracy, they fall behind in

recall, indicating reduced sensitivity to minority classes. SMOTE improves data balance but yields lower precision and recall, reflecting possible overfitting to synthetic samples. The Biased Support Vector Machine demonstrates strong recall (0.912) but lower accuracy compared to the proposed method. Overall, these results confirm that the proposed method provides a more consistent and robust classification performance.

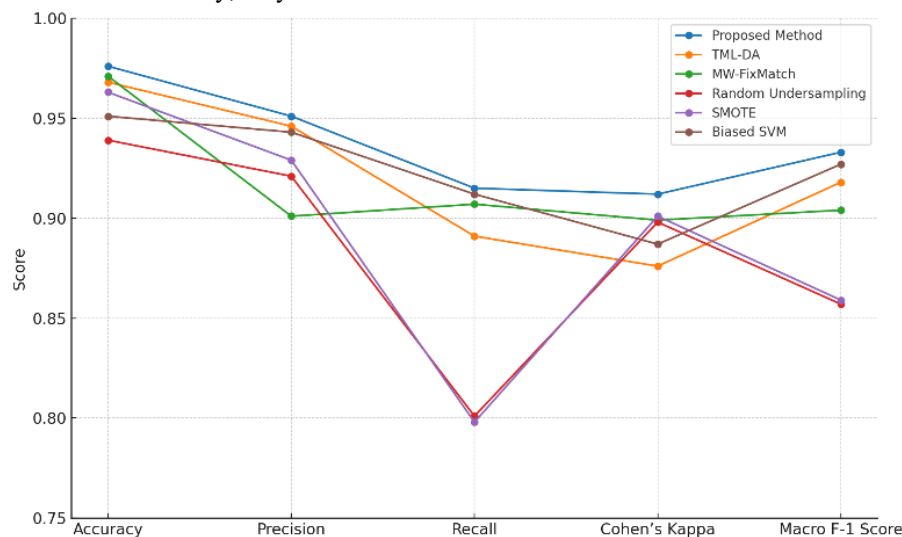


Figure 1. Performance for Ganoderma Disease Detection Dataset

Table 6. Performance for Palm Oil Leaf Disease Dataset

	Proposed Method	TML-DA	MW-FixMatch	Random Undersampling	SMOTE	Biased Support Vector Machine
Accuracy	0.980	0.92	0.93	0.960	0.940	0.960
Precision	0.972	0.969	0.957	0.961	0.941	0.921
Recall	0.957	0.939	0.926	0.945	0.931	0.945
Cohen's Kappa	0.911	0.891	0.901	0.906	0.897	0.895
Macro F1-Score	0.964	0.954	0.941	0.953	0.936	0.933

The results for the Palm Oil Leaf Disease Dataset, as shown in Table 6 and Figure 2, indicate that the Proposed Method outperforms all comparative techniques across most evaluation metrics. It achieves the highest accuracy (0.980), precision (0.972), recall (0.957), Cohen's Kappa (0.911), and Macro F1-score (0.964), demonstrating its superior balance between correctly identifying positive and negative instances.

While TML-DA and MW-FixMatch show competitive precision and recall, their overall performance is slightly lower. Traditional resampling methods, such as Random Undersampling and SMOTE, as well as the Biased Support Vector Machine, deliver reasonable results but fail to match the robustness and consistency of the Proposed Method.

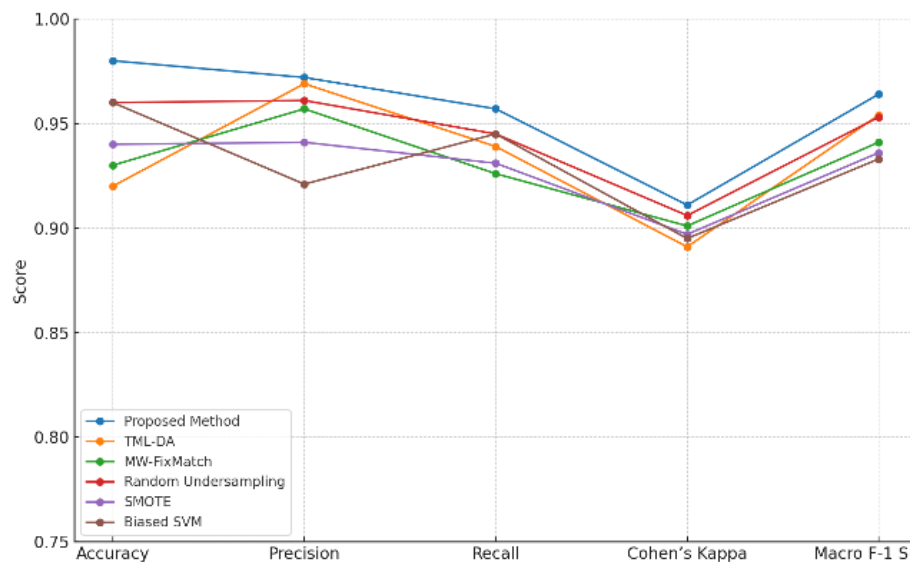


Figure 2. Performance for Palm Oil Leaf Disease Dataset

Table 7. Performance for Leaf Nutrient Detection (Boron) Dataset

	Proposed Method	TML-DA	MW-FixMatch	Random Undersampling	SMOTE	Biased Support Vector Machine
Accuracy	0.995	0.971	0.963	0.967	0.975	0.981
Precision	0.967	0.981	0.965	0.943	0.978	0.965
Recall	0.973	0.964	0.961	0.967	0.941	0.968
Cohen's Kappa	0.921	0.901	0.911	0.931	0.918	0.904
Macro F1-Score	0.969	0.963	0.968	0.955	0.959	0.966

The comparative results for the Leaf Nutrient Detection (Boron) Dataset, as shown in Figure 3 and Table 8, demonstrate that the Proposed Method outperforms all other techniques across most evaluation metrics. It achieves the highest accuracy (0.995), recall (0.973), and macro F1-score (0.969), confirming superior overall classification performance and stability. Although TML-DA records the highest precision (0.981) and Random Undersampling achieves the highest Cohen's Kappa (0.931), these gains are limited. The higher Kappa from Random Undersampling

reflects the metric's sensitivity to error distribution across classes, as undersampling enforces label balance but at the cost of discarding valuable majority-class information, reducing generalisation capacity. Similarly, SMOTE and Biased SVM produce competitive results, but both methods exhibit either reduced accuracy or weaker overall robustness. In contrast, the Proposed Method consistently delivers the strongest and most balanced outcomes, highlighting its effectiveness in preserving both sensitivity and precision while maintaining overall reliability.

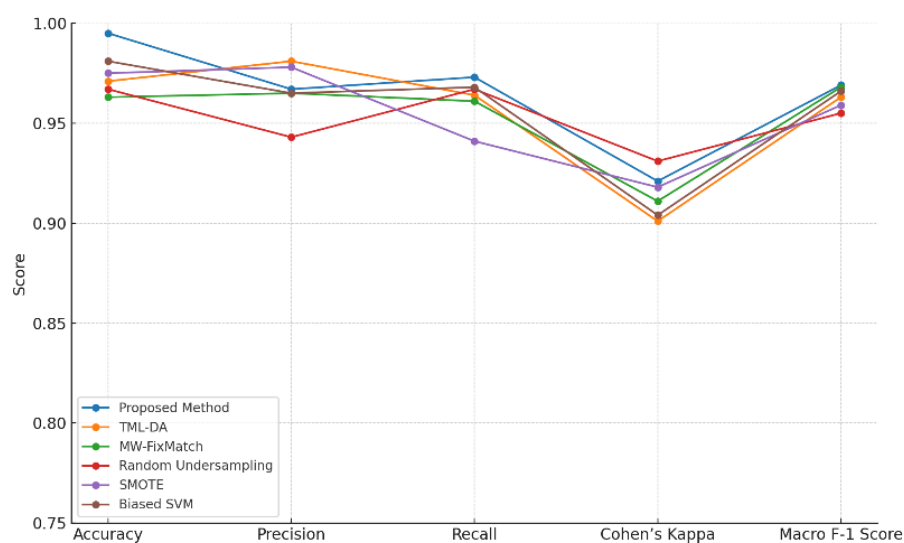


Figure 3. Performance for Leaf Nutrient Detection (Boron) Dataset

Table 8. Performance for Leaf Nutrient Detection (Magnesium) Dataset

	Proposed Method	TML-DA	MW-FixMatch	Random Undersampling	SMOTE	Biased Support Vector Machine
Accuracy	0.995	0.981	0.961	0.968	0.965	0.948
Precision	0.957	0.915	0.921	0.945	0.917	0.934
Recall	0.961	0.908	0.913	0.935	0.921	0.932
Cohen's Kappa	0.919	0.901	0.911	0.901	0.911	0.921
Macro F1-Score	0.959	0.911	0.917	0.94	0.919	0.933

The experimental results for the Leaf Nutrient Detection (Magnesium) Dataset, as presented in Table 8 and Figure 4, demonstrate that the proposed method outperforms all other techniques across most evaluation metrics. It achieves the highest accuracy (0.995), indicating superior overall classification performance, and the highest macro F1-score (0.959), reflecting balanced effectiveness in handling both classes. The precision (0.957) and recall (0.961) values further confirm the model's robustness in minimizing both false positives and false negatives. Compared to advanced methods such as TML-DA (accuracy 0.981, F1-score 0.911) and MW-FixMatch (accuracy 0.961, F1-score 0.917), the proposed method consistently delivers better results, highlighting its ability to improve predictive reliability.

An interesting observation emerges in terms of Cohen's Kappa: the Biased SVM records a slightly higher value (0.921) than the Proposed Method (0.919). This difference stems from the intrinsic nature of Cohen's Kappa, which corrects for chance agreement and places greater weight on how errors are distributed across classes. In this case, the Proposed Method—despite achieving the highest accuracy (0.995), precision (0.957), recall (0.961), and macro F1-score (0.959)—still produces minor misclassifications in the minority class, which reduce the proportional agreement relative to chance. Nevertheless, the Proposed Method remains superior overall, offering stronger and more balanced predictive performance compared to Biased SVM.

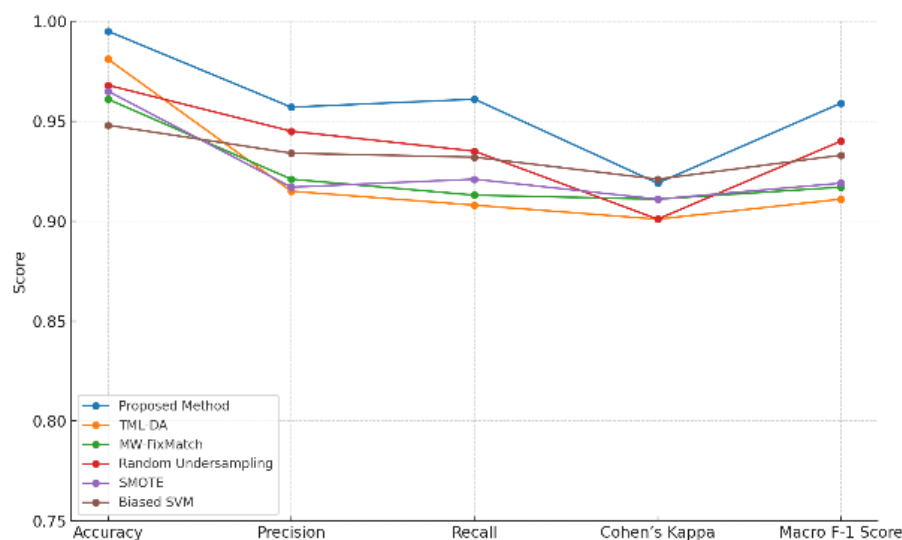


Figure 4. Performance for Leaf Nutrient Detection (Magnesium) Dataset

3.2 Discussions

The experimental results across all four datasets consistently demonstrate that the proposed MetaTMLDA method surpasses both conventional and state-of-the-art techniques in addressing the dual challenges of class imbalance and domain shift. On small-scale datasets such as Ganoderma Disease Detection and Palm Oil Leaf Disease, MetaTMLDA achieves the highest scores in accuracy, precision, recall, and macro F1-score, indicating superior sensitivity to minority class instances without sacrificing precision. This performance suggests that the combination of metric learning-based domain alignment and meta-learned adaptive weighting effectively mitigates the common tendency of models

to bias toward majority classes in severely imbalanced conditions.

On large-scale datasets such as Leaf Nutrient Detection (Boron) and Leaf Nutrient Detection (Magnesium), the proposed method achieves near-perfect accuracy (0.995) and balanced macro F1-scores, demonstrating strong generalisation across both majority and minority classes. Compared to TML-DA, which excels in distribution alignment but lacks adaptive weighting, and MW-FixMatch, which addresses imbalance but not domain shift, MetaTMLDA benefits from the complementary strengths of both approaches. The meta-weight network plays a crucial role by dynamically reweighting losses based on sample difficulty and class representation, thus improving

robustness against pseudo-label noise and enhancing model stability.

Traditional resampling methods, such as Random Undersampling and SMOTE, show limited effectiveness in these experiments. While SMOTE occasionally improves Cohen's Kappa, it risks overfitting to synthetic samples, particularly in small datasets. Random Undersampling, meanwhile, can sometimes produce relatively high Kappa values on Boron, as it enforces a more balanced label distribution. However, this comes at the cost of discarding valuable majority-class data, which reduces overall robustness and generalisation. Similarly, Biased SVM achieves competitive recall and occasionally higher Kappa values, but struggles to maintain high global accuracy, indicating a trade-off between sensitivity and overall predictive strength. In contrast, MetaTMLDA consistently delivers balanced performance across all metrics, highlighting its ability to preserve both sensitivity and specificity without introducing synthetic data artefacts or excessive data reduction.

It is also important to note that on the Boron and Magnesium datasets, the Cohen's Kappa values of the proposed method are slightly lower than those of Random Undersampling and Biased SVM in certain cases. This can be explained by the intrinsic nature of Cohen's Kappa, which corrects for chance agreement and is highly sensitive to how errors are distributed across classes. Even with near-perfect accuracy and strong macro F1-scores, the proposed method may still produce small but systematic misclassifications concentrated in the minority class. These misclassifications reduce the proportional agreement relative to chance, slightly lowering the Kappa value. In contrast, Random Undersampling and Biased SVM, being explicitly oriented toward balancing class contributions, sometimes achieve a more proportional distribution of correct classifications, resulting in marginally higher Kappa values—despite inferior accuracy and macro F1 performance. This finding reinforces that a model with higher accuracy and F1-score does not always guarantee the highest Cohen's Kappa, especially when class imbalance persists.

Overall, the results validate that integrating transfer metric learning with meta-weighted semi-supervised learning provides a robust and scalable solution for precision agriculture applications. The ability to address both imbalance and domain distribution mismatch is particularly beneficial for real-world oil palm monitoring, where data heterogeneity and scarcity of minority classes are common. This approach not only improves predictive accuracy but also enhances early detection capabilities, potentially leading to more sustainable plantation management practices.

4. Conclusions

This study introduces MetaTMLDA, a novel hybrid framework that integrates Transfer Metric Learning-Based Distribution Alignment (TML-DA) with the

meta-weighted semi-supervised learning approach, specifically designed to address the dual challenges of class imbalance and domain shift in oil palm disease and micronutrient deficiency detection. By combining domain-invariant feature learning with adaptive instance reweighting, the proposed method effectively enhances the sensitivity to minority classes without sacrificing overall predictive accuracy.

Experimental evaluations on four diverse datasets—Ganoderma Disease Detection, Palm Oil Leaf Disease, Leaf Nutrient Detection (Boron), and Leaf Nutrient Detection (Magnesium)—demonstrated that MetaTMLDA consistently outperforms both conventional resampling methods and state-of-the-art imbalance-oriented algorithms across multiple performance metrics, including Accuracy, Precision, Recall, Cohen's Kappa, and Macro F1-Score. The results confirm that integrating TML-DA's distribution alignment with MW-FixMatch's adaptive weighting mechanism yields a robust and scalable solution capable of maintaining high classification performance even under severe imbalance and heterogeneous domain conditions.

Beyond its empirical superiority, the proposed framework holds strong practical relevance for precision agriculture, enabling earlier and more reliable detection of critical oil palm health issues. This capability can support timely intervention strategies, reduce potential yield losses, and promote more sustainable plantation management practices. Future work may focus on extending the approach to multi-class and multi-label scenarios, incorporating additional domain adaptation techniques, and exploring its applicability to other agricultural and environmental monitoring tasks.

Acknowledgements

This research was funded by the Ministry of Higher Education, Science, and Technology under Grant Number: 122/C3/DT.05.00/PL/2025 and 7/SPK/LL1/AL.04.03/PL/2025.

References

- [1] H. Hartono and E. Ongko, "Hybrid Approach with Distance Feature for Multi-Class Imbalanced Datasets," *JOIV: International Journal on Informatics Visualization*, vol. 7, no. 1, pp. 131–138, Feb. 2023, doi: 10.30630/joiv.7.1.1292.
- [2] M. Błaszczyk and J. Jędrzejowicz, "Framework for imbalanced data classification," *Procedia Computer Science*, vol. 192, pp. 3477–3486, Jan. 2021, doi: 10.1016/j.procs.2021.09.121.
- [3] K. Dabrock, J. Ulken, N. Pflugrad, J. M. Weinand, and D. Stolten, "Generating a nationwide residential building types dataset using machine learning," *Building and Environment*, vol. 274, p. 112782, Apr. 2025, doi: 10.1016/j.buildenv.2025.112782.
- [4] Hartono and R. B. Y. Syah, "Hybrid Approach with Membership-Density Based Oversampling for handling multi-class imbalance in Internet Traffic Identification with overlapping and noise," *ICT Express*, vol. 10, no. 5, pp. 1094–1102, Oct. 2024, doi: 10.1016/j.ict.2024.04.007.
- [5] H. Ghazouani, W. Barhoumi, E. Chakroun, and A. Chehri, "Dealing with Unbalanced Data in Leaf Disease Detection: A Comparative Study of Hierarchical Classification, Clustering-

- based Undersampling and Reweighting-based Approaches,” *Procedia Computer Science*, vol. 225, pp. 4891–4900, Jan. 2023, doi: 10.1016/j.procs.2023.10.489.
- [6] K. Sai, N. Sood, and I. Saini, “Classification of various nutrient deficiencies in tomato plants through electrophysiological signal decomposition and sample space reduction,” *Plant Physiology and Biochemistry*, vol. 186, pp. 266–278, Sept. 2022, doi: 10.1016/j.plaphy.2022.07.022.
- [7] B. Talukdar, “Handling of Class Imbalance for Plant Disease Classification with Variants of GANs,” in *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)*, Nov. 2020, pp. 466–471. doi: 10.1109/ICIIS51140.2020.9342728.
- [8] A. A. Khan, O. Chaudhari, and R. Chandra, “A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation,” *Expert Systems with Applications*, vol. 244, p. 122778, June 2024, doi: 10.1016/j.eswa.2023.122778.
- [9] M. Vasconcelos and L. Cavique, “Mitigating false negatives in imbalanced datasets: An ensemble approach,” *Expert Systems with Applications*, vol. 262, p. 125674, Mar. 2025, doi: 10.1016/j.eswa.2024.125674.
- [10] Y. W. Khoo and K. P. Chong, “Ganoderma boninense: general characteristics of pathogenicity and methods of control,” *Front Plant Sci*, vol. 14, p. 1156869, July 2023, doi: 10.3389/fpls.2023.1156869.
- [11] X. Song et al., “Effect of boron deficiency on the photosynthetic performance of sugar beet cultivars with contrasting boron efficiencies,” *Front. Plant Sci.*, vol. 13, Jan. 2023, doi: 10.3389/fpls.2022.1101171.
- [12] M. Ariza-Sentis, S. Vélaz, R. Martínez-Peña, H. Baja, and J. Valente, “Object detection and tracking in Precision Farming: a systematic review,” *Computers and Electronics in Agriculture*, vol. 219, p. 108757, Apr. 2024, doi: 10.1016/j.compag.2024.108757.
- [13] P. Goel and A. Ganatra, “Unsupervised Domain Adaptation for Image Classification and Object Detection Using Guided Transfer Learning Approach and JS Divergence,” *Sensors*, vol. 23, no. 9, Art. no. 9, Jan. 2023, doi: 10.3390/s23094436.
- [14] D. Cacciarelli and M. Kulahci, “Active learning for data streams: a survey,” *Mach Learn*, vol. 113, no. 1, pp. 185–239, Jan. 2024, doi: 10.1007/s10994-023-06454-2.
- [15] X. Wang et al., “Dual adversarial network with meta-learning for domain-generalized few-shot text classification,” *Applied Soft Computing*, vol. 146, p. 110697, Oct. 2023, doi: 10.1016/j.asoc.2023.110697.
- [16] N. Moniz and V. Cerqueira, “Automated imbalanced classification via meta-learning,” *Expert Systems with Applications*, vol. 178, p. 115011, Sept. 2021, doi: 10.1016/j.eswa.2021.115011.
- [17] H. Chamlal, H. Kamel, and T. Ouaderhman, “A hybrid multi-criteria meta-learner based classifier for imbalanced data,” *Knowledge-Based Systems*, vol. 285, p. 111367, Feb. 2024, doi: 10.1016/j.knsys.2024.111367.
- [18] Y. Guo et al., “Meta-ensemble learning with adaptive sampling for imbalanced medical Raman spectroscopy data,” *Applied Soft Computing*, vol. 176, p. 113142, May 2025, doi: 10.1016/j.asoc.2025.113142.
- [19] M. Jindal and B. Singh, “Class imbalance-aware domain specific transfer learning approach for medical image classification: Application on COVID-19 detection,” *Engineering Applications of Artificial Intelligence*, vol. 150, p. 110583, June 2025, doi: 10.1016/j.engappai.2025.110583.
- [20] M. Isnain, A. A. Hidayat, and B. Pardamean, “Indonesian Agricultural-crops Classification Using Transfer Learning Model,” *Procedia Computer Science*, vol. 227, pp. 128–136, Jan. 2023, doi: 10.1016/j.procs.2023.10.510.
- [21] M. Yan, N. Jiang, and N. Li, “TML-DA: Transfer metric learning based distribution alignment framework for domain class imbalanced classification,” *Knowledge-Based Systems*, vol. 311, p. 113086, Feb. 2025, doi: 10.1016/j.knsys.2025.113086.
- [22] J. Wang and X.-L. Zhang, “Improving pseudo labels with intra-class similarity for unsupervised domain adaptation,” *Pattern Recognition*, vol. 138, p. 109379, June 2023, doi: 10.1016/j.patcog.2023.109379.
- [23] Y. Li et al., “Imbalanced complemented subspace representation with adaptive weight learning,” *Expert Systems with Applications*, vol. 249, p. 123555, Sept. 2024, doi: 10.1016/j.eswa.2024.123555.
- [24] X. Zheng, W. Hong, D. Chen, A. Xue, and Y. Kong, “MW-FixMatch: A class imbalance semi-supervised learning algorithm based on re-weighting,” *Neurocomputing*, vol. 626, p. 129385, Apr. 2025, doi: 10.1016/j.neucom.2025.129385.
- [25] R. Roboflow, “Palm Leaves Disease Detection,” Roboflow. Accessed: June 14, 2025. [Online]. Available: <https://universe.roboflow.com/firstproject-ofadf/palm-leaves-disease-detection>
- [26] J. Guo, H. Wu, X. Chen, and W. Lin, “Adaptive SV-Borderline SMOTE-SVM algorithm for imbalanced data classification,” *Applied Soft Computing*, vol. 150, p. 110986, Jan. 2024, doi: 10.1016/j.asoc.2023.110986.
- [27] A. E. Yilmaz and H. Demirhan, “Weighted kappa measures for ordinal multi-class classification performance,” *Applied Soft Computing*, vol. 134, p. 110020, Feb. 2023, doi: 10.1016/j.asoc.2023.110020.
- [28] B. Liu and G. Tsoumakas, “Dealing with class imbalance in classifier chains via random undersampling,” *Knowledge-Based Systems*, vol. 192, p. 105292, Mar. 2020, doi: 10.1016/j.knsys.2019.105292.
- [29] B. S. Raghuwanshi and S. Shukla, “SMOTE based class-specific extreme learning machine for imbalanced learning,” *Knowledge-Based Systems*, vol. 187, p. 104814, Jan. 2020, doi: 10.1016/j.knsys.2019.06.022.
- [30] J.-J. Zhang and P. Zhong, “Learning Biased SVM with Weighted Within-Class Scatter for Imbalanced Classification,” *Neural Process Lett*, vol. 51, no. 1, pp. 797–817, Feb. 2020, doi: 10.1007/s11063-019-10096-8.