

植物病害診断システムにおけるDiffusion-Basedデータ拡張と補完情報の効果

西井, 彩葉 / NISHII, SAYO

(出版者 / Publisher)

法政大学大学院理工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学研究科編

(巻 / Volume)

65

(開始ページ / Start Page)

1

(終了ページ / End Page)

7

(発行年 / Year)

2024-03-24

(URL)

<https://doi.org/10.15002/00030752>

植物病害診断システムにおける Diffusion-Based データ拡張と補完情報の効果

EFFECTS OF DIFFUSION-BASED DATA EXTENSION AND COMPLEMENTARY INFORMATION ON
PLANT DISEASE DIAGNOSIS SYSTEMS

西井 彩葉
SAYO NISHII
指導教員 彌富仁

法政大学大学院理工学研究科応用情報工学専攻修士課程

Plant diseases are a serious problem in agriculture, with the FAO reporting that up to 40% of the world's annual crop production is lost to pests. This causes losses to the global economy in excess of \$220 billion annually, necessitating rapid detection and accurate treatment of plant diseases. However, current inspection methods are mainly visual and genetic, which are costly in terms of manpower, money, and time. To address this issue, research has developed inexpensive and convenient automatic plant disease diagnosis systems. 2015 saw the development of an automatic diagnosis system using deep learning, followed by the release of plant disease datasets and the proposal of improved versions, which have achieved high discrimination performance. However, these automatic diagnosis systems have the problem that the leaf images included in the training images are pre-cropped with the same background, which significantly reduces the diagnosis performance for images taken in actual fields or agricultural fields. As a means of addressing this problem, it is essential to ensure the diversity of training images and improve upon realistic datasets. Many data extensions using GAN-based image generation models have been proposed in the past, but these methods also suffer from a lack of diversity due to a biased distribution of training images of healthy leaves, as well as from the enormous amount of time and effort required. On the other hand, research on Diffusion Models (DMs) has been progressing, and it is now possible to generate images with rich diversity using large-scale pre-trained models. In this study, we applied DreamBooth to Stable diffusion to generate disease-specific images using a small number of plant disease images, and evaluated the effectiveness of Diffusion-based data expansion using the generated images. The results showed that the +Prompt(v2)-augmentation method achieved the best discrimination performance for cucumber leaves in HE, MYSV, MD, Macro avg, and Micro avg, and for eggplant leaves in HE, PM, and BW. This means that we can use the information complemented by prompt to identify the leaves. We confirmed that domain-adaptation is also feasible by considering complementary information from prompts. Future work is expected to improve the quality of the generated images by addressing problems in converting latent representations to images in the text-to-image generation model, adjusting the loss function of DreamBooth, and improving the attention mechanism of U-Net. Translated with DeepL.com (free version)

Keywords : *plant disease diagnosis, Stable diffusion, DreamBooth*

1. はじめに

植物病害は農業における深刻な問題であり、国際連合食糧農業機関（FAO）の報告 [1] によると、年間世界の作物生産量の最大 40% が害虫によって失われていると推定されている。そのため、植物病害の迅速な検出とそれに続く的確な処置が必要不可欠である。しかしながら、現行の検査手法は主として農家や専門家による目視検査や、必要に応じて遺伝子検査が主流となっており、これらには人的・金銭的・時間的コストが伴う。

この課題に対処すべく、安価で手軽な植物病害自動診断システムの開発が要請され、それに向けた研究が進展している。2015 年に、深層学習技術を活用したキュー

リ葉病害画像の自動診断システムが開発され、[2] その後、深層学習技術の急速な進展と、植物病害データセット「PlantVillage dataset(PVD)[3]」の公開により、多くの深層学習ベースの自動診断システムが提案され、高い識別性能を達成している [4][5]。しかし、これらの自動診断システムは、学習画像に含まれる葉画像があらかじめ同一の背景で刈り取られたものであり、実際の畑や農地で撮影された画像に対する診断性能が顕著に低下するという課題が存在している [6][7]。これらに対処するため、実際の農業現場で撮影された画像を使用した植物病害データセットである PVD の改良版「PlantDoc[8]」が公開され、これにより実用性を考慮した条件下での研究が増加している [9][10]。しかし、これらの研究の多くは、広範な圃場で撮影された画像を混ぜて学習画像とテスト画像に

分割し、背景、画角、日照環境などが学習画像と極めて類似した画像で評価している。2021年に実施された研究[11]では、異なる農地から収穫されたキュウリ、トマト、ナス、イチゴの画像を用いて大規模な解析が行われた。この研究では、同じ手法で構築された学習画像と同じ農地データを含むテスト画像に対して非常に高い診断性能が示されたが、異なる農地から厳密に未知のデータを評価した結果、診断性能が低下することが報告された。この研究成果より、過去において数万件以上の学習を行ったモデルであっても、このような学習画像に対する過学習と、病気に対する汎化性の獲得が難しいことを明示的に示している。そのため、葉領域が本来の対象であるべきなのに対し、背景領域の情報を基に診断している可能性が浮上した[12]。この問題に対処する手段として、学習画像の多様性を確保し、現実的なデータセットへの改良が不可欠である。そこで、過去には多くのGAN (Generative Adversarial Network) [13] ベースの画像生成モデルを使用したデータ拡張が提案されている [14][15]。これらの手法は一定の成果を収めているが、異なる圃場で撮影された画像を混ぜて学習画像とテスト画像に分割し、背景や画角、日照環境などが学習画像と極めて類似した状態で評価が行われていたため、実際の圃場環境を考慮したデータ拡張が実施されていない。また、CycleGAN[16]を使用した研究では、実際の圃場環境を考慮し、CycleGANのスタイル変換機能を使用し、健全な葉の画像から病害葉の画像を生成している [17][18]。しかし、CycleGANは画像全体を変換するため、病害の特徴だけでなく、生成される画像の背景も変化させてしまうという課題が存在するため、適切な画像が生成されない可能性がある。また、生成された画像に対応する分布が制約されるため、データ拡張の効果も限られている。一方で、CycleGANを改良した LeafGAN[19]では、注目する葉領域にのみ病徴を付与する独自の学習機構を備えており、自然な病害画像を生成することができ、データ拡張として効果的に活用できた。結果として平均7.4%の性能向上が得られたが、CycleGANは1対1の変換を行うモデルであるため、生成可能な画像の枚数は健全な葉の枚数に制限される。そこで、無数の健全画像を生成し、それらに病徴を付与することで、さらに多様な画像を生成を行う productive and pathogenic image generation (PPIG)[20]が提案された。生成された画像を学習する識別器は背景などへの過学習を低減し、平均識別率を71.5%から81.2%まで向上させることに成功したが、これらの手法には、生成モデルの学習に使用された健全な葉の学習画像の分布に偏りが見られ、それにより他のバリエーションを効果的に網羅できていないという課題が存在する。また、PPIGの健全な葉画像および病害葉画像の生成過程では、最大43,137枚の画像生成と高コストな style transfer の実装が行われており、これには画像枚数の多さによる手間や、モデルの構築に要する時間が膨大であるといった課題が伴う。

一方で、近年は Diffusion Model (DM) [21] の研究が進み、大規模な事前学習モデルを使用した高品質な画像

Table 1 データセットの詳細 [画像枚数 (撮影圃場数)]

Class	Cucumber		Eggplant	
	Train	Test	Train	Test
Healthy(HE)	11,895(6)	5,576(2)	12,431(5)	1,122(1)
Cucurbit chlorotic yellow virus (CCYV)	3,775(1)	179(1)		
Melon yellow spot virus (MYSV)	16,098(4)	1,004(3)		
Mosaic diseases(MD)	23,866(3)	1,626(2)		
Corynespora leaf spot (CLS)	6,141(4)	1,813(3)		
Powdery mildew (PM)	5,642(4)	1,898(3)	7,936(5)	938(1)
Gray mold(GM)			1,024(1)	166(1)
Leaf mold(LM)			3,188(3)	732(2)
Verticillium wilt(VW)			3,176(3)	354(2)
Bacterial wilt(BW)			3,415(3)	462(1)
Total	67,417	12,093	31,170	3,774

生成モデルが提案されている。これらのモデルは多くのタスクに適応可能であり、GANよりも多様な画像生成が可能である。特に、text-to-image 技術をテキスト (= prompt) と併用した場合、より多くの多様な画像を生成できる可能性がある。

Stable Diffusion は論文の Latent Diffusion Model(LDM)[22] をベースとしたノイズから高解像度な画像を生成する text-to-image モデルであり、DMの両端に VAE[23] を配置し、潜在空間での生成処理を組み合わせることで、学習の効率とデータ生成速度を大幅に向上させた。さらに、テキストエンコーダと cross-attention 機構を通じて、テキストなどの情報を潜在空間で画像と組み合わせることで、text-to-image の画像生成を可能にしている。

そこで、本研究では Stable Diffusion に DreamBooth[?] を適用し、学習用画像データが不足している植物病に対して、その植物病の少量の学習画像を用いて病徴に特化した画像を生成を行い、Diffusion-based のデータ拡張を行う。Stable Diffusion は事前学習に用いたデータに基づきテキストから画像を生成可能だが、専門的なタスクにはそのままでは適用できない。そこで、DreamBoothを活用して Stable Diffusion に新しい植物病害の概念を追加する。この手法により、植物病に特有の症状や環境に関連する概念を柔軟に追加でき、生成される画像の多様性を向上させることが期待できる。これによって、データ拡張としての有効性を高め、識別モデルの性能向上に寄与することが本研究の目的である。

2. 実験

本実験で使用したデータセットは、各植物について専門家が病害を接種または虫を介して感染させ、共感染を避けるために隔離された施設で厳密に管理された作物を撮影した。このデータセットには、典型的な症状だけでなく、感染初期の症状画像も含まれている。性能評価のため、評価用画像には学習画像と別の圃場で撮影された画像である。各作物の病気とクラス数の内訳は以下の通りである。各データセットの内訳を Table 1 に示す。きゅうり葉データセットは5種類の植物病害 (退緑黄化病 (CCYV), 黄化えそ病 (MYSV), モザイク病 (MD), すずかび病 (CLS), うどんこ病 (PM)) と健全 (HE) の計6クラスを使用。なす葉データセットは5種類の植物病害 (うどんこ病 (PM), 灰色かび病 (GM), すずかび病 (LM), 半

身萎凋病 (VW), 青枯病 (BW)) と健全 (HE) の計 6 クラスを使用した。

3. 病害識別器の構築

各種生成画像の追加によるデータ拡張の効果を検証するため、以下の (1)~(4) の条件で各データセットを作成し、ImageNet21k[25] データセットで学習済みの EfficientNetV2[26] で学習させた。画像生成モデルは、LAION-5B で事前学習済みの stable diffusion2. 1 モデルを使用し、DreamBooth で追加学習を行った。

病害識別器のデータ拡張には、512 × 512 ピクセルにリサイズし、画像の中心を切り抜き、上下左右反転、-180° から 180° までの回転を行った。最後に ColorJitter で brightness = 0.1, contrast = 0.1, saturation = 0.1 と設定。epoch 数は 100 に設定し、生成された画像の効果を評価した。

a) DreamBooth 設定

instance prompt(画像生成時に追加学習させたデータと呼び出す prompt) には "`<*> leaf`" に設定。instance prompt の学習用画像には、各病害ごとの学習画像データセットの中から 30 枚選択した。"`<*> leaf`" の "`<*>`" の部分に、instance prompt 学習時に使用した 30 枚を考慮した概念が学習される。"`<*>`" が、"leaf" というクラスと対になる固有の識別子に相当する。また、各データ拡張手法で使用する 30 枚は全て同じものとする。* の中身にはきゅうり葉は HE, CCYV, MYSV, MD, CLS, PM の順に 00,..., 05 と命名し、なす葉には HE, PM, GM, LM, VW, BW の順に 00,..., 05 と命名した。

+Locally-augmentation で使用する HE の圃場の instance prompt には "`<*> leaf`" に設定。* の中身にはきゅうり葉は埼玉, 福島, 愛知, 宮崎, 高知, 農研(農研機構)の順に 00_Saitama,..., 05_Noken と命名し、なす葉には富山, 山梨, 三重, 京都, 高知, BW の順に 00_Toyama,..., 04_Kochi と命名した。

新たに学習したい単語の大まかなクラス名の入力として class prompt には "leaf" と設定。class prompt の学習用画像(正規化画像)には、各病害ごとの学習画像データセットを使用した。"`<*> leaf`" の "leaf" の部分に、class prompt 学習時に使用した正規化画像を考慮した概念が学習される。

トレーニングのバッチサイズは 1.0, gradient accumulation steps(バックワード/更新パスを実行する前に、勾配を累積する更新ステップの数)は 2.0 に設定。prior preservation loss(正規化画像の loss の重み)(PPL)の損失関数を有効にして、重みを 1.0 に設定した。train text encoder を有効にしてテキストエンコーダを fine tuning した。学習率は 5e-6 に設定、そして scale lr を導入し、lr scheduler(Learning Rate を調整するための Scheduler)には "constant" を採用した。lr warmup steps は 0 に指定。解像度は 768 × 768 ピクセルに設定し、max train steps(学習させるステップ数)800 に設定した。

b) Stable diffusion 設定

DreamBooth で学習した instance prompt を Stable diffusion の prompt に入力することで画像の生成を行った。num inference steps を 80 回に、guidance scale は 10 に設定した。

c) +Normal-augmentation

DreamBooth に入力する学習用画像は各病害ごとに学習用画像から 30 枚選択した。そして学習した各植物病害概念を prompt に入力し、各病害ごとに 3,000 枚ずつ生成した。Baseline のデータセットに生成画像を追加し、きゅうり葉データセットを計 85,417 枚、なす葉データセットを 49,170 枚にした。

d) +Locally-augmentation

+Locally-augmentation では健全葉が持つ圃場に対しそれぞれの病害を掛け合わせて画像生成を行う。

学習データセットに含まれる圃場数と種類についてまとめた表を Table f) に示す。+Locally-augmentation では、その病害が学習データセットに持たない圃場を HE の圃場と混合させて画像生成を行う。例えば、きゅうり葉 HE の学習データセットは埼玉, 福島, 愛知, 宮崎, 高知, 農研の計 6 つの圃場から集計されている。きゅうり葉 PM の場合、HE の圃場が持つ高知, 農研の学習データセットを持ち合わせていない。そのため、PM には

1. HE の高知と PM の概念を学習
 - "`<04_Kochi > leaf and <05 > leaf`"
2. HE の農研と PM の概念を学習
 - "`<05_Noken > leaf and <05 > leaf`"

の計 2 フェーズの生成が必要となる。Fig. ?? に記述されている Filed num の () の中身の数字は、各病害に必要なフェーズ数を表している。1 フェーズに 1,000 枚生成を行うため、きゅうり葉 PM の場合だと、2 フェーズ × 1,000 枚 = 計 2,000 枚生成、きゅうり葉全体だと、16 フェーズ × 1,000 枚 = 計 16,000 枚生成する。

DreamBooth に入力する学習用画像は HE の各圃場画像 30 枚と植物病害の画像 30 枚を DreamBooth の concept_list を用いて同時に学習を行い、各植物病害概念を 2 つ学習した。そして各植物病害概念 2 つを prompt に入力し、1 フェーズ 1,000 枚生成した。Baseline のデータセットに生成画像を追加し、きゅうり葉データセットを計 83,417 枚、なす葉データセットを 45,170 枚にした。

e) +Prompt-augmentation

DreamBooth に入力する学習用画像は各病害ごとに学習用画像から 30 枚選択した。そして学習した各植物病害概念と 3 つの文脈を組み合わせて入力した。植物病害概念と組み合わせる prompt を 1. "`<*>leaf, sunlight`", 2. "`<*>leaf in a flowerpot`", 3. "`<*> leaf on the soil`" の 3 つの文脈を prompt に入力し、各文脈ごとに 1,000 枚生成した。各病害ごとに学習し生成した計 3,000 枚を Baseline のデータセットに追加し、きゅ

うり葉データセットを計 85,417 枚, なす葉データセットを 49,170 枚にした。

f) +Prompt(v2)-augmentation

DreamBooth に入力する学習用画像は各病害ごとに学習用画像から 30 枚選択した。そして学習した各植物病害概念 prompt と病害の評価画像の背景や症状に似た条件に合わせた文脈を組み合わせて入力し、各文脈ごとに 1,000 枚生成した。各病害ごとに学習し生成した計 3,000 枚を Baseline のデータセットに追加し、きゅうり葉データセットを計 85,417 枚, なす葉データセットを 49,170 枚にした。

4. 結果・考察

5. 病害葉画像生成と効果

(1) +Normal-augmentation

+Normal-augmentation で生成された葉は、宙に浮いてるような葉が複数生成される画像が確認できた。これは、DreamBooth に用いられる損失関数 Reconstruction Loss (RL) の最小化による影響で、DreamBooth 学習時に用いた 30 枚の画像に大きく依存した画像が生成されたためであると考えられる。今回使用した 30 枚は、葉が一部見切れているものや、背景に複数葉が映り込んでいるものが多かったため、今回のような画像が生成されたと推測する。また、葉の色が変色して黒、赤や紫色に生成されるものが病害全体で見られた。これらの原因として、Stable diffusion の VAE decoder の潜在表現から画像に変換する際に色味の調整が出来ず、不具合が生じたためであると推察する。PM の生成画像が、葉ではなく、葉の背景に症状が付与されて生成された。これは、U-Net の Self-Attention 機構で注目すべき画像の画像の物体 (葉) 配置が誤っており、葉の病徴が付加されていないような画像が生成されたと推測する。GM は、奇形な生成画像が多く見られた。また、学習画像に用いなかった車の画像が映り込んで生成されていた。これは、LAION-5B による膨大な数で学習された text-to-image 生成モデルにより、prompt と何かしら車の概念が紐づけられて画像が生成されたと推察する。全体的には比較的学習画像に近い自然な植物病害画像が生成され、きゅうり葉の識別性能では CLS が 88.0% と最も高い結果を示し、Baseline と比較して Micro avg が 69.2% まで向上した。また、なす葉では HE への GM の誤分類率が 0.6%、HE への VW の誤分類率が 4.2% 減少しており、改善がみられた。これは、RL の最小化により、選択した病害画像 (使用した 30 枚) に関連付けながら、text-to-image が備える病害の多様な生成を可能にしたからであると考えられる。

(2) +Locally-augmentation

+Locally-augmentation で生成された葉は、病害と HE が混合されずに、生成されたように見られた。また、CCYV の症状が一部のみ生成されたような画像が生成された。これは、U-Net の Cross-Attention 機構でテキ

ストから画像に変換する際に注目すべき箇所が誤っていたり、物体の大きさを把握しておらず、病害 (葉のサイズ小) と HE (葉のサイズ大) がそれぞれ同サイズのまま混合されたためであると考えられる。HE の各圃場と病害を混合して生成された画像は、掛け合わせる HE の圃場によって大きく見た目が異なる画像を生成した。これは、全体的に HE 圃場の持つドメイン情報 (背景や葉など) と病害画像の持つドメイン情報が融合することで、新たな形状をした葉 (新たなドメイン情報) が作成されたためと考える。しかし、HE の最大 5 圃場を学習データセットに追加した CCYV のきゅうり葉識別性能が、Baseline と比べて HE への CCYV の誤分類率が 2.7% 増加した。これは、病徴が付加されていないような生成画像 (HE とほぼ同様) が多く作成されたため、HE への誤認識率が増加したと推定する。また、なす葉では HE の最大 4 圃場を学習データセットに追加した GM が、Baseline と比べ、HE への GM の誤分類率が 1.2% 減少し若干の改善がみられたものの、劇的な識別性能の向上は見られなかった。

(3) +Prompt-augmentation

prompt の入力による生成結果は”<*> leaf, sunlight”と”<*> leaf on the soil”と指定した際の生成画像は、全体的に文脈通りに自然な植物病害画像を生成することが出来た。”<*> leaf in a flowerpot”と指定した際の生成画像は、全体の半分以上が植木鉢に入っていない画像や植木鉢の形状が奇形なもの、一部葉が植木鉢の外に生成された画像が確認できた。これは、CLIP による Text Encoder のテキストと画像間の Contrastive learning が適切に行われておらず、文脈通りの画像が生成されなかったと推定する。また Cross-Attention 機構でテキストから画像に変換する際に注目すべき箇所が誤っているためであると推測する。

きゅうり葉の識別性能は Baseline と比べ、HE への MD の誤分類率が 4.9%、HE への PM の誤分類率が 2.7% 減少し、なす葉では、HE への VW の誤分類率が 3.1% 減少し、改善がみられたが、評価画像とのドメイン差を埋めるような生成画像が +Prompt-augmentation では出来なかったと推定する。

(4) +Prompt(v2)-augmentation

”<*> leaf in the field”と指定した際の生成画像は、本来目的としていた圃場の中にある生成画像ではなく、野原や原野の中にある植物病害画像が生成された。これは、Stable diffusion の CLIP によるテキストと画像の紐付けている意味合いによって生成される画像が変わってくるためであると考えられる。field は圃場という意味を持つが、CLIP では野原や原野として画像に紐付けられて学習されていたため、今回の画像が生成されたと推定する。”<*> leaf with water droplets on it”と指定した際の生成画像は、全体的に水滴が病害場に付与された生成画像が作成できたが、水滴が大粒でイラスト風に生成されており、自然な外観ではない。GM の生成画像では、葉の表面や背景に灰のようなものが見られ、GM の病徴が反映された画像が生成された。また、”<*> leaf with

Table 2 +Locally-aug の使用データセットの説明

Class	Train			
	Cucumber		Eggplant	
	Field num	Field species	Field num	Field species
HE	6	埼玉, 福島, 愛知, 宮崎, 高知, 農研	5	富山, 山梨, 三重, 京都, 高知
CCYV	1(+5)	埼玉		
MYSV	4(+2)	埼玉, 愛知, 宮崎, 農研		
MD	3(+5)	埼玉, 長野, 法政		
CLS	4(+2)	埼玉, 福島, 愛知, 宮崎		
PM	4(+2)	埼玉, 福島, 愛知, 宮崎	5(+0)	富山, 山梨, 三重, 京都, 高知
GM			1(+4)	三重
LM			3(+2)	三重, 京都, 高知
VW			3(+2)	富山, 山梨, 京都
BW			3(+2)	富山, 山梨, 三重

Table 3 きゅうり葉 F1-score の比較

Dataset	F1-score [%]							
	HE	CCYV	MYSV	MD	CLS	PM	Macro avg	Micro avg
Baseline	70.5	68.5	52.1	48.0	87.8	66.1	65.5	65.4
+Normal-augmentation	75.4	67.3	49.6	56.9	88.0	64.8	67.0	69.2
+Locally-augmentation	70.3	64.4	48.5	46.4	87.0	64.9	63.6	64.1
+Prompt-augmentation	72.7	48.1	46.3	56.8	85.9	73.5	63.9	68.1
+Prompt(v2)-augmentation	80.4	62.5	54.8	59.7	87.0	72.8	69.5	74.3

Table 4 なす葉 F1-score の比較

Dataset	F1-score [%]							
	HE	PM	GM	LM	VW	BW	Macro avg	Micro avg
Baseline	80.8	89.9	63.7	85.5	63.3	59.2	73.8	79.8
+Normal-augmentation	79.5	87.4	62.5	82.7	56.8	57.9	71.1	77.3
+Locally-augmentation	78.7	87.1	66.0	84.3	61.6	55.7	72.2	77.8
+Prompt-augmentation	79.9	88.9	60.8	84.9	59.0	50.6	70.7	77.9
+Prompt(v2)-augmentation	81.2	90.0	62.8	85.0	60.0	61.1	73.3	79.7

eggplants”と指定した際の生成画像は、イラストで生成された画像が確認でき、葉の中に収まる程度の小さななすの実が添えて生成された。これは、LAION-5Bによる膨大な数で学習されたデータセットの中から、promptに関連した多様な概念を生成に組み込み、画像が作成されたためであると考えられる。求めていたpromptのsub情報が追加されずに生成されたため、今後は、prompt engineeringによる生成条件の制御を行うことや、PPLにより正規化画像の重みを調整する必要があると考えられる。+Prompt(v2)-augmentationで生成されたなす葉の画像の例をFig. 1に示す。全体的には、植物病害の持つ症状や重要な視覚的特徴への高い忠実度を維持しながら、promptの文脈の環境や条件に沿った生成画像を実現した。きゅうり葉の識別性能はHE, MYSV, MD, Macro avg, Micro avgにて、なす葉の識別性能は、HE, PM, BWにて他のデータ拡張手法の中で一番高い精度を叩き出した。これにより、promptに病害評価画像に似た条件の症状や環境を入力することで、評価画像によ

る domain-adaptation が可能であることが確認できた。

6. おわりに

本研究ではStable diffusionにDreamBoothを適用し、少量の植物病害の学習画像を用いて病徴に特化した画像を生成を行い、4つのDiffusion-basedのデータ拡張を行った。その結果、データ拡張の中でPrompt(v2)-augmentationが最も優れた性能を発揮し、promptによる補完情報を考慮することによるdomain-adaptationが実現できることを確認した。

今後の展望として、text-to-image生成モデルの潜在表現から画像に変換する際の不具合やDreamBoothの損失関数の調整、U-NetのAttention機構に対する改善を進め、生成画像の品質向上が望まれる。

謝辞

本研究はJSPS科研費基盤研究(C)17K08033, (2017-2020)の助成を受けたものである。

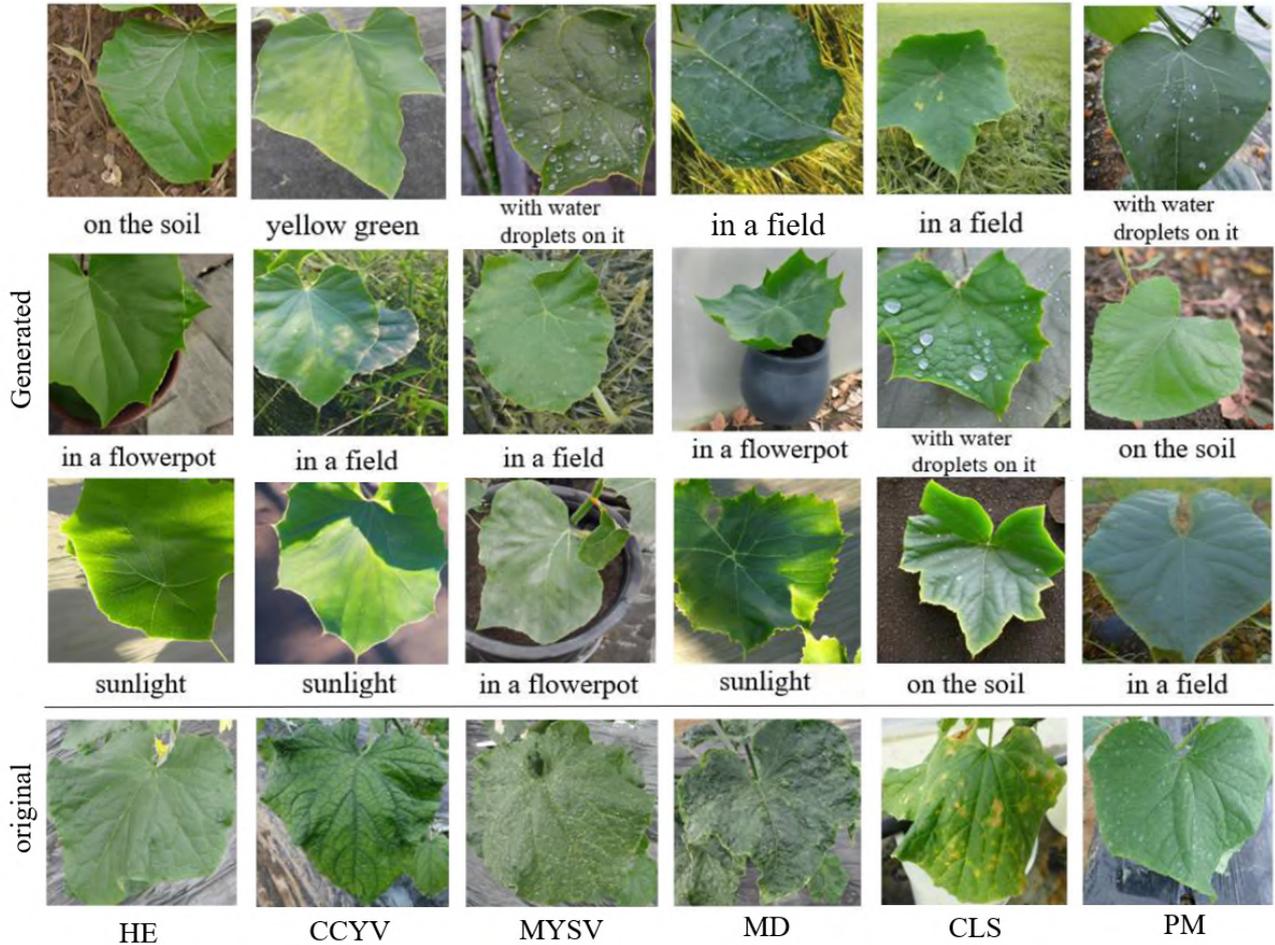


Fig. 1 +Prompt(v2)-augmentation : 本物の画像と生成画像比較 (なす葉)

参考文献

- [1] I. Secretariat, “Scientific review of the impact of climate change on plant pests,” tech. rep., 2021.
- [2] Y. Kawasaki, H. Uga, S. Kagiwada, and H. Iyatomi, “Basic study of automated diagnosis of viral plant diseases using convolutional neural networks,” in *Advances in Visual Computing: 11th International Symposium, ISVC 2015, Las Vegas, NV, USA, December 14-16, 2015, Proceedings, Part II 11*, pp. 638–645, Springer, 2015.
- [3] S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using deep learning for image-based plant disease detection,” *Frontiers in plant science*, vol. 7, p. 1419, 2016.
- [4] L. Kumar and D. K. Singh, “Analyzing computational response and performance of deep convolution neural network for plant disease classification using plant leave dataset,” in *2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT)*, pp. 549–553, IEEE, 2021.
- [5] V. Pandey, U. Tripathi, V. K. Singh, Y. S. Gaur, and D. Gupta, “Survey of accuracy prediction on the plantvillage dataset using different ml techniques,” *EAI Endorsed Transactions on Internet of Things*, vol. 10, 2024.
- [6] S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using deep learning for image-based plant disease detection,” *Frontiers in plant science*, vol. 7, p. 1419, 2016.
- [7] K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” *Computers and electronics in agriculture*, vol. 145, pp. 311–318, 2018.
- [8] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, “Plantdoc: A dataset for visual plant disease detection,” in *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, pp. 249–253, 2020.
- [9] M. Chandra, S. Redkar, S. Roy, and P. Patil, “Classification of various plant diseases using deep siamese network,” 2020.
- [10] M. Sarada, “Comparative analysis of ai techniques for plant disease detection and classification on plantdoc dataset,” in *Artificial Intelligence Tools and Technologies for Smart Farm-*

- ing and Agriculture Practices*, pp. 233–261, IGI Global, 2023.
- [11] S. Shibuya, Q. H. Cap, S. Nagasawa, S. Kagiwada, H. Uga, and H. Iyatomi, “Validation of prerequisites for correct performance evaluation of image-based plant disease diagnosis using reliable 221k images collected from actual fields,” in *AI for Agriculture and Food Systems*, 2022.
- [12] T. Saikawa, Q. H. Cap, S. Kagiwada, H. Uga, and H. Iyatomi, “Aop: An anti-overfitting pretreatment for practical image-based plant diagnosis,” in *2019 IEEE International Conference on Big Data (Big Data)*, pp. 5177–5182, IEEE, 2019.
- [13] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” *Advances in neural information processing systems*, vol. 27, 2014.
- [14] E. Olaniyi, D. Chen, Y. Lu, and Y. Huang, “Generative adversarial networks for image augmentation in agriculture: a systematic review,” *arXiv preprint arXiv:2204.04707*, 2022.
- [15] M. Xin, L. W. Ang, and S. Palaniappan, “A data augmented method for plant disease leaf image recognition based on enhanced gan model network,” *Journal of Informatics and Web Engineering*, vol. 2, no. 1, pp. 1–12, 2023.
- [16] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.
- [17] Y. Tian, G. Yang, Z. Wang, E. Li, Z. Liang, *et al.*, “Detection of apple lesions in orchards based on deep learning methods of cyclegan and yolov3-dense,” *Journal of Sensors*, vol. 2019, 2019.
- [18] H. Nazki, S. Yoon, A. Fuentes, and D. S. Park, “Unsupervised image translation using adversarial networks for improved plant disease recognition,” *Computers and Electronics in Agriculture*, vol. 168, p. 105117, 2020.
- [19] Q. H. Cap, H. Uga, S. Kagiwada, and H. Iyatomi, “Leafgan: An effective data augmentation method for practical plant disease diagnosis,” *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 2, pp. 1258–1267, 2020.
- [20] K. Satoi, N. Shunta, Q. H. Cap, H. Uga, S. Kagiwada, and H. Iyatomi, “Ppig: Productive pathogenic image generation for plant disease diagnosis,” p. 554–559, 2021.
- [21] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, “Deep unsupervised learning using nonequilibrium thermodynamics,” in *International conference on machine learning*, pp. 2256–2265, PMLR, 2015.
- [22] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-resolution image synthesis with latent diffusion models,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10684–10695, 2022.
- [23] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” *arXiv preprint arXiv:1312.6114*, 2013.
- [24] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, and K. Aberman, “Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 22500–22510, 2023.
- [25] T. Ridnik, E. Ben-Baruch, A. Noy, and L. Zelnik-Manor, “Imagenet-21k pretraining for the masses,” *arXiv preprint arXiv:2104.10972*, 2021.
- [26] M. Tan and Q. Le, “Efficientnetv2: Smaller models and faster training,” in *International conference on machine learning*, pp. 10096–10106, PMLR, 2021.