

### The Phase Augmentation

伊藤, 大貴 / ITO, Daiki

---

(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)

4

(発行年 / Year)

2024-03-24

(URL)

<https://doi.org/10.15002/00030732>

# The Phase Augmentation

THE PHASE AUGMENTATION

伊藤 大貴  
Daiki ITO  
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

To build robust discriminators, it is desirable to represent the source data in the lowest possible meaningful dimension and to have a large amount of data in the feature space. data augmentation (DA) has the effect of increasing the number of pseudo-data, and appropriate DA contributes to the creation of robust discriminant boundaries. In various fields, DA contributes to improving the performance of machine learning models (ML). Fourier transforms the original data into a frequency-space representation and has been widely used as a fundamental technology in engineering. In recent years, it has been actively utilized in fields such as video compression. On the other hand, if the dimensionality of the transformed representation is used as it is, it is possible to transform the original data into a different representation without loss. Here, for low-dimensional representations constructed by deep learning model (DL) such as CNNs for classification, it is possible to obtain a lossless frequency space representation using the fourier transform, and by changing the phase angle in various ways, it is expected to create an infinite number of data representations that can be completely restored to the original data.

In this paper, we propose The Phase Augmentation, a method for learning such data, and examine its effectiveness and issues. We expected that the proposed DA method would improve the performance, but we could not confirm the performance improvement even when using ensembles.

**Keywords** : data augmentation, fast fourier transform, convolutional neural networks, latents, plant disease diagnosis

## 1. はじめに

頑健な識別器の構築のためには、なるべく意味のある低い次元で元データを表現し、かつその特徴量空間で多くのデータがあることが望ましい。data augmentation (DA) は入力空間、あるいはデータ空間での表現を変える事で、疑似的にデータ数を増やす効果をもたらす。適切な DA はなめらかな、つまり頑健な識別境界の作成に貢献する。computer vision に限らず、NLP(natural language processing) や一般的な signal processing など、ML の様々な分野で、DA は機械学習モデル (ML) の性能向上に貢献している。ところで、元の情報を周波数空間に変換するフーリエ変換 [1] は、元のデータを周波数空間表現に変換し、実時間あるいは実空間ではコストの高い計算を効率よく行うことが可能であり、工学の根拠を成す技術として広く用いられてきた。

特に近年においては、動画圧縮などの分野において、一般の画像や動画データは、低周波数成分に主要な成分を持ち、高周波成分の情報がない、あるいは限られている偏っている特性を生かし、フーリエ変換を用いることで、サンプリング定理を超えるデータの高圧縮を実現する compressed sensing [2] が注目され積極的に活用されている。一方変換された表現の次元数をそのまま用

いれば、元のデータを損失なしに異なる表現に変換することが可能である。ここで、CNN などの深層学習モデル (DL) が分類のために構築した低次元表現に対して、フーリエ変換を用い損失のない周波数空間表現を得た上で、位相角を様々に変化させれば、元データに完全に復元可能な、データ表現を無数に作る事が期待できる。

本稿では、その 1 例として植物のデータセットに対し、損失のない周波数空間表現を得た上で、学習データの位相角を様々に変化させて学習する手法である The Phase Augmentation を提案するとともに、その効果の検証と課題を述べる。ここで、DL から得られる低次元表現  $x$  をフーリエ変換して得られる低次元表現を  $z$  とする。このとき、The Phase Augmentation は、各エポックごとに一様乱数  $\theta (-\pi < \theta \leq \pi)$  を決定 (固定) し、 $z$  の位相を  $\theta$  だけ回転させて得られる低次元表現  $z_{rotate}$  を学習に用いる手法である。この手法により、学習する低次元空間上の特徴点を増やすことで、 $z$  の学習モデルに比べ、精度が向上しないか検証を行った。

## 2. 方法

### (1) 提案手法: The Augmentation

提案手法である The Phase Augmentation(Fig.1) について説明する。その前に、低次元表現  $z$  が必要とな

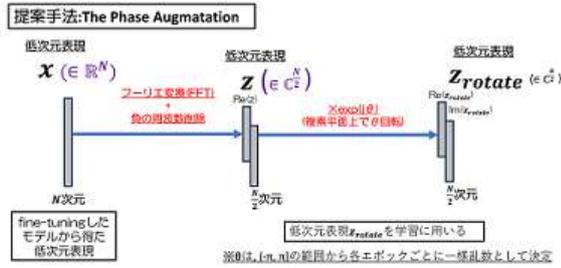


Fig. 1 提案手法: The Phase Augmentation

る。そのため、その獲得過程について説明する。まず、fine-tuningされたモデルに画像を入力し、低次元表現  $x \in \mathbb{R}^N$  を得る。その低次元表現  $x$  をフーリエ変換したのち、負の周波数の削除を行った。このことにより、低次元表現  $z \in \mathbb{C}^{\frac{N}{2}}$  を得る。提案手法である The Phase Augmentation では、各エポックごとに一様乱数  $\theta (-\pi < \theta \leq \pi)$  を決定(固定)し、 $z$  の位相を  $\theta$  だけ回転させて得られる低次元表現  $z_{rotate} (= z \cdot e^{j\theta})$  を学習に用いる手法である。

## (2) データセット

本稿で用いたデータセットについて説明する。本研究では、計 21,036 枚 (学習用:18,729 枚、検証用:2,307 枚) のいちごの葉表の画像 (Table1)、計 40,572 枚 (学習用:36,680 枚、検証用:3,892 枚) のなすの葉表の画像 (Table2)、計 115,840 枚 (学習用: 97,900 枚、検証用:17,940 枚) のきゅうりの葉表の画像 (Table3) 及び計 66,610 枚 (学習用:51,471 枚、検証用:15,139 枚) のとまとの葉表の画像 (Table4) を用いた。

## (3) 前処理

本研究で使用する低次元表現  $z$  を得るための前処理について説明する。まず、各植物画像で再学習 (fine-tuning) した EfficientNet-B4 [3] を用いて、FC 層直前の低次元表現  $x \in \mathbb{R}^N$  を獲得した。具体的には、 $512 \times 512$  の画像を EfficientNet-B4 に入力することで、 $N = 1792$  次元の低次元表現  $x$  を得た。

次に、EfficientNet-B4 に入力した画像に対する前処理を説明する。学習用画像には、RandomResizedCrop(画像をランダムに切り抜いた後、画像サイズを  $512 \times 512$  にする処理)、RandomHorizontalFlip(ランダムに画像の左右を反転する処理)、RandomVerticalFlip(ランダムに画像の上下を反転させる処理)、RandomRotation(ランダムに画像の回転を行う処理) を用いた。また、検証用画像には、RandomResizedCrop(画像をランダムに切り抜いた後、画像サイズを  $512 \times 512$  にする処理) のみを用いた。

# 3. 実験

## (1) 実験の概要

本稿では、低次元表現 (latents) を用いた Augmentation の効果と課題について検証を行った。具体的には、提案手法 The Phase Augmentation を適用することで精度が向上しないか検証を行った。尚、低次元表現  $x$  と

Table 1 いちごのデータセット (葉表) [枚]

ID_病名略称 (病名正式名称)	学習用	検証用	合計
00.HE (HEalthy)	10,472	578	11,050
01.PM (Powdery Mildew)	1,948	893	2,841
03.ANT (ANThracnose)	3,701	609	4,310
15.FW (Fusarium Wilt)	2,608	227	2,835

Table 2 なすのデータセット (葉表) [枚]

ID_病名略称 (病名正式名称)	学習用	検証用	合計
00.HE (HEalthy)	12,431	1,122	13,553
01.PM (Powdery Mildew)	7,936	938	8,874
02.GM (Gray Mold)	1,024	166	1,190
06.LM (Leaf Mold)	3,188	732	3,920
11.LS (Leaf Spot)	5,510	118	5,628
18.VW (Verticillium Wilt)	3,176	354	3,530
19.BW (Bacterial Wilt)	3,415	462	3,877

$z$  の間の関係は、可逆変換であるため、 $z$  を学習した時よりも提案手法の精度が向上すれば、 $x$  よりも精度が本質的に向上したと考えることができる。そこで、提案手法 The Phase Augmentation を用いたモデルが、 $z$  を用いたモデルに対し、精度向上しないか検証を行った。

また、検証時には、検証データをアンサンブルすることで、入力データの予測を行った。具体的には、複数の一様乱数  $\theta (-\pi < \theta \leq \pi)$  を用意し、1つの検証データに対し、複数の回転させた表現を用意する。その表現をモデルに入力し、得られる出力に Softmax 関数を掛けたものを、各クラスの確率と見て、確率の積が大きいものをモデルの予測結果とした。

さらに、比較として、 $\text{Re}(z_{rotate})$ ,  $\text{Im}(z_{rotate})$ ,  $\text{Arg}(z_{rotate})$ ,  $\text{Log}(z_{rotate})$  を学習したモデルの精度についても調査した。尚、 $\text{Arg}(z_{rotate})$  及び  $\text{Log}(z_{rotate})$  は、それぞれ多価関数  $\arg(z)$  及び  $\log(z)$  の主値を返す関数とする。

## (2) 比較に用いたモデルの構造

比較に用いたモデルの構造について説明する。Fig.2 に示したように、実部と虚部を分離させ、実部と虚部からなる実ベクトルを入力するモデルである (尚、一部、実部、または、虚部のみの場合がある)。また、活性化関数には SELU 関数を用いた。この理由は、入力データが明らかに、負の値を持つため Relu 関数のような、負の値を切り落とす活性化関数は適さないと考えたためである。

## (3) 評価方法

提案手法 The Phase Augmentation を用いたモデルの精度は、すべて調べるのが困難である。そこで、エポック数については 1, 3, 5, 998, 1,000 の 5 段階とし、アンサンブル数については、1, 2, 5, 100 の 4 段階とした。その中で最大となる各種精度 (精度, 再現率, F1 スコア) を各アンサンブル手法の精度とし、 $z$  で学習したモデルに比べ精度が向上しないか検証を行った。

Table 3 きゅうりのデータセット (葉表) [枚]

ID_病名略称 (病名正式名称)	学習用	検証用	合計
00_HE (HEalthy)	16,023	5,576	21,599
01_PM (Powdery Mildew)	7,764	1,898	9,662
02_GM (Gray Mold)	643	167	810
03_ANT (ANThracnose)	3,038	76	3,114
08_DM (Downy Mildew)	6,953	2,579	9,532
09_CLS (Corynespora Leaf Spot)	7,565	1,813	9,378
17_GSB (Gummy Stem Blight)	1,483	374	1,857
20_BS (Bacterial Spot)	4,362	2,648	7,010
22_CCYV	5,969	179	6,148
23_MD (Mosaic Diseases)	26,861	1,626	28,487
24_MYSV	17,239	1,004	18,243

Table 4 とまとのデータセット (葉表) [枚]

ID_病名略称 (病名正式名称)	学習用	検証用	合計
00_HE (HEalthy)	8,120	2,994	11,114
01_PM (Powdery Mildew)	4,490	4,250	8,740
02_GM (Gray Mold)	9,327	571	9,898
05_CLM (Cercospora Leaf Mold)	4,078	1,809	5,887
06_LM (Leaf Mold)	2,761	151	2,912
07_LB (Late Blight)	2,049	808	2,857
10_CTS (Corynespora Target Spot)	1,732	1,350	3,082
19_BW (Bacterial Wilt)	2,259	412	2,671
21_BC (Bacterial Canker)	4,369	128	4,497
27_ToMV	3,453	49	3,502
28_ToCV	4,320	871	5,191
29_YLC (Yellow Leaf Curl)	4,513	1,746	6,259

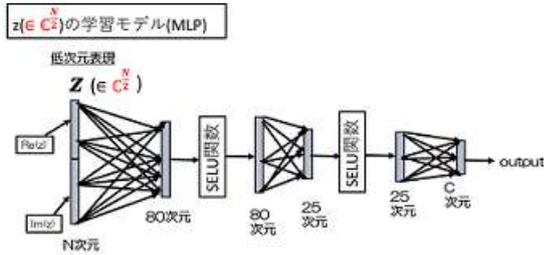


Fig. 2 比較実験に用いたモデルの構造

#### 4. 結果・考察

各植物における結果を, Table 5, Table6, Table7, Table8にそれぞれ示した. 提案手法は, ベースライン  $z$  を明らかに向上させることができなかった. これは, EfficientNet-B4 が優れた CNN であり, もともと十分な性能が得られていた可能性が高いことに起因していると考えられる. また, 今回評価に用いたモデルが, 複素数の回転に対応したモデルとは限らないため, 適切なモデルを検討していく必要があると考えられる. また, もし適切なモデルが見つかったとしても, 今回は低周波から高周波まですべての周波数にパワーが存在していたため,  $|z|^2 \neq 0$ , 即ち,  $z \neq 0$  となり,  $\text{Log}(z)$  や  $\text{Arg}(z)$  が定義できたが,  $z = 0$  が存在するデータではどのように扱うか検討する必要があると考えられる.

Table 5 精度比較結果 (イチゴ,  $z$  のみアンサンブルなし) [%]

学習データ	精度	再現率	F1 スコア
$z$ (ベースライン)	86.8	85.6	86.1
$z_{rotate}$	85.5	84.9	85.1
$\text{Re}(z_{rotate})$	86.8	85.7	86.1
$\text{Im}(z_{rotate})$	<b>87.1</b>	<b>86.6</b>	<b>86.8</b>
$\text{Arg}(z_{rotate})$	85.3	85.2	85.2
$\text{Log}(z_{rotate})$	87.0	85.1	85.9

Table 6 精度比較結果 (ナス,  $z$  のみアンサンブルなし) [%]

学習データ	精度	再現率	F1 スコア
$z$ (ベースライン)	78.4	<b>76.0</b>	75.2
$z_{rotate}$	77.8	75.3	74.2
$\text{Re}(z_{rotate})$	79.8	<b>76.0</b>	<b>75.4</b>
$\text{Im}(z_{rotate})$	80.2	75.7	74.9
$\text{Arg}(z_{rotate})$	79.0	<b>76.0</b>	75.3
$\text{Log}(z_{rotate})$	<b>81.3</b>	75.4	75.3

Table 7 精度比較結果 (きゅうり,  $z$  のみアンサンブルなし) [%]

学習データ	精度	再現率	F1 スコア
$z$ (ベースライン)	48.0	<b>56.9</b>	50.2
$z_{rotate}$	49.6	56.8	51.0
$\text{Re}(z_{rotate})$	49.4	<b>56.9</b>	51.4
$\text{Im}(z_{rotate})$	49.4	56.7	<b>51.5</b>
$\text{Arg}(z_{rotate})$	49.2	56.2	50.7
$\text{Log}(z_{rotate})$	<b>52.4</b>	55.7	50.2

Table 8 精度比較結果 (とまと,  $z$  のみアンサンブルなし) [%]

学習データ	精度	再現率	F1 スコア
$z$ (ベースライン)	62.3	<b>64.6</b>	<b>59.0</b>
$z_{rotate}$	63.2	63.2	58.2
$\text{Re}(z_{rotate})$	62.0	63.7	57.6
$\text{Im}(z_{rotate})$	62.1	63.4	56.8
$\text{Arg}(z_{rotate})$	61.3	63.6	57.7
$\text{Log}(z_{rotate})$	<b>64.0</b>	62.2	57.3

#### 5. おわりに

コストを最小限にし, 頑健な識別器を構築するためには, 汎用性の高い低次元表現における DA を考えることは極めて重要である. 今回の検証においては, フーリエ変換を用いた DA の効果は確認できなかったが, 今後も, 低次元表現を用いた様々な DA を提案するとともに, 検証していきたい.

## 参考文献

- [1] J. W. Cooley and J. W. Tukey, “An algorithm for the machine calculation of complex fourier series,” *Mathematics of computation*, vol. 19, no. 90, pp. 297–301, 1965.
- [2] M. Lustig, D. Donoho, and J. M. Pauly, “Sparse mri: The application of compressed sensing for rapid mr imaging,” *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.
- [3] M. Tan and Q. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” in *International Conference on Machine Learning*, pp. 6105–6114, PMLR, 2019.