

博士論文

小規模データ条件下における  
深層学習を用いたハイパースペクトル画像復元  
に関する研究

2024年3月22日 提出

指導教員 奥田正浩 教授

同志社大学 理工学研究科 情報工学専攻

板坂 樹生



# 論文要旨

ハイパースペクトル画像 (HSI: Hyperspectral Image) は、通常の RGB 画像と比較してより細密なスペクトル情報を持つ 3 次元データである。HSI はそのスペクトル帯域の狭さ故に、十分な照度の確保が困難になることが多く、観測される画像はノイジーになることが多い。この画質劣化は、HSI を用いたクラス分類, Unmixing, 変化検知などの応用タスクにおいてしばしば性能を低下させる要因となる。そのため、HSI 復元は実応用上において重要なプロセスとなる。

この問題を解決するために、凸最適化ベースの HSI ノイズ除去手法がこれまでに多数提案されている。HSI のノイズ除去の多くは逆問題を解くことで達成される。これらの逆問題に基づくノイズ除去手法は先験情報を正則化項としてモデル化され、それは多くの場合凸最適化問題に帰着される。凸最適化ベースの画像復元手法では、正則化項に従って画像復元を行うため、大規模な学習データセットを必要としない一方で、精度の面では学習ベースの手法に劣る。

近年の深層学習ベースの画像復元手法の発展に伴い、無劣化画像を必要としない RGB 画像復元手法がいくつか提案されている。自己教師あり学習とよばれるこれらの手法では、劣化 RGB 画像のみからなる大規模データセットから学習ペアを作成することで、画像復元モデルの学習を可能にしている。これらの手法の多くは十分なデータが利用可能である場合、モデルベースの最適化手法と比較して高い復元精度を持つ。

しかし、HSI は RGB 画像と比較して撮像に時間がかかることや、撮像に使用するカメラによって HSI の特性が変わってしまうことなどから、条件の整った大量の劣化 HSI の収集には膨大なコストがかかるため、自己教師あり RGB 画像復元手法を HSI に対して直接適用することは困難である。また、観測劣化 HSI のみから学習する深層学習ベースの HSI 復元手法も提案されているが、従来のモデルベースの手法と比較して復元精度が低い。

本論文ではこの問題を解決すべく、小規模な学習データから高精度な復元を行う HSI

ノイズ除去手法を提案する。他ドメインのデータを用いた事前学習、凸最適化ベース手法の深層展開、ゼロショット学習の三つのアプローチを軸とした手法をそれぞれ提案する。

一つ目の手法では、HSI ノイズ除去のための深層学習モデルの事前学習において、グレースケール画像、およびスペクトルデータを用いた事前学習手法を提案する。深層学習モデルにおいて使用されている Separable convolution が、空間方向の畳み込みを行う Depth-wise convolution と、チャンネル方向の畳み込みを行う Point-wise convolution に分離されていることを利用し、Depth-wise convolution カーネルをグレースケール画像、Point-wise convolution カーネルをスペクトルデータによって学習する。HSI と比較して容易かつ大量に利用可能なデータで事前学習することで、劣化が激しい条件下においてより優れた復元精度を示す。

二つ目の手法として、深層学習モデルと凸最適化ベース手法の深層展開モデルを組み合わせる手法を提案する。凸最適化ベースの画像復元では、先験情報をもとに設計された正則化項を満たすように画像を復元する。この特性を利用し、深層学習モデルに凸最適化ベース手法をもとに作成された深層展開モデルを連結させることで、画像復元モデル全体の出力が正則化項をもとに復元するよう制約する。前段の深層学習モデルは、画像復元モデル全体の表現力を向上させると同時に、後段の深層展開モデルで使用されている正則化項に適応した変換を学習する。小規模なデータセット条件下において、深層学習モデルのみの場合と比較して優れた汎化性能を持つことを示す。

三つ目の手法として、観測劣化 HSI のみから学習と復元を行うゼロショットノイズ除去手法を提案する。ゼロショット学習では観測劣化 HSI と、観測劣化 HSI に何らかの変更を加えたデータのペアを用いて学習を行う。本論文では観測劣化 HSI の劣化過程を用いる手法、Blind-spot 戦略を用いる手法を提案した。ガウシアンノイズなどの合成ノイズ除去実験、および現実世界の劣化 HSI ノイズ除去実験において、優れた復元精度を有することを示す。また同時に、Separable convolution における空間方向とチャンネル方向を分離して畳み込む構造が、HSI の構造的特徴と合致し、優れた復元結果が得られることを示す。

本論文では、各手法の小規模データ条件下における有効性に関する実験を行う。事前学習、および深層展開をベースとした2つの手法では、過学習しやすい小規模データ条件下における汎化性能の向上に寄与し、ゼロショット学習をベースとした手法では、観測劣化 HSI の劣化過程によらず高い復元精度を示す。

# 目次

<b>第 1 章</b>	<b>序論</b>	<b>15</b>
1.1	研究背景 . . . . .	15
1.2	研究目的 . . . . .	16
1.3	論文構成 . . . . .	17
<b>第 2 章</b>	<b>凸最適化および深層学習による画像復元</b>	<b>19</b>
2.1	問題設定 . . . . .	19
2.2	凸最適化による画像復元 . . . . .	20
2.2.1	凸最適化とは . . . . .	20
2.2.2	画像復元で用いられる正則化項 . . . . .	21
2.2.3	凸最適化問題を解くためのソルバ . . . . .	23
2.3	深層学習による画像復元 . . . . .	25
2.3.1	教師あり学習による画像復元 . . . . .	25
2.3.2	ノイズ除去におけるモデルアーキテクチャおよび学習手法 . . . . .	26
<b>第 3 章</b>	<b>事前学習を用いたハイパースペクトル画像復元</b>	<b>31</b>
3.1	研究背景および関連手法 . . . . .	31
3.1.1	事前学習 . . . . .	31
3.1.2	Separable convolution . . . . .	32
3.2	提案手法 . . . . .	34
3.2.1	モデルアーキテクチャ . . . . .	34
3.2.2	提案手法における事前学習手順 . . . . .	35
3.2.3	カーネル拡張 . . . . .	36
3.2.4	Point-wise convolution カーネルの事前学習 . . . . .	37
3.3	実験 . . . . .	37

3.3.1	実験設定 . . . . .	37
3.3.2	ガウシアンノイズ除去実験 . . . . .	38
3.4	結論 . . . . .	46
3.4.1	本章のまとめ . . . . .	46
3.4.2	提案手法の課題 . . . . .	47
<b>第 4 章</b>	<b>深層学習モデルと深層展開モデルを組み合わせたグレースケール画像復元</b>	<b>49</b>
4.1	研究背景および関連手法 . . . . .	49
4.1.1	深層展開 (Deep unfolding, Deep unrolling) . . . . .	50
4.1.2	深層展開によって拡張された Anisotropic TV . . . . .	51
4.2	提案手法 . . . . .	51
4.2.1	Deep unfolding モデル (DUM) . . . . .	52
4.2.2	Deep learning モデル (DLM) . . . . .	53
4.3	実験 . . . . .	53
4.3.1	実験設定 . . . . .	53
4.3.2	ガウシアンノイズ, およびラプラシアンノイズ除去実験 . . . . .	54
4.3.3	ガウシアンノイズ付きピクセル補間の画像復元実験 . . . . .	55
4.3.4	Deep learning モデルにおける前処理の効果 . . . . .	56
4.4	結論 . . . . .	56
4.4.1	本章のまとめ . . . . .	56
4.4.2	提案手法の課題 . . . . .	57
<b>第 5 章</b>	<b>HSI における Separable convolution の構造的優位性</b>	<b>59</b>
5.1	研究背景および関連手法 . . . . .	59
5.1.1	自己教師あり学習による RGB 画像復元 . . . . .	59
5.1.2	ゼロショット学習による RGB 画像復元 . . . . .	61
5.2	提案手法 . . . . .	62
5.2.1	HSI の構造的特徴について . . . . .	63
5.2.2	Separable convolution の構造的特徴について . . . . .	64
5.2.3	提案手法における学習方法 . . . . .	65
5.2.4	モデルアーキテクチャ . . . . .	66
5.3	実験 . . . . .	67
5.3.1	実験設定 . . . . .	67

5.3.2	二重劣化 HSI を用いたゼロショット HSI ノイズ除去におけるガウシアンノイズ除去実験 . . . . .	70
5.3.3	ノイズ除去タスクにおける Separable convolution の有効性 . . .	70
5.3.4	ピクセル補間タスクにおける Separable convolution の有効性 . .	74
5.3.5	クラス分類タスクにおける Separable convolution の有効性 . . .	76
5.4	結論 . . . . .	77
5.4.1	本章のまとめ . . . . .	77
5.4.2	提案手法の課題 . . . . .	77
<b>第 6 章</b>	<b>3D パターン化マスクを使用したゼロショット HSI ノイズ除去</b>	<b>79</b>
6.1	研究背景および関連手法 . . . . .	79
6.1.1	Blind2Unblind . . . . .	80
6.1.2	ゼロショット学習による HSI 復元 . . . . .	80
6.2	3D パターン化マスクを用いたゼロショット HSI ノイズ除去 . . . . .	81
6.2.1	3D パターン化マスクを用いた学習におけるモデルアーキテクチャ	82
6.2.2	3D パターン化マスクを用いた Blind-spot 戦略 . . . . .	83
6.2.3	損失関数 . . . . .	83
6.3	3D パターン化マスクを用いたゼロショットノイズ除去に関する実験 . . .	84
6.3.1	実験設定 . . . . .	84
6.3.2	ガウシアンノイズ除去実験 . . . . .	85
6.3.3	3D パターン化マスクの有効性に関する検証 . . . . .	86
6.4	ランダム 3D パターン化マスクを用いたゼロショット HSI 復元 . . . . .	88
6.4.1	ランダム 3D パターン化マスクを用いた学習におけるモデルアーキテクチャ . . . . .	88
6.4.2	ランダム 3D パターン化マスクを用いた Blind-spot 戦略 . . . . .	88
6.5	ランダム 3D パターン化マスクに関する実験 . . . . .	90
6.5.1	実験設定 . . . . .	90
6.5.2	ノイズ除去実験 . . . . .	90
6.5.3	ランダム 3D パターン化マスクの有効性に関する検証 . . . . .	98
6.6	結論 . . . . .	102
6.6.1	本章のまとめ . . . . .	102
6.6.2	提案手法の課題 . . . . .	102

第 7 章 結論	103
参考文献	106



# 表目次

3.1	標準偏差 $\sigma = 0.2$ におけるガウシアンノイズ除去結果 . . . . .	39
3.2	標準偏差 $\sigma = 0.3$ におけるガウシアンノイズ除去結果 . . . . .	39
3.3	標準偏差 $\sigma = 0.4$ におけるガウシアンノイズ除去結果 . . . . .	39
4.1	ガウシアンノイズ除去結果 . . . . .	54
4.2	ラプラシアンノイズ除去結果 . . . . .	55
4.3	ガウシアンノイズとピクセル補間の混合ノイズにおける実験結果 . . . . .	55
5.1	各データにおけるモデルの学習可能パラメータ数 . . . . .	68
5.2	ガウシアンノイズ除去実験結果 . . . . .	69
5.3	畳み込み手法変更におけるガウシアンノイズ除去結果 (PSNR) . . . . .	70
5.4	畳み込み手法変更におけるガウシアンノイズ除去結果 (SSIM) . . . . .	71
5.5	畳み込み手法変更におけるラプラシアンノイズ除去結果 (PSNR) . . . . .	72
5.6	畳み込み手法変更におけるラプラシアンノイズ除去結果 (SSIM) . . . . .	72
5.7	畳み込み手法変更におけるピクセル補間結果 (PSNR) . . . . .	74
5.8	畳み込み手法変更におけるピクセル補間結果 (SSIM) . . . . .	74
5.9	クラス分類モデルにおける学習可能パラメータ数 . . . . .	76
5.10	クラス分類結果 . . . . .	76
6.1	実験に使用する HSI のサイズ . . . . .	84
6.2	標準偏差 $\sigma = 0.1$ における 3D パターン化マスクを用いたガウシアンノイズ除去結果 . . . . .	86
6.3	標準偏差 $\sigma = 0.15$ における 3D パターン化マスクを用いたガウシアンノイズ除去結果 . . . . .	87
6.4	3D パターン化マスクによる Blind-spot 戦略の有効性に関する検証 . . . . .	88

6.5	標準偏差 $\sigma = 0.1$ におけるランダム化 3D パターン化マスクを用いたガウシアンノイズ除去実験結果 . . . . .	91
6.6	標準偏差 $\sigma = 0.15$ におけるランダム化 3D パターン化マスクを用いたガウシアンノイズ除去実験結果 . . . . .	92
6.7	標準偏差 $\sigma = 0.1$ におけるランダム化 3D パターン化マスクを用いたラプラシアンノイズ除去実験結果 . . . . .	93
6.8	標準偏差 $\sigma = 0.15$ におけるランダム化 3D パターン化マスクを用いたラプラシアンノイズ除去実験結果 . . . . .	94
6.9	ノイズ強度定数 $s = 15$ におけるランダム化 3D パターン化マスクを用いたポアソンノイズ除去実験結果 . . . . .	95
6.10	ノイズ強度定数 $s = 6$ におけるランダム化 3D パターン化マスクを用いたポアソンノイズ除去実験結果 . . . . .	96
6.11	$K = 2$ における畳み込みの比較 . . . . .	98
6.12	$K = 3$ における畳み込みの比較 . . . . .	98
6.13	ランダム 3D パターン化マスクとその他の Blind-spot 戦略の比較 . . . .	100

# 図目次

1.1	RGB 画像と HSI の違い . . . . .	16
2.1	画像復元 モデル図 . . . . .	20
2.2	凸の例 . . . . .	21
2.3	TV における差分画像のイメージ図 . . . . .	22
2.4	Soft thresholding . . . . .	25
2.5	Residual learning モデル図 . . . . .	28
3.1	通常の畳み込み層における処理のモデル図 . . . . .	33
3.2	Depth-wise convolution における処理のモデル図 . . . . .	33
3.3	Point-wise convolution における処理のモデル図 . . . . .	33
3.4	提案手法におけるモデルアーキテクチャ . . . . .	34
3.5	提案手法における学習手順 . . . . .	35
3.6	カーネル拡張手順 . . . . .	36
3.7	標準偏差 $\sigma = 0.2$ における Frisco データに対するノイズ除去結果画像 . . . . .	40
3.8	標準偏差 $\sigma = 0.2$ における PaviaC データに対するノイズ除去結果画像 . . . . .	41
3.9	標準偏差 $\sigma = 0.3$ における Frisco データに対するノイズ除去結果画像 . . . . .	42
3.10	標準偏差 $\sigma = 0.3$ における PaviaC データに対するノイズ除去結果画像 . . . . .	43
3.11	標準偏差 $\sigma = 0.4$ における Frisco データに対するノイズ除去結果画像 . . . . .	44
3.12	標準偏差 $\sigma = 0.4$ における PaviaC データに対するノイズ除去結果画像 . . . . .	45
4.1	深層展開による反復アルゴリズムの拡張モデル図 . . . . .	50
4.2	提案手法 モデルアーキテクチャ . . . . .	51
4.3	Deep unfolding モデル (DUM) のモデルアーキテクチャ . . . . .	52
4.4	Deep learning モデル (DLM) のモデルアーキテクチャ . . . . .	53
4.5	表 4.1 における $N_{\text{train}} = 50$ かつ $\sigma = 0.15$ の場合の各モデルの結果画像 . . . . .	57

5.1	教師あり学習と Noise2Noise における学習用データペアの違い . . . . .	60
5.2	Deep image prior におけるデータペア . . . . .	62
5.3	HSI に対するタッカー分解の特異値プロット . . . . .	63
5.4	HSI の各軸方向差分のヒストグラム . . . . .	64
5.5	提案手法における学習用データペア作成方法 . . . . .	65
5.6	モデルアーキテクチャ モデル図 . . . . .	66
5.7	実験で使用される畳み込み手法による処理の違い . . . . .	67
5.8	ガウシアンノイズ除去における畳み込み手法による復元結果の違い . . . .	71
5.9	ラプラシアンノイズ除去における畳み込み手法による復元結果の違い . .	73
5.10	ピクセル補間タスクにおける畳み込み手法による復元結果の違い . . . .	75
6.1	Blind2Unblind における Blind-spot 戦略 . . . . .	80
6.2	Deep hs prior におけるデータペア . . . . .	81
6.3	3D パターン化マスクを用いたゼロショット HSI ノイズ除去におけるモ デルアーキテクチャ . . . . .	82
6.4	3D パターン化マスクを用いたゼロショット HSI ノイズ除去における Blind-spot 戦略 . . . . .	83
6.5	ランダム 3D パターン化マスクを用いたゼロショット HSI ノイズ除去に おけるモデルアーキテクチャ . . . . .	89
6.6	3D パターン化マスクを用いたゼロショット HSI ノイズ除去における Blind-spot 戦略 . . . . .	89
6.7	実世界の劣化 HSI に対するノイズ除去結果 . . . . .	97
6.8	標準偏差 0.15 のガウシアンノイズ除去におけるそれぞれの畳み込み層を 用いたモデルの PSNR プロット . . . . .	99
6.9	標準偏差 0.1 のガウシアンノイズ除去における Blind-spot 戦略ごとの PSNR のプロット . . . . .	101

# 論文中の文字，記号，および数式について

論文中で使用される記号，および数式は原則として以下のルールに従って表記される。ルールから外れて文字，記号，数式が使用される場合は本文中に明記する。

$x$  のような小文字のイタリック体はスカラー， $\mathbf{x}$  のような小文字のボールド体はベクトル， $\mathbf{X}$  のような大文字のボールド体は行列，および 3 次元以上のテンソルを表す。ベクトルのサイズは要素数  $p$  かつ要素が実数で構成されるベクトルの場合， $\mathbf{x} \in \mathbb{R}^p$  のように表記し，特に言及がない限り列ベクトルとする。行列のサイズは  $p$  行  $q$  列の場合， $\mathbf{X} \in \mathbb{R}^{p \times q}$  のように表記する。3 次元テンソルより大きな次元の場合も同様に表記する。また，ベクトル  $\mathbf{x}$ ，および行列  $\mathbf{X}$  の転置は，それぞれ  $\mathbf{x}^T$ ， $\mathbf{X}^T$  で表記される。

$\|\cdot\|_p$  は  $l_p$  ノルムを表しており， $\|\mathbf{x}\|_2$  の場合， $l_2$  ノルム（ユークリッドノルム）を表す。集合は，大文字のイタリック体で  $S$  のように表記する。

関数は，イタリック体の文字で表す。パラメータを持たない関数の場合，関数は  $f(\mathbf{x})$  のように表される。また，パラメータ  $\Theta$  を持つ関数（深層学習モデルを含む）は， $f(\mathbf{x}; \Theta)$  のように表記する。ここで  $\mathbf{x}$  は関数に与えられる引数であり， $n$  個の引数を与える関数の場合は  $f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n; \Theta)$  のように表記する。

$\min$  は目的関数の最小化を表し， $\underset{\mathbf{x}}{\operatorname{argmin}}$  は目的関数を最小化する  $\mathbf{x}$  を表す。



# 第 1 章

## 序論

### 1.1 研究背景

近年、センシング技術が著しく発展している。それに伴い、ハイパースペクトル画像、イベントデータ、深度マップなどに代表される情報源の高度化が進んでいる。これらの情報源は、一般的な RGB 画像や動画データが持ちえない、高いスペクトル分解能、高い時間分解能、カメラから撮像対象の物体までの距離情報などを保持しており、それぞれのデータの特性に応じたさまざまな応用タスクで用いられている。

その中でも、ハイパースペクトル画像 (HSI: Hyperspectral Image) は、図 1.1 に示すように、通常の RGB 画像と比較してより細密なスペクトル情報を持つ 3 次元データである。RGB 画像では赤、緑、青の 3 バンドの色情報を保持しているのに対し、HSI では数十から数百ものバンド数を持つ。物質ごとに、特定の波長の光をどの程度吸収、反射、透過するかが決まっているため、波長ごとの情報をより細かく取得できる HSI では、RGB 画像と比較してより高精度な物体識別を行うことができる。この特性を活かし、HSI は地球探査 [1,2]、資源探査 [3,4]、医療診断 [5,6]、食品検査 [7,8] など、様々な目的で使用されている。しかし、HSI はそのスペクトル帯域の狭さ故に、十分な照度の確保が困難になることが多く、観測される画像はノイジーになることが多い。この画質劣化は、HSI を用いたクラス分類 [9–12]、Unmixing [13]、変化検知 [14–16] などの応用タスクにおいてしばしば性能を低下させる要因となる。そのため、HSI 復元は実応用上において重要なプロセスとなる。

この問題を解決するために、HSI ノイズ除去 (Denoising) 手法が多数提案されている。HSI のノイズ除去の多くは逆問題 (Inverse problem) を解くことで達成される。これらの逆問題に基づくノイズ除去手法は先験情報 (Prior knowledge) を正則化項 (Reg-

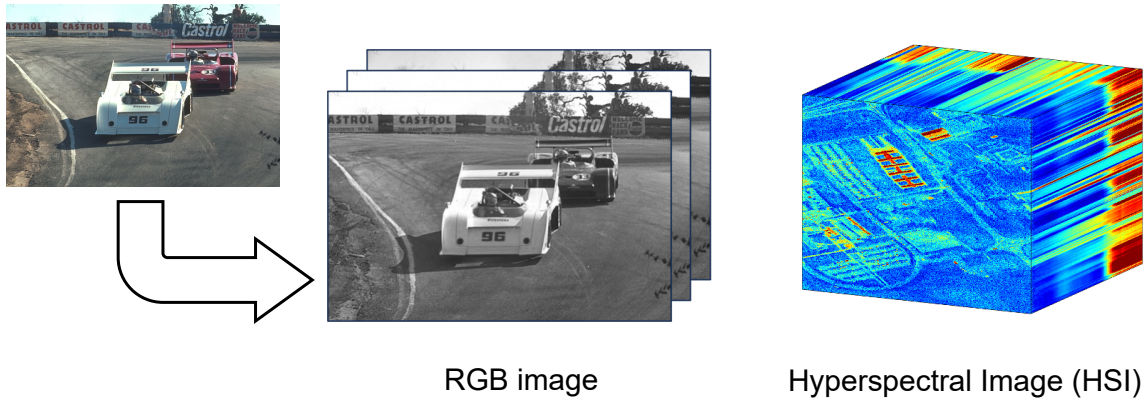


図 1.1: RGB 画像と HSI の違い

ularization) としてモデル化し、それは多くの場合凸最適化問題 (Convex optimization problem) に帰着される. 最もよく用いられる正則化の一つに Total Variation (TV) がある. HSI やマルチチャンネル画像に TV を応用することでノイズ除去を達成する手法が多く提案されている [17–21]. さらに HSI の低ランク性 (Low-rank property) を利用した手法も提案されており、スパース性 (Sparsity) と低ランク性を同時に考慮することで、高い精度のノイズ除去が実現できる [22–27]. 凸最適化ベースの画像復元手法では、正則化項に従って画像復元を行うため、大規模な学習データセットを必要としない一方で、精度の面では学習ベースの手法に劣る.

また、近年では RGB 画像に対する深層学習ベースの画像復元手法が多く提案されている [28–36]. 深層学習を用いた一般的な RGB 画像復元では、モデルに入力する劣化画像と、モデルが目標とする無劣化画像 (教師画像) のデータペアからなる大規模データセットで学習することで、画像復元モデルを得る. これらの手法の多くは十分なデータが利用可能である場合、モデルベースの最適化手法と比較して高い復元精度を達成することが多い. しかし、十分なデータを利用できない場合、深層学習モデルが過学習してしまうため未知のデータを復元することができず、結果として最適化ベース手法の精度が優れている場合がある.

## 1.2 研究目的

深層学習を用いた HSI 復元では、深層学習モデルの高い関数近似能力から、優れた復元精度が期待できる. しかし、深層学習を用いた HSI 復元を行うには、いくつかの問題点が存在する. 1 つ目は、無劣化画像を得ることが困難である点である. 第 1.1 節でも述べた



ように HSI は撮像時に劣化しやすいため、劣化 HSI と無劣化 HSI のペアを作成することが困難である。2つ目は、HSI は RGB 画像と比較して撮像に時間がかかる上、データサイズが大幅に大きい点である。撮像に時間がかかることは、大規模データセットを用意するためのコストが増大することを意味する。加えて大きなデータサイズは、CPU、GPU 上にデータを大量に読み込むことを困難にするため、モデルの学習コストを増大させる。3つ目は、撮像に使用するハイパースペクトルカメラによって、空間分解能、およびスペクトル分解能が異なる点である。ハイパースペクトルカメラに使用されているセンサによって、取得するスペクトル範囲やバンド数が異なる。バンド数が異なる場合、深層学習モデルへ入力するデータの形状が一意に定まらないため、モデルアーキテクチャに制約がかかる。4つ目は、深層学習モデルのパラメータ数が大規模になりやすい点である。HSI は RGB 画像と比較して膨大なバンド数を保有しているため、モデルが十分な性能を発揮するだけのチャンネル数を確保することが難しい。深層学習モデルの規模はリソースの制約を受ける上、大きくなり過ぎると過学習の危険性が高まる。

本研究では整備された大規模データセットを用意することが困難なデータである HSI に対し、小規模な学習データから高精度な復元を行うノイズ除去手法を提案することを目的とする。上述した問題を解決するために、大きく分けて3つのアプローチを提案する。各アプローチの概要は第 1.3 節で説明する。

## 1.3 論文構成

本論文における章構成は以下のようになっている。第 1 章では、HSI の特性、HSI 復元に関する従来研究、深層学習モデルを用いた RGB 画像復元に関する研究について簡単に紹介した。

第 2 章では、凸最適化や深層学習を用いた従来の画像復元手法に関して解説する。凸最適化に関する説明では、凸最適化問題として帰着することの有用性、画像復元において使用される正則化項、凸最適化を解くために使用されるアルゴリズムについて説明する。深層学習に関する説明では、基本的な教師あり学習、および教師あり学習ベースの RGB 画像復元手法について説明する。

第 3 章では、他ドメインのデータを用いた事前学習を活用した HSI ノイズ除去について述べる。深層学習モデルの表現力を維持しつつ必要パラメータ数を減らすために、Separable convolution を導入する。また、Separable convolution が空間方向の畳み込みとチャンネル方向の畳み込みをそれぞれ独立に行っていることに着目し、それぞれのカーネルを容易に入手可能なグレースケール画像、およびスペクトル情報で事前学習する。事前

学習において HSI とグレースケール画像でデータの形状が異なるため、グレースケール画像で学習後に空間方向の畳み込みのカーネルを拡張することで対処する。

第 4 章では、畳み込みニューラルネットワークを用いた前処理用深層学習モデルと、先験情報を組み込んだ深層展開モデルを組み合わせる手法について述べる。最適化ベースの手法では先験情報をもとに画像を復元することで、学習データなしでの復元を可能にする。この手法ではその仕組みを利用し、画像復元モデルに先験情報を組み込むことで、モデル全体の出力を制約し過学習を抑制する。また、同時に最適化ベースの手法より高い表現力を持つニューラルネットワークを前処理に利用することで、画像復元モデルの表現力を向上させる。

第 5 章では、ゼロショット学習ベースの HSI 復元手法、およびゼロショット HSI 復元に適した畳み込み構造について述べる。観測した劣化 HSI を教師画像、教師画像をさらに劣化させた二重劣化 HSI を入力とするゼロショット学習アーキテクチャを新たに提案する。また、ゼロショット HSI 復元における Separable convolution の構造的な優位性をいくつかの実験によって示す。

第 6 章では、追加データを必要としないゼロショット HSI ノイズ除去手法について述べる。従来手法や第 5 章で提案した手法では、モデルを学習するための追加情報を必要とする。提案手法では、Blind-spot 戦略を用いることでこれらの必要性を排除し、観測劣化 HSI のみからノイズ除去を行う手法を提案する。3D パターン化マスクを用いることで、空間方向とスペクトル方向両方の情報から、Blind-spot に対応する画素を復元するように学習する。また Blind-spot を特定のパターンではなくランダムに選択するランダム 3D パターン化マスクも同時に提案する。ランダム化することによって過学習を抑制し、よりロバストな学習を可能にする。

第 7 章では、本論文についてまとめる。

## 第2章

# 凸最適化および深層学習による画像復元

### 2.1 問題設定

画像復元では、図 2.1 に示すようにピクセルのランダムな欠損、ノイズ、ライン欠損などによって劣化した画像から、元の無劣化画像に近い画像を復元する。様々な画像劣化モデルが存在するが、本論文では以下の式に示す画像劣化モデルを主に使用する。

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \quad (2.1)$$

ここで、 $\mathbf{y} \in \mathbb{R}^p$  は要素数  $p$  のベクトル化した劣化画像、 $\mathbf{A} \in \mathbb{R}^{p \times q}$  は線形変換を行う  $p$  行  $q$  列の行列、 $\mathbf{x} \in \mathbb{R}^q$  は要素数  $q$  のベクトル化した無劣化画像、 $\mathbf{n} \in \mathbb{R}^q$  は画像に付加されるノイズである。式 (2.1) では、 $\mathbf{A}$ 、および  $\mathbf{n}$  を変化させることで、様々な劣化を表現することができる。画像復元タスクでは、与えられた  $\mathbf{y}$  に対し、無劣化画像  $\mathbf{x}$  にできるだけ近い復元画像  $\hat{\mathbf{x}}$  を得ること、あるいはそのような  $\hat{\mathbf{x}}$  を出力する関数  $f(\cdot; \Theta)$  を得ることを目的とする。

本論文では、小規模データセットしか扱うことができない条件下における高精度なノイズ除去手法を提案することを目標とする。ノイズ除去タスクの場合、式 (2.1) は以下のように表される。

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \quad (2.2)$$

$\mathbf{n}$  は主に標準偏差  $\sigma$ 、平均 0 のガウシアンノイズ（白色ガウシアンノイズ）が用いられる。

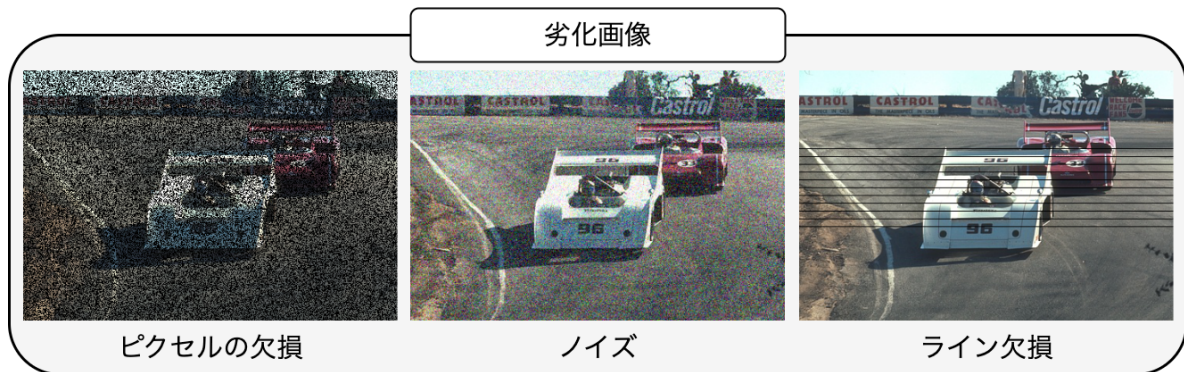


図 2.1: 画像復元 モデル図

## 2.2 凸最適化による画像復元

### 2.2.1 凸最適化とは

$P$  次元ベクトルを  $(-\infty, \infty]$  にマッピングする目的関数  $\Delta : \mathbb{R}^P \rightarrow (-\infty, \infty]$  に対して制約条件を表す集合を  $S \subset \mathbb{R}^P$  とした場合、連続最適化問題の一般形は

$$\min_{\mathbf{x} \in \mathbb{R}^P} \Delta(\mathbf{x}) \quad s.t. \quad \mathbf{x} \in S \quad (2.3)$$

で表すことができる。ここで、関数  $\Delta$  が凸関数、集合  $S$  が凸集合であるとき、式 (2.3) は凸最適化問題である。ここで、凸集合とは集合  $G \subset \mathbb{R}^P$  において、 $\alpha \in [0, 1]$  かつ  $\mathbf{x}_1, \mathbf{x}_2 \in G$  のとき、 $\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2 \in G$  を満たすような集合のことをいう。これは図 2.2a に示すように、集合  $G$  の任意の 2 点間を結ぶ線分全体が全て  $G$  に含まれるような集合であると解釈できる。また、凸関数とは  $\alpha \in [0, 1]$  かつ  $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^P$  のとき、 $\alpha \Delta(\mathbf{x}_1) + (1 - \alpha) \Delta(\mathbf{x}_2) \leq \Delta(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2)$ 、または  $\alpha \Delta(\mathbf{x}_1) + (1 - \alpha) \Delta(\mathbf{x}_2) \geq$

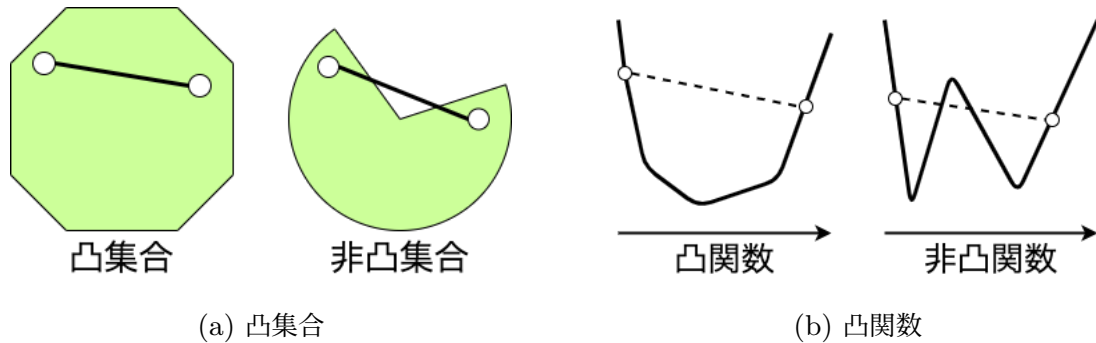


図 2.2: 凸の例

$\Delta(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2)$  を満たすような関数である。条件式によって関数が出っ張っている方向が異なり、前者の条件式は上に凸、後者の条件式は下に凸な凸関数を表す。上に凸な関数は符号を逆にすることで、下に凸な関数として扱うことができる。以降は特に言及がない限り、凸関数とは下に凸な関数を指す。これは図 2.2b に示すように、関数  $\Delta$  中の任意の 2 点間を結んだときできる線分全体が、 $\Delta$  のグラフよりも常に下（あるいは上）にあるような関数であると解釈できる。凸最適化問題では局所解と大域最適解が一致するため、初期解を逐次的に更新することにより大域最適解を求めることができる。しかし、未知のデータにうまく適合するハイパーパラメータを見つけることが難しい、深層学習モデルのように事前学習済みモデルを使用することができないため、正則化項が複雑になるほど推論時に計算量を必要とするなどの問題点がある。

### 2.2.2 画像復元で用いられる正則化項

凸最適化による信号復元では、復元対象となる信号が本来持つ特性（先験情報）を正則化項としてモデル化し、目的関数を最小化することで復元された信号を得る。凸最適化を用いた画像復元における最適化式の多くは以下ようになる。

$$\operatorname{argmin}_{\hat{\mathbf{x}}} \frac{1}{2} \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|_2^2 + \lambda f(\hat{\mathbf{x}}) \quad (2.4)$$

ここで、 $\frac{1}{2} \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|_2^2$  は観測した劣化信号と復元信号との差を測る項（Fidelity 項）であり、 $f(\hat{\mathbf{x}})$  は先験情報を元に設計された正則化項（Regularization 項）、 $\lambda$  は正則化の強さを決定するハイパーパラメータである。設計した正則化項が無劣化画像の特性を的確に捉えているほど、復元精度が向上する。深層学習ベースの手法とは異なり、復元対象データのドメインに関する知識が必要となるものの、大量の学習用データは必要としない。その

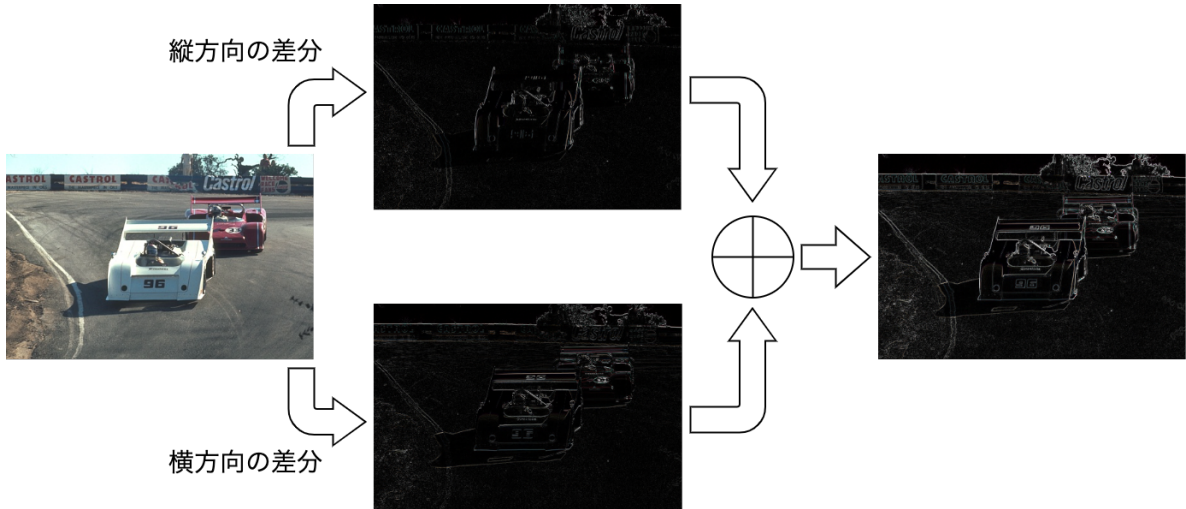


図 2.3: TV における差分画像のイメージ図

ため、HSI のように大規模データセットを用意することが困難なドメインに対しても、効果的な復元が可能である。

画像は主に低周波数成分で構成されている（ほとんどの部分で滑らかに変化する）という先験情報から、画像復元における正則化項では Total Variation (TV) が良く用いられる。TV では隣接画素差分のノルムを計算し総和をとるため、入力した画像の滑らかさを評価することができる。TV には Isotropic TV と Anisotropic TV の 2 種類が存在する。それぞれの定義は以下のようにになっている。

$$\text{TV}_{\text{iso}}(\mathbf{v}) = \sum_{i,j} \sqrt{(v_{i+1,j} - v_{i,j})^2 + (v_{i,j+1} - v_{i,j})^2} \quad (2.5)$$

$$\text{TV}_{\text{aniso}}(\mathbf{v}) = \sum_{i,j} |v_{i+1,j} - v_{i,j}| + |v_{i,j+1} - v_{i,j}| \quad (2.6)$$

ここで、 $v_{i,j}$  は画像  $\mathbf{v} \in \mathbb{R}^{H \times W}$  の  $i$  行  $j$  列目のピクセルの値を表す。縦方向と横方向に隣接した画素の差分を計算し、それぞれ  $l_2$  ノルムと  $l_1$  ノルムを計算後、総和をとっている。図 2.3 に示すように、自然画像においてノイズによって劣化していない画像の場合、物体の輪郭部分の差分が大きくなり、それ以外の場所では差分がほとんどゼロになる。ノイズは主に高周波成分で構成されているため、ノイズが多く含まれる場合は隣接画素との差分値が大きくなり、TV 正則化項の値が大きくなる。HSI 復元においては、上述した TV や、スパース性、低ランク性を利用した様々な正則化項を用いた手法がこれまでに多数提案されている [17–27, 37, 38]。

### 2.2.3 凸最適化問題を解くためのソルバ

凸最適化問題を解くためのアルゴリズムとして、交互方向乗数法 [39] (ADMM: Alternating Direction Method of Multipliers) や主-双対近接分離法 [40] (Primal-dual splitting method) がよく用いられる。本小節では ADMM について主に説明する。ADMM は以下のような最適化問題に対して適用可能である。

$$\min_{\hat{\mathbf{x}} \in \mathbb{R}^P, \mathbf{z} \in \mathbb{R}^Q} g(\hat{\mathbf{x}}) + h(\mathbf{z}) \quad s.t. \mathbf{z} = \mathbf{G}\hat{\mathbf{x}} \quad (2.7)$$

ここで、 $g(\hat{\mathbf{x}})$  は二次関数程度の簡単な凸関数、 $h(\mathbf{z})$  は近接写像が計算可能な凸関数であり、 $\mathbf{G}$  は列フルランクである必要がある。ADMM では式 (2.7) の最適化を、 $\hat{\mathbf{x}}$  に関する最適化、 $\mathbf{z}$  に関する最適化、双対変数である  $\mathbf{b}$  の調整の 3 つのプロセスに分割し、それらを繰り返し解く。ADMM のアルゴリズムを Algorithm1 に示す。ここで、 $\text{prox}_{\gamma h}(\cdot)$  は

---

**Algorithm 1** Alternating Direction Method of Multipliers (ADMM)

---

```
input  $t = 0, \mathbf{x}^{(0)}, \mathbf{z}^{(0)}, \mathbf{b}^{(0)}$ 
while A stopping criterion is not satisfied do
   $\hat{\mathbf{x}}^{(t+1)} = \underset{\hat{\mathbf{x}}}{\text{argmin}} g(\hat{\mathbf{x}}) + \frac{1}{2\gamma} \|\mathbf{z}^{(t)} - \mathbf{G}\hat{\mathbf{x}} - \mathbf{b}^{(t)}\|^2$ 
   $\mathbf{z}^{(t+1)} = \text{prox}_{\gamma h}(\mathbf{G}\hat{\mathbf{x}}^{(t+1)} + \mathbf{b}^{(t)})$ 
   $\mathbf{b}^{(t+1)} = \mathbf{b} + \mathbf{G}\hat{\mathbf{x}}^{(t+1)} - \mathbf{z}^{(t+1)}$ 
   $t \leftarrow t + 1$ 
end while
Output :  $\hat{\mathbf{x}}^{(t)}$ 
```

---

近接写像である。近接写像では入力されたベクトルとユークリッドノルムの意味で距離が近いベクトルから、関数  $h(\cdot)$  の値を最小にするようなベクトルを選択する操作である。近接写像の定義は以下の式で与えられる。

$$\text{prox}_{\gamma h}(\mathbf{v}) = \underset{\mathbf{u}}{\text{argmin}} h(\mathbf{u}) + \frac{1}{2\gamma} \|\mathbf{u} - \mathbf{v}\|^2 \quad (2.8)$$

ここでハイパーパラメータ  $\gamma$  は、入力されたベクトル  $\mathbf{v}$  に対してどの程度離れることを許容するかを決定するパラメータである。 $\gamma$  が非常に大きい場合、式 (2.8) における右側の項の重みが小さくなるため、 $\mathbf{v}$  から大きく異なる  $\mathbf{u}$  を選択することを許容するようになる。

具体例として式 (2.6) の Anisotropic TV を正則化項とした場合の画像復元の例を示す。ここでは単純化のため、画素値に関する制約はつけていない。Anisotropic TV を式 (2.4) の  $f(\hat{\mathbf{x}})$  に代入すると、以下のような式になる。

$$\operatorname{argmin}_{\hat{\mathbf{x}}} \frac{1}{2} \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|_2^2 + \lambda \text{TV}_{\text{aniso}}(\hat{\mathbf{x}}) \quad (2.9)$$

この  $\text{TV}_{\text{aniso}}(\cdot)$  関数を、縦方向の差分を取る行列  $\mathbf{D}_x$  と横方向の差分を取る行列  $\mathbf{D}_y$  を縦方向に連結した行列  $\mathbf{D}$  に置き換えると以下のようなになる。

$$\hat{\mathbf{x}} = \operatorname{argmin}_{\hat{\mathbf{x}}} \frac{1}{2} \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{D}\hat{\mathbf{x}}\|_1 \quad (2.10)$$

ここで、 $g(\hat{\mathbf{x}}) = \frac{1}{2} \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|_2^2$ ,  $h(\mathbf{z}) = \lambda \|\mathbf{z}\|_1$ ,  $\mathbf{z} = \mathbf{D}\mathbf{x}$  とすると、式 (2.7) の形式と一致する。

次に ADMM のアルゴリズムに適用する形で、式 (2.10) を変形すると以下のような式になる。

$$\hat{\mathbf{x}}^{(k+1)} = \operatorname{argmin}_{\hat{\mathbf{x}}} \frac{1}{2} \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|_2^2 + \frac{1}{2\gamma} \|\mathbf{z}^{(k)} - \mathbf{D}\hat{\mathbf{x}} - \mathbf{b}^{(k)}\|_2^2 \quad (2.11)$$

$$\mathbf{z}^{(k+1)} = \operatorname{argmin}_{\mathbf{z}} \lambda \|\mathbf{z}\|_1 + \frac{1}{2\gamma} \|\mathbf{z} - \mathbf{D}\hat{\mathbf{x}}^{(k+1)} - \mathbf{b}^{(k)}\|_2^2 \quad (2.12)$$

$$\mathbf{b}^{(k+1)} = \mathbf{b} + \mathbf{D}\hat{\mathbf{x}}^{(k+1)} - \mathbf{z}^{(k+1)} \quad (2.13)$$

ここで、式 (2.11) はどちらの項も二次関数の形であるため、容易に解析解を求めることができる。また、式 (2.13) はベクトルの和、および差の計算であるため、こちらも容易に計算することができる。しかし、式 (2.12) は  $l_1$  ノルムの近接写像が必要になる。

式 (2.12) における  $\operatorname{argmin}_{\mathbf{z}} \lambda \|\mathbf{z}\|_1 + \frac{1}{2\gamma} \|\mathbf{z} - \mathbf{D}\hat{\mathbf{x}}^{(k+1)} - \mathbf{b}^{(k)}\|_2^2$  は全体を  $\lambda$  で割っても解が変わらないことと、式 (2.8) から、近接写像は以下のようなになる。

$$\operatorname{prox}_{\gamma \|\cdot\|_1}(\mathbf{v}) = \operatorname{argmin}_{\mathbf{u}} \|\mathbf{u}\|_1 + \frac{1}{2\gamma\lambda} \|\mathbf{u} - \mathbf{v}\|^2 \quad (2.14)$$

ここで、 $l_1$  ノルムも  $l_2$  ノルムも他のベクトルの要素に対して独立しているため、式 (2.14) の計算はベクトルの  $i$  番目の要素に対してそれぞれ分けて行うことができる。その場合の近接写像の各要素は以下の式で求められる。

$$[\operatorname{prox}_{\gamma \|\cdot\|_1}(\mathbf{v})]_i = \operatorname{argmin}_{u_i \in \mathbb{R}} |u_i| + \frac{1}{2\gamma\lambda} (u_i - v_i)^2 \quad (2.15)$$

この式は容易に解析解を求めることが可能であり、最終的な近接写像は以下の式で与えられる。

$$[\operatorname{prox}_{\gamma \|\cdot\|_1}(\mathbf{v})]_i = \operatorname{sgn}(v_i) \max\{|v_i| - \gamma\lambda, 0\} \quad (2.16)$$



ここで、 $\text{sgn}(\cdot)$  は符号関数であり、以下のように定義される。

$$\text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases} \quad (2.17)$$

$\gamma\lambda = 1$  の場合における式 (2.16) をプロットすると、図 2.4 のようになる。  $l_1$  ノルムの近

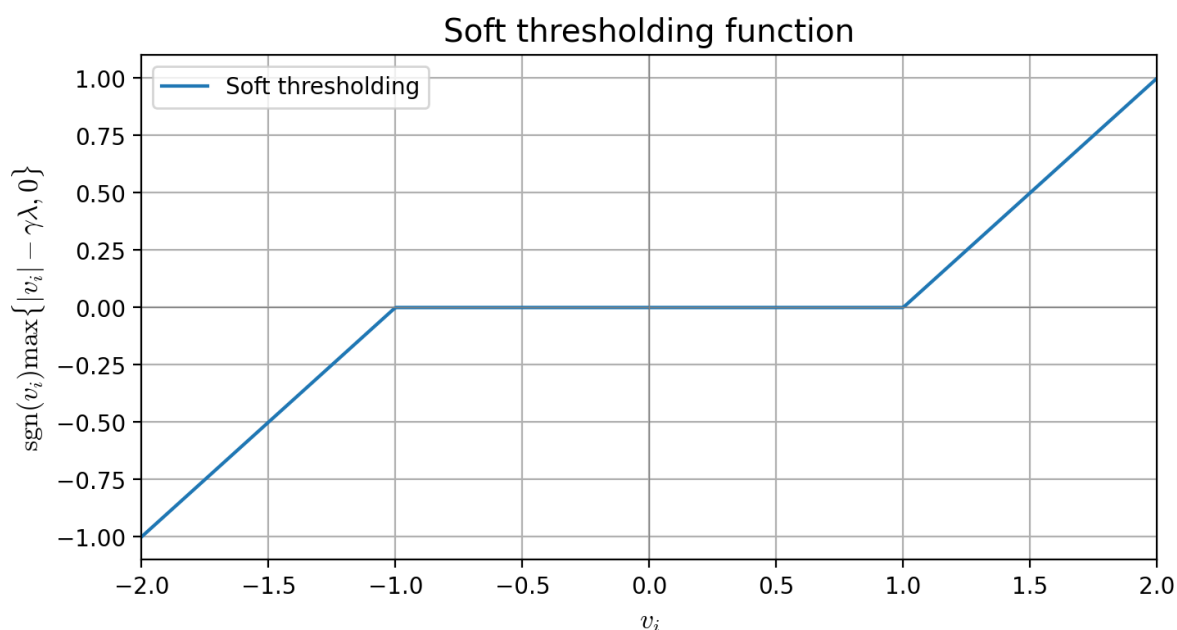


図 2.4: Soft thresholding

接写像は、図 2.4 に示すように Soft thresholding と呼ばれる処理で与えられる。このように微分不可能な幾つかの凸関数において、近接写像が効率的に計算できることが知られており、 $l_1$  ノルムはそのうちの一つである。また非凸ではあるものの、フィルタ設計において ADMM を用いた  $l_0$  ノルム最小化が試みられており、実験的には収束することが示されている [41]。

## 2.3 深層学習による画像復元

### 2.3.1 教師あり学習による画像復元

教師あり学習をベースとした画像復元のための深層学習モデルでは、復元したい対象ドメインのデータにおいて、劣化画像と理想とする出力である教師画像のペアからなる大規

模データセット  $D_{\text{train}} = \{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots, (\mathbf{x}_n, \mathbf{y}_n), \dots, (\mathbf{x}_N, \mathbf{y}_N)\}$  を用いて学習する。ここで、 $\mathbf{x}_n$  は  $n$  枚目の無劣化画像、 $\mathbf{y}_n$  は  $n$  枚目の劣化画像、 $N$  は学習用データセット全体の枚数である。解決したいタスクに合わせて適切に  $D_{\text{train}}$  を作成することで、ボケ除去、モーションブラー除去、超解像、雨粒除去、デノイジングなど各タスクに対応する深層学習モデルを得ることができる。

深層学習モデルにおける学習では、与えられた損失関数  $L(\mathbf{y}, \mathbf{x})$  を最小化するような出力が得られるようにモデルが持つ重みパラメータ  $\Theta$  を調整する。学習により調整されたパラメータ  $\Theta^*$  は以下の最適化式で求められる。

$$\Theta^* = \underset{\Theta}{\operatorname{argmin}} \frac{1}{N} \sum_{n=1}^N L(\mathbf{y}_n, \mathbf{x}_n; \Theta) \quad (2.18)$$

画像復元では MSE などに代表される画素値の間の距離を測ることができる損失関数がよく用いられる。劣化画像から無劣化画像への直接的なマッピングを MSE を用いて学習する場合、 $\Theta^*$  は以下の式によって与えられる。

$$\Theta^* = \underset{\Theta}{\operatorname{argmin}} \frac{1}{2N} \sum_{n=1}^N \|f(\mathbf{y}_n; \Theta) - \mathbf{x}_n\|_2^2 \quad (2.19)$$

テスト画像  $\mathbf{y}$  に対する復元画像  $\hat{\mathbf{x}}$  は、学習済みの深層学習モデル  $f(\cdot; \Theta^*)$  に  $\mathbf{y}$  を入力することによって得られる。観測劣化画像以外のデータを入力しない画像復元手法の場合、復元画像  $\hat{\mathbf{x}}$  は以下の式で得られる。

$$\hat{\mathbf{x}} = f(\mathbf{y}; \Theta^*) \quad (2.20)$$

深層学習モデルでは、モデルパラメータを学習によって調整し固定するため、比較的高速に推論を行うことができる。しかし、 $D_{\text{train}}$  における劣化画像  $\mathbf{y}_n$  が復元対象の劣化画像  $\mathbf{y}$  と大きく異なる場合、うまく復元することができない。

### 2.3.2 ノイズ除去におけるモデルアーキテクチャおよび学習手法

これまでに提案されたノイズ除去ネットワークのモデルアーキテクチャは、主に全結合層や畳み込み層、Batch normalization や Instance normalization などに代表される正規化層 [42, 43]、Sigmoid 関数や ReLU などの活性化関数 [44] を階層的に積み重ねることで構築されている。これに加えて、過学習を抑制するために、Dropout [45] などの正則化を導入している場合もある。Jain らの研究 [28] 以降、様々なモデルアーキテクチャを持つノイズ除去手法が提案されている。FFDNet [29] では、入力画像を 4 つのサブ画像に

ダウンサンプリングし深層学習モデルへ入力することで、高速なノイズ除去を実現している。また同時に、ノイズレベルマップを入力することで、ノイズ強度に幅がある場合やノイズが空間的に変化する場合でも柔軟にノイズ除去を行うことができる。

深層学習モデルの学習では、層数が増加するにつれて勾配爆発や勾配消失などの学習における問題が発生する可能性が高くなる。そのため、ResNet [46] で提案されている Skip connection を導入したモデルが提案されている [30, 31]。Skip connection によって、畳み込み層だけでは難しい恒等写像の学習（入力を変更せずに出力することを学習）が容易になる上、勾配消失を回避することができる。RED-Net [30] では、畳み込みと逆畳み込みからなる Encoder-Decoder 構造のノイズ除去モデルにおいて、Skip connection を導入している。さらに、Memnet [31] では Skip connection 構造を発展させ、現在の情報と過去の情報をどれだけ維持するべきかを判断する Gate unit を導入することによりモデルの表現能力を增強し、復元精度を向上させている。

また、特徴マップに対して明示的に重み付けを行う注意機構を持つモデルも提案されている [32, 47]。RIDNet [32] では、Short skip connection と Long skip connection を組み合わせた手法を提案している。この手法では、特徴マップに対して Global pooling, 畳み込み層, Sigmoid 関数による  $[0, 1]$  の正規化によって得られた重みを特徴マップに掛け合わせることで、それぞれの特徴マップに対して重み付けを行っている。

また、近年の CV 分野での Vision transformer [48, 49] の発展に伴い、画像復元においても Transformer をベースとした画像復元手法が提案されている [33]。Transformer 内で用いられている Self Attention(SA) では、計算量が空間分解能に対して 2 次関数的に増加するため、サイズの大きい画像に適用することができない。この問題に対し、Restormer [33] では、チャンネル間において SA を適用することで、計算量を抑えつつ長距離の画素相互作用を捉える手法を提案した。

初期の教師あり学習ベースの画像復元手法に対して定量的評価を行う場合、無劣化画像が必要となるため実験の多くはノイズの少ないデータセット [50, 51] に対し合成ノイズ (白色ガウシアンノイズ, ラプラシアンノイズ, ポアソンノイズなど) によって劣化させた画像を入力画像として画像復元モデルの精度を評価してきた。しかし、合成ノイズと実ノイズではノイズの分布に差があることから、現実のノイズ画像に対してうまく適応できないなどの問題があり、露光時間や ISO 感度を変更して撮像された実世界における劣化画像と無劣化画像のペアからなるデータセット [52-55] を活用することで、これらの問題に対処しようとしてきた [56, 57]。Kim らの手法 [56] では、合成ノイズで学習したモデルを転移学習することで、少量のラベル付きデータを用いた学習によってこの問題を解決しようとしている。CBDNet [57] では、ノイズ推定サブネットワークと非ブラインドノイ

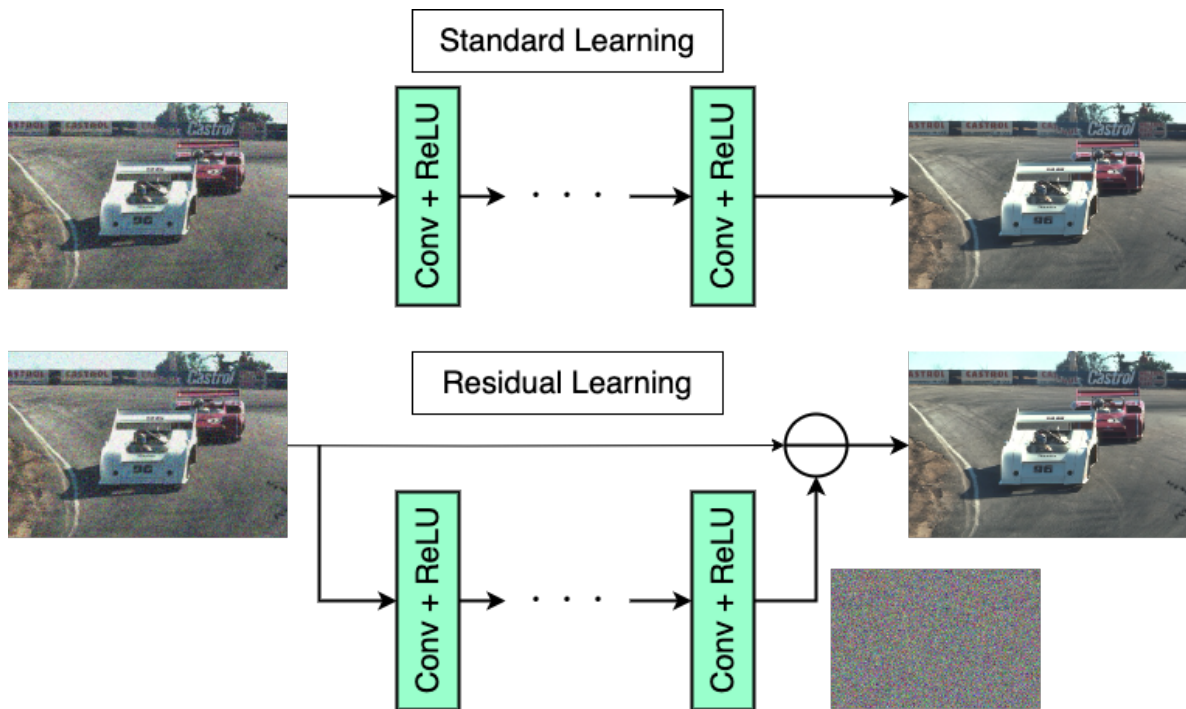


図 2.5: Residual learning モデル図

ズ除去器の2つのネットワークで構成される実ノイズ除去手法を提案している。

また、モデルアーキテクチャだけでなく、劣化画像から教師画像へのマッピングを行う一般的な教師あり学習ベース手法を発展させた手法がいくつか提案されている。DnCNN [34] によって提案された残差学習 (Residual learning) では、図 2.5 に示すように劣化画像から無劣化画像への直接的なマッピングを学習するのではなく、劣化画像からノイズを推定し、取り除くことで画像復元を行う。直接的なマッピングを行う式 (2.19) に対し、残差学習は以下の式で表すことができる。

$$\Theta^* = \operatorname{argmin}_{\Theta} \frac{1}{2N} \sum_{n=1}^N \|f(\mathbf{y}_n; \Theta) - (\mathbf{y}_n - \mathbf{x}_n)\|_F^2 \quad (2.21)$$

また、敵対的生成ネットワーク (GAN: Generative Adversarial Network) を使用したノイズ除去モデルも提案されている。Jingwen らの研究 [58] では、ノイズ生成器の学習とノイズ除去器の学習の2段階からなる手法が提案されている。劣化画像から生成したノイズパッチを学習し、ノイズパッチに類似するノイズを作成できるように Generator を訓練する。その後、無劣化画像に対し Generator によって生成されたノイズを付加することで劣化画像を作成し、劣化画像と無劣化画像のペアでモデルを学習する。

最適化ベースの手法に対して深層学習モデルを組み込むことができるような手法も幾つ

か提案されている [59–62]. 第 2.2 節でも説明したように, 最適化ベースの手法では特定の先験情報をモデル化しその近接写像を用いることで, 画像復元を行っていた. これに対し Plug-and-play では, 先験情報から近接写像を求めるのではなく, 近接写像を定めることによって対応する先験情報が定まると逆説的に考える. ADMM の正則化項には画像を滑らかにするような正則化項を指定することが多いため, これに対応する近接写像として学習済みノイズ除去器などを使用することができる.



## 第3章

# 事前学習を用いたハイパースペクトル画像復元

### 3.1 研究背景および関連手法

第 1.2 節で述べたように、深層学習を用いたハイパースペクトル画像 (HSI: Hyperspectral Image) 復元では、モデルの規模に対して学習データが不足することによる汎化性能の低下 (過学習)、および HSI が持つ膨大なバンド数によるモデルパラメータの増大が問題となる。HSI のバンド数に起因するパラメータ数の増大は、深層学習モデル全体を構築するための必要パラメータ数の増加に繋がるため、こちらも過学習を引き起こす原因となる。そのため、学習データセットを増加させることによる過学習の抑制、およびモデルアーキテクチャを変更することによる必要パラメータ数の削減が重要となる。本章では、事前学習による学習データの増強、および Separable convolution を用いた畳み込み処理における必要パラメータ数の削減を行う教師あり学習ベースの HSI ノイズ除去手法を提案する。

#### 3.1.1 事前学習

学習データの不足を原因とする過学習には、様々な対処方法が存在する。その中でも事前学習と呼ばれる手法では、大量に利用可能な別ドメインの大規模データセットを用いて学習し、実際に使用したいドメインのデータセットで調整 (Fine-tuning) することで、データの不足を補う。事前学習は特に自然言語処理分野においてよく利用されており、BERT [63] や GPT-3 [64] に代表される大規模な自然言語モデルで活用されている。

これらの手法では、マスクされた単語を予測する、次に出現する単語を予測するなどのラベルなしテキストデータを用いた自己教師あり学習によって、大規模データセットを用いた性能向上を可能にしている。

同様に画像に関するタスクにおいて、事前学習を用いた精度向上を行う手法がいくつか提案されている [65–67]。Chen らの研究 [65] では、低レベルコンピュータービジョンタスクで事前学習された Transformer ベース事前学習済みモデルを提供している。この事前学習済みモデルは、ImageNet データセット [68] から作成された合成データセットを用いて学習され、超解像、ノイズ除去、雨粒除去などのタスクで優れた精度を示すとともに、学習していない劣化に対しても優れた復元精度を示した。画像復元タスクではないものの、Kataoka らの研究 [66] ではクラス分類タスクにおいて、数式から人工的に生成した画像を用いて事前学習を行っている。ランダムに決定されたパラメータから作成された人工生成画像、およびラベルを用いて学習した場合に、同一ドメインの大規模データセットで学習した場合と遜色ない予測精度が得られることを実験的に示した。また、Brempong らの研究 [67] では、低レベルコンピュータービジョンタスクによる事前学習の有効性について検証している。セマンティックセグメンテーションのタスクにおける事前学習において、深層学習モデルの Decoder 部をノイズ除去タスクで学習した場合に、少ないラベル付きデータによる Fine-tuning でも優れた精度が得られることを実験によって示した。このように、小規模データセットしか利用することができない条件下において、別の大規模データセットを用いた深層学習モデルの事前学習により、精度を向上させることができる可能性が示唆されている。

### 3.1.2 Separable convolution

畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) における畳み込み層を複数の畳み込み処理に分割することで、畳み込み層 1 層あたりの必要パラメータ数を削減する試みが行われている [69, 70]。必要パラメータ数が減少することで、計算量の減少、推論の高速化などの利点を得られる。Separable Convolution [69] では通常の畳み込み層の処理を、空間方向の畳み込みとチャンネル方向の畳み込みに分離することで必要パラメータ数を削減する。通常の畳み込み処理のモデル図を図 3.1 に示す。図 3.1 に示すように、カーネルサイズ  $k \times k$ 、入力チャンネル数  $C_{in}$  の場合、通常の畳み込み層におけるバイアスを除く必要パラメータ数は 1 チャンネルあたり  $k^2 C_{in}$  である。出力チャンネル数を  $C_{out}$  とすると、畳み込み層 1 層あたりの必要パラメータ数はバイアスを除けば  $k^2 C_{in} C_{out}$  となる。



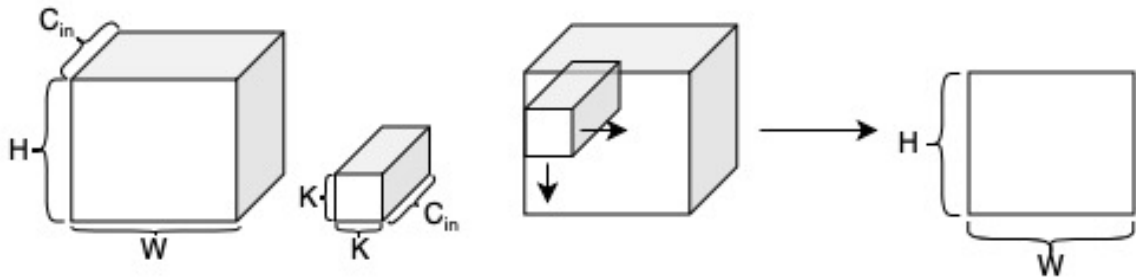


図 3.1: 通常の畳み込み層における処理のモデル図

これに対し, Separable convolution では, 図 3.2 に示す Depth-wise convolution と, 図 3.3 に示す Point-wise convolution の処理を組み合わせることで1つの畳み込み層の処理を実現する. Depth-wise convolution では図 3.2 のように, 入力の各チャンネルに対し

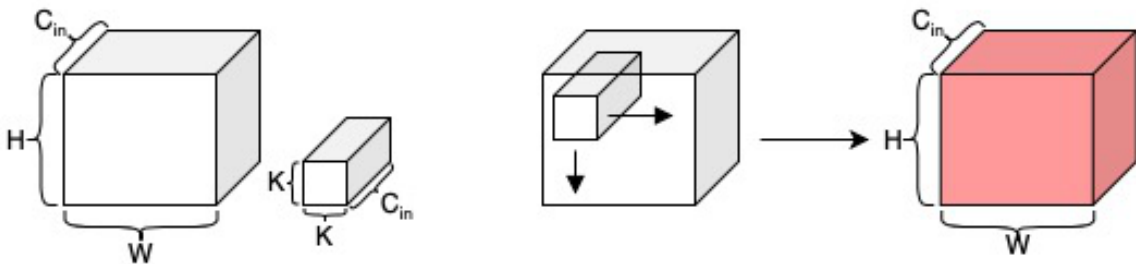


図 3.2: Depth-wise convolution における処理のモデル図

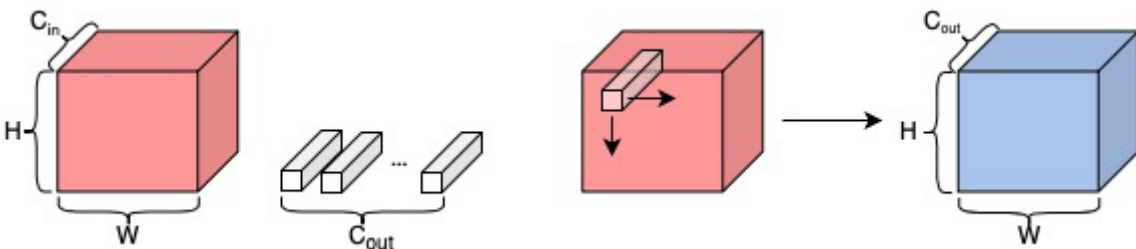


図 3.3: Point-wise convolution における処理のモデル図

て独立に畳み込み処理が行われる. そのため,  $k \times k \times C_{in}$  のサイズのカーネルを用いた畳み込みにより, 入力サイズと同サイズの出力を得る. その後, Depth-wise convolution

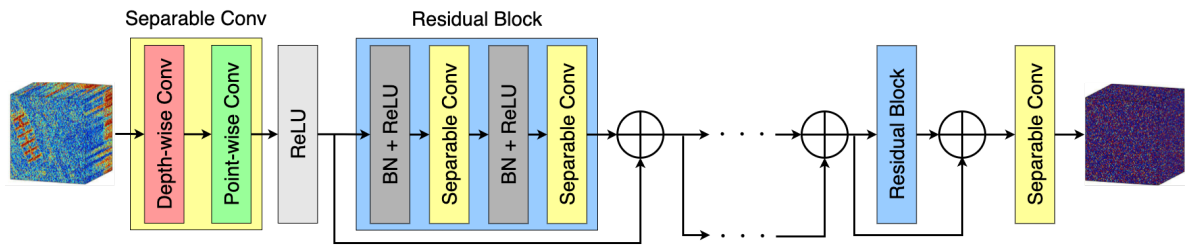


図 3.4: 提案手法におけるモデルアーキテクチャ

の出力に対して Point-wise convolution の処理を行う。

Point-wise convolution では図 3.3 のように、入力に対してカーネルサイズ  $1 \times 1 \times C_{in}$  のカーネルを用いた通常の畳み込み処理を行う。カーネルサイズが  $1 \times 1$  のため、チャンネル方向のみの情報を結合させる処理となる。Point-wise convolution における出力チャンネル数を  $C_{out}$  とすると、必要パラメータ数は  $C_{in}C_{out}$  となる。

結果として、バイアスを除いた場合の Separable convolution に必要なパラメータ数は、Depth-wise convolution では  $k^2C_{in}$ 、Point-wise convolution では  $C_{in}C_{out}$  となるため、合計で  $k^2C_{in} + C_{in}C_{out}$  となる。ここで、入力と出力のチャンネル数が同数、かつ  $C_{out} \gg k$  の場合、Separable convolution の必要パラメータ数は、通常の畳み込み層の約  $1/k^2$  となる。必要パラメータ数が大幅に減少することにより、少ないメモリ消費量で大規模なモデル作成が可能になり、モデルの表現力を高く維持しつつも過学習を抑制することが期待できる。

## 3.2 提案手法

### 3.2.1 モデルアーキテクチャ

第 3.1.2 小節で示したように、通常の畳み込み層を使用した場合、1 層につき  $k^2C_{in}C_{out}$  のパラメータが必要となる。特に HSI は通常数百ものバンドを有するため、チャンネル方向に密な結合を持つ畳み込みカーネルを用いる場合、中間層における出力チャンネル数の拡張にあたり膨大なパラメータが必要となる。第 2.3 節で説明している DnCNN [34] では、1 チャンネルまたは 3 チャンネルの入力画像に対して 64 枚のフィルタによる約 20 層の畳み込み層を有するが、HSI において DnCNN と同水準の倍率での出力チャンネル数の拡張は現実的ではない。第 3.1.2 小節で説明したように、Separable convolution では計算量が削減されると同時に、パラメータ数が大きく削減されるため、通常の畳み込みを使用した

ネットワークに比べて層の深いネットワークを構築することが可能となる。上記の理由から、提案手法では全ての畳み込み計算に Separable convolution を採用する。

本提案手法におけるモデルアーキテクチャを図 3.4 に示す。本提案手法は第 2.3 節で説明した、DnCNN をベースとしている。ネットワークは主に Separable convolution, ReLU [44], Batch normalization [42] からなる処理群によって構成され、Skip connection [46] によって接続される。まず、入力された HSI に対し Separable convolution, ReLU を用いて中間層の出力を生成する。得られた出力はその後に続く Residual block へ入力される。ここで、Residual block は図 3.4 中の青色の枠で示すように、Batch normalization, ReLU, および Separable convolution を二度繰り返すモジュールである、また、Skip connection によって、Residual block への入力とその Block の出力とが接続されており、二つの出力の和が次の層への入力となる。いくつかの Residual block を経た後に、Separable convolution を用いて出力チャンネル数を調節する。本提案手法ではすべての学習において、第 2.3.2 小節で説明した手法である、劣化 HSI からノイズを予測する残差学習 (Residual learning) を使用している。そのため、入力 HSI と深層学習モデルが予測したノイズの差を取ることで、復元された HSI を得る。

### 3.2.2 提案手法における事前学習手順

本手法では、Separable convolution における Depth-wise convolution と Point-wise convolution のそれぞれのカーネルが空間方向及びチャンネル方向で独立していることを利用した事前学習手法を提案する。トレーニングの流れを図 3.5 に示す。HSI が持つ空間方



図 3.5: 提案手法における学習手順

向の特徴はどのバンドもおおよそ類似しており、特定のバンドに対して学習されたカーネルは別のバンドに対しても有効であると考えられる。すなわち、HSI における特定のバンドをグレースケール画像とみなし、学習したカーネルは学習していない他のバンドに対しても適用可能であると考えることが出来る。その場合、Depth-wise convolution カーネルの学習にはデータの取得が容易なグレースケール画像を用いればよく、タスクの難度を

下げることが可能となる。HSI で取得されるバンド情報は撮影対象の物質が大きく影響する。バンド方向の特徴とみなせる物質のスペクトルデータも大量に用意することが出来るため、これらを用いて Separable convolution における Depth-wise convolution カーネル  $\mathbf{W}_d$  と Point-wise convolution カーネル  $\mathbf{W}_p$  に対し個別の事前学習を行うことでデータ不足を補い、汎化性能を向上させる。

### 3.2.3 カーネル拡張

HSI 復元におけるグレースケール画像、および RGB 画像を用いた事前学習では、保有バンド数が大きく異なることから、深層学習モデルのカーネルの形状を一意に決定できない問題点がある。例えば、RGB 画像を用いて深層学習モデルを事前学習する場合、RGB が保有するバンド数は 3 であるため、入力層に畳み込み層のを用いた場合のカーネルの形状は、 $k \times k \times 3 \times C_{out}$  となる。これに対し、バンド数  $C_{in}$  の HSI を事前学習済みモデルに入力する場合、深層学習モデルのチャンネル数と入力する HSI のバンド数が合致しないため、出力を計算することができない。また、3 チャンネルから  $C_{out}$  チャンネルへの中間特徴量の拡張と、 $C_{in}$  チャンネルから  $C_{out}$  チャンネルへの中間特徴量の拡張では、情報の冗長性が大きく異なる。

本提案手法では、グレースケール画像を用いて学習した Depth-wise convolution カーネル  $\mathbf{W}_{pre} \in \mathbb{R}^{k \times k \times l}$  を、カーネル拡張を用いて  $\mathbf{W}_d \in \mathbb{R}^{k \times k \times l \cdot B}$  に拡張する。ここで、 $B$  は Fine-tuning に用いる HSI のバンド数である。カーネル拡張のモデル図を図 3.6 に示す。カーネル拡張では、Depth-wise convolution カーネルの  $i$  チャンネル目を、チャンネル

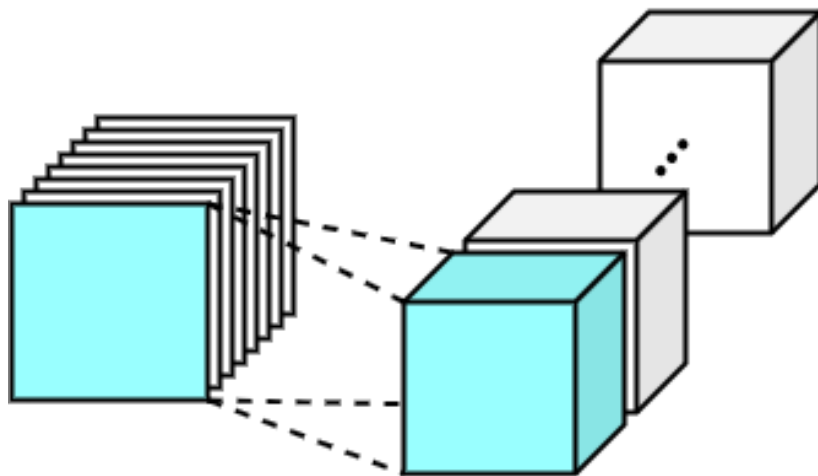


図 3.6: カーネル拡張手順

方向にコピーすることで  $B$  チャンネルの 3 次元カーネルに拡張する。この操作をすべてのチャンネルに対してそれぞれ行うことで、 $\mathbf{W}_d$  が得られる。カーネル拡張は、グレースケール画像で学習した各カーネルを、HSI のすべてのチャンネルに対してそれぞれ適用できるようにするために行われる。

### 3.2.4 Point-wise convolution カーネルの事前学習

Point-wise convolution カーネルの学習では、Depth-wise convolution カーネルが学習によって更新されないよう固定する。チャンネル方向の事前学習には 1 次元スペクトルデータセットである USGS Library<sup>\*1</sup>を用いる。このデータセットには 224 バンドのスペクトルデータが含まれている。このデータに対しランダムに重み付けを行って線形和をとることで擬似スペクトルベクトルを合成し、学習に使用する。この際、これらのベクトルの値域が  $[0, 1]$  となるように正規化を行う。学習の際は、作成した擬似スペクトルベクトルから連続する  $B$  個の要素をランダムに抽出し、抽出されたベクトルを  $H \times W \times B$  のサイズになるように空間方向にコピーする。ここで、 $H$ ,  $W$ ,  $B$  はそれぞれ Fine-tuning の際に用いるハイパススペクトル画像の縦幅、横幅、バンド数である。この操作によって空間方向の値はバンドごとに同じ値となり、Depth-wise Convolution による空間方向への畳み込みの影響を抑制することが出来る。最後に、学習された Depth-wise convolution カーネルと Point-wise convolution カーネルを初期値とし、HSI を用いて Fine-tuning を行う。

## 3.3 実験

### 3.3.1 実験設定

観測されるノイズの標準偏差は既知であるとし、標準偏差  $\sigma = 0.2, 0.3, 0.4$  の白色ガウシアンノイズ除去実験を行なった。劣化 HSI は式 (2.2) に従って与えられる。深層学習モデルにおける Residual block の数は 10 とし、ミニバッチサイズは全ての学習において共通で 32 とした。空間方向の事前学習に用いたグレースケール画像の数は 110 枚であり、パッチサイズを  $40 \times 40 \times 1$  としてランダムクロップ、反転および回転によって Data augmentation を行っている。実験で用いる HSI のデータセットに航空写真が多く含まれていることから、ドメインを近づける為に Web サイトから取得した航空写真をグレース

---

<sup>\*1</sup> <https://www.usgs.gov/core-science-systems/usgs-library/>

ケール画像に変換したデータセットを作成し学習に使用する。また、このネットワークでは中間層の出力チャンネル数を  $l = 10$  と設定した。最適化手法として Adam [71] を採用し、学習率  $1.0 \cdot 10^{-3}$  で 48 エポック、その後、学習率  $1.0 \cdot 10^{-4}$  で 2 エポック事前学習を行なった。

チャンネル方向の事前学習用に作成した擬似スペクトルデータベクトルの数は 20000 パターンであり、連続する  $B = 50$  要素をランダムに抽出、その  $1 \times 1 \times 50$  のベクトルを空間方向にコピーし  $40 \times 40 \times 50$  のパッチを作成し学習に使用した。 $B, l$  はここではそれぞれ  $B = 50, l = 10$  であるため、カーネル拡張により中間層の出力チャンネル数は 500 となる。最適化手法として Adam を採用し、学習率  $1.0 \cdot 10^{-3}$  で 7 エポック、その後、学習率  $1.0 \cdot 10^{-4}$  で 3 エポック事前学習を行なった。

Fine-tuning の際に使用した HSI の枚数は 19~20 枚で、パッチサイズを  $40 \times 40 \times 50$  とし、トレーニングデータセットを作成した。最適化手法として Adam を採用し、学習率  $1.0 \cdot 10^{-3}$  で 16 エポック、その後、学習率  $1.0 \cdot 10^{-4}$  で 4 エポック学習を行なった。

テストデータには、劣化したチャンネルを取り除いた WashingtonDC, Cuprite-cor, Suwannee-cor, PaviaC, PaviaU, Frisco, Salinas, Dioni を使用し、表中に記載されている Average は以上の 8 枚の画像の結果を平均したものである。テストデータはあらかじめ  $[0, 1]$  に正規化し、切り出したものを使用した。また、先行研究のうち、ASTV [21], SSTV [23], HSSTV [25], BM4D [72], LRTDTV [37] を比較手法として使用した。ここで、ASTV, SSTV, HSSTV, LRTDTV は HSI のスパース性や低ランク性を考慮したモデルベースの復元手法であり、BM4D は BM3D を三次元画像に拡張したノンローカルベースのノイズ除去手法である。これらの性能評価指標として Peak Signal-to-Noise Ratio (PSNR) を用いる。

### 3.3.2 ガウシアンノイズ除去実験

評価画像のうち 6 枚の実験結果と全評価画像の結果の平均 PSNR を、表 3.1, 表 3.2, 表 3.3 に示す。ASTV, SSTV, HSSTV, BM4D, LRTDTV と比較し、提案手法は平均的に見た場合、従来手法を上回る性能が得られた。また、ノイズ強度が高くなるほど従来手法との性能差が大きくなっており、本手法は劣化が激しい画像に対してより有効になることがわかる。

Frisco データ, PaviaC データに対するノイズ除去結果を図 3.7, 図 3.8, 図 3.9, 図 3.10, 図 3.11, 図 3.12 に示す。

表 3.1: 標準偏差  $\sigma = 0.2$  におけるガウシアンノイズ除去結果

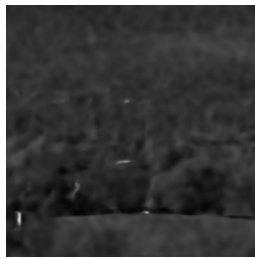
Images	ASTV	SSTV	HSSTV	BM4D	LRTDTV	Ours
WashingtonDC	26.66	30.78	32.52	31.85	<b>33.70</b>	31.88
Cuprite	28.65	28.52	32.08	31.87	33.13	<b>33.33</b>
Suwannee	30.01	30.04	<b>33.86</b>	33.60	32.68	32.12
PaviaU	26.78	29.57	30.83	31.17	30.41	<b>35.43</b>
Frisco	30.88	30.32	34.28	33.38	33.64	<b>35.67</b>
Dioni	31.71	30.16	34.75	33.62	33.24	<b>36.00</b>
Overall average	28.90	30.07	33.12	32.64	32.58	<b>34.13</b>

表 3.2: 標準偏差  $\sigma = 0.3$  におけるガウシアンノイズ除去結果

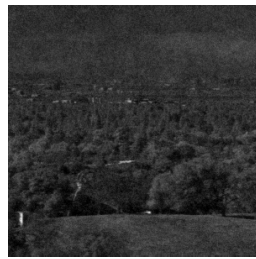
Images	ASTV	SSTV	HSSTV	BM4D	LRTDTV	Ours
WashingtonDC	24.46	28.10	<b>30.80</b>	29.93	30.02	30.78
Cuprite	26.91	25.61	30.66	30.70	30.09	<b>32.16</b>
Suwannee	28.45	27.36	<b>31.97</b>	31.54	28.69	30.83
PaviaU	25.63	26.98	28.99	29.12	27.13	<b>33.48</b>
Frisco	29.47	27.57	32.43	31.45	29.97	<b>33.41</b>
Dioni	30.10	27.42	32.93	31.86	29.37	<b>34.95</b>
Overall average	27.81	27.26	31.48	30.93	29.29	<b>32.70</b>

表 3.3: 標準偏差  $\sigma = 0.4$  におけるガウシアンノイズ除去結果

Images	ASTV	SSTV	HSSTV	BM4D	LRTDTV	Ours
WashingtonDC	22.79	26.21	29.51	28.64	27.48	<b>30.00</b>
Cuprite	25.70	23.52	29.62	29.86	27.76	<b>31.51</b>
Suwannee	27.39	25.58	30.48	30.11	26.04	<b>31.05</b>
PaviaU	24.81	25.15	27.54	27.80	24.67	<b>31.90</b>
Frisco	28.36	25.70	31.15	30.11	27.41	<b>31.89</b>
Dioni	28.90	25.57	31.65	30.57	26.69	<b>34.29</b>
Overall average	26.61	25.28	30.17	29.56	26.74	<b>31.90</b>



(a) ASTV



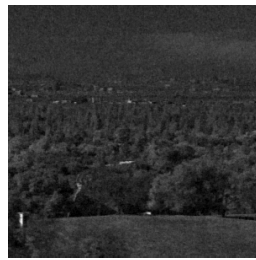
(b) SSTV



(c) HSSTV



(d) BM4D



(e) LRTDTV



(f) Ours



(g) オリジナル画像



(h) ノイジー画像

図 3.7: 標準偏差  $\sigma = 0.2$  における Frisco データに対するノイズ除去結果画像





(a) ASTV



(b) SSTV



(c) HSSTV



(d) BM4D



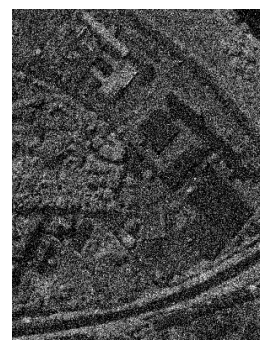
(e) LRTDTV



(f) Ours



(g) オリジナル画像



(h) ノイジー画像

図 3.8: 標準偏差  $\sigma = 0.2$  における PaviaC データに対するノイズ除去結果画像

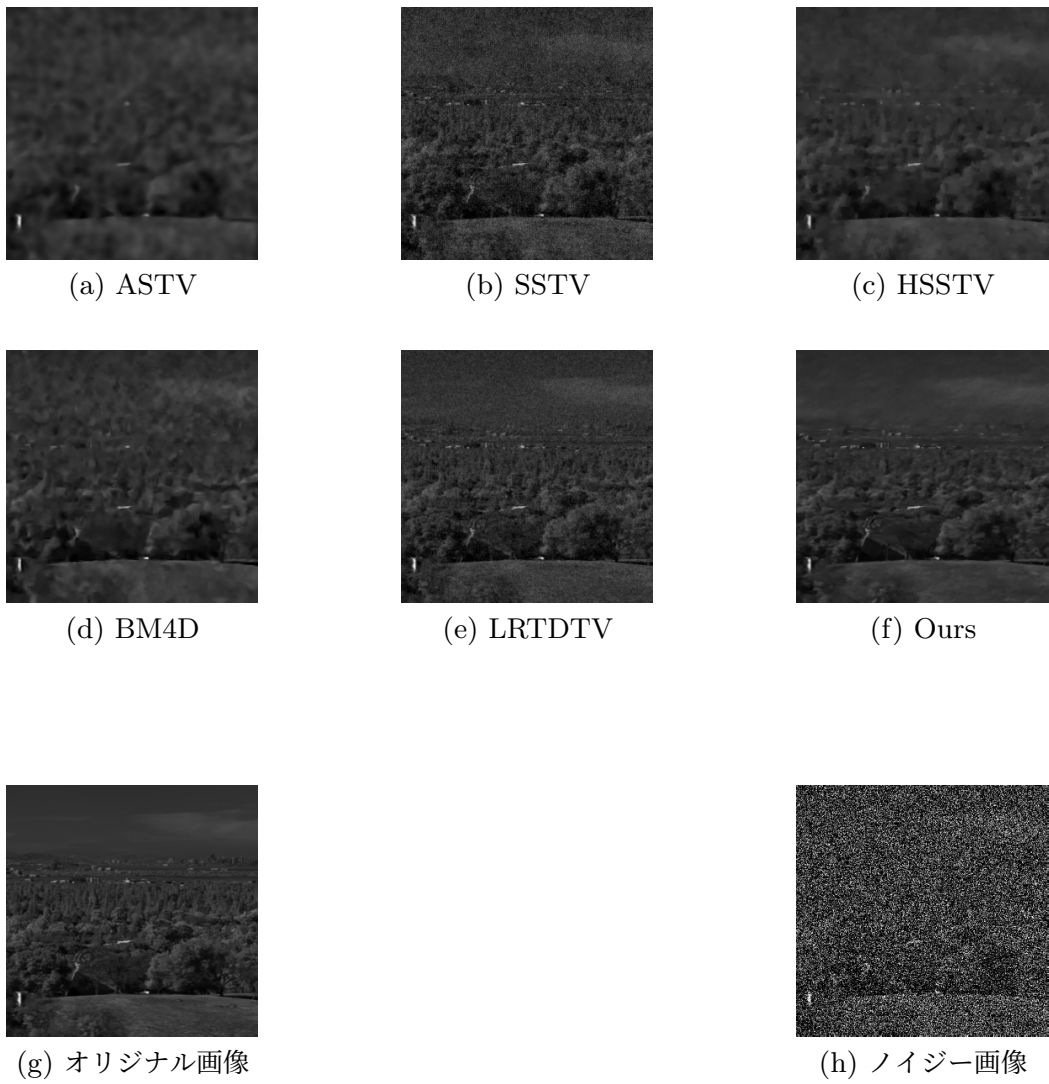


図 3.9: 標準偏差  $\sigma = 0.3$  における Frisco データに対するノイズ除去結果画像



(a) ASTV



(b) SSTV



(c) HSSTV



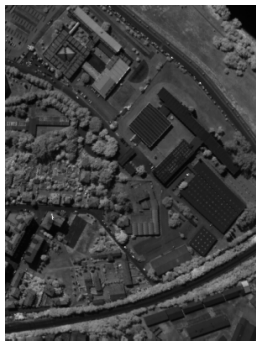
(d) BM4D



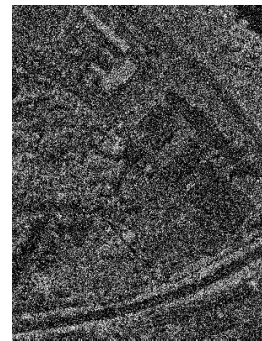
(e) LRTDTV



(f) Ours



(g) オリジナル画像

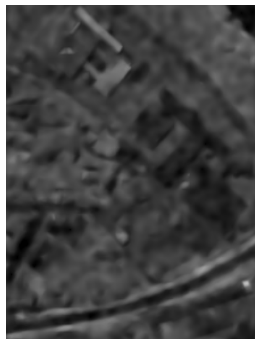


(h) ノイジー画像

図 3.10: 標準偏差  $\sigma = 0.3$  における PaviaC データに対するノイズ除去結果画像



図 3.11: 標準偏差  $\sigma = 0.4$  における Frisco データに対するノイズ除去結果画像



(a) ASTV



(b) SSTV



(c) HSSTV



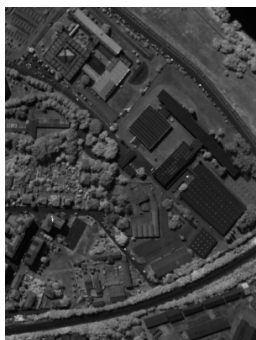
(d) BM4D



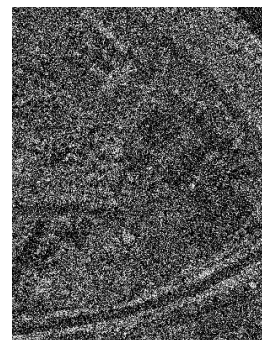
(e) LRTDTV



(f) Ours



(g) オリジナル画像



(h) ノイジー画像

図 3.12: 標準偏差  $\sigma = 0.4$  における PaviaC データに対するノイズ除去結果画像

すべての標準偏差において、ASTV では画像全体がぼやける傾向にあり、SSTV、LRTDTV は標準偏差が上がるにつれてノイズが残る傾向にあった。HSSTV は過平滑された部分が組み合わされたような見た目になってしまっている。BM4D が最も従来手法の中で違和感がないような画像になっているが、本提案手法の方がより自然な画像になっている。

このような結果になったことへの考察として、従来手法は入力が高ランクかつスパースになるような尤もらしい画像を出力するが、ノイズ強度が大きくなってくると、エッジが増え、低ランクな解の選択肢が広がるために最適解に近い解を得るのが難しくなる。残差学習ベースの手法の場合、学習したデータから予測されたノイズ強度のノイズを出力とするので、ノイズ強度に関係なく良い結果が得やすい。すなわち従来手法に比べて最適解に近い解を得やすい。このため、従来手法よりも良い結果が得られていると考えられる。

また、データが十分にある場合、ノイズの分布が全て似通うことは稀であるため、様々な画素値にノイズを加えた分布を満遍なく学習することが可能である。そのため、従来手法に比べて良い結果が得られる。しかし、本提案手法で扱った HSI のようにデータが少ない場合、得られる元の画素値の分布に限られるため、ノイズを加えた後の画素値の分布も限られた分布となる。元のデータの分布に限られているため、Data augmentation を行なったとしても、元の画素値の分布は大きく変わらず、それほど大きな効果は得られない。Fine-tuning で空間方向とチャンネル方向を学習させることで、各重みが様々な分布のデータを学習することができ、汎化性能が向上する。標準偏差が高いと元の画素値から大きく離れるため、ノイズのない元の画像の画素値を推測するのに必要なデータの数が多くなる。そのため、ノイズを推測する今回のネットワークでも同様に良い結果が得づらい。Fine-tuning による事前学習は増加した必要なデータ数を補っていると考えられるため、性能が向上しているとも考えられる。

## 3.4 結論

### 3.4.1 本章のまとめ

本研究では、Separable convolution における空間方向のカーネルとチャンネル方向のカーネルをそれぞれ独立に事前学習する HSI ノイズ除去手法を提案した。Separable convolution の採用により、畳み込み層 1 層あたりの必要パラメータ数が削減され、数百ものバンドを有する HSI に対しても比較的規模の大きいネットワークの学習が可能となった。また、Separable convolution の畳み込みカーネルが空間方向の畳み込みである

Depth-wise convolution と、チャンネル方向の畳み込みである Point-wise convolution に分離されていることに着目し、Depth-wise convolution カーネルをグレースケール画像、Point-wise convolution カーネルを合成した擬似スペクトルデータで学習する。従来手法との比較では、平均的に見た場合に提案手法のノイズ除去性能の方が優れており、ノイズ強度が高いほど従来手法よりも提案手法の方が良い結果が得られる傾向にあった。

### 3.4.2 提案手法の課題

事前学習に使用するグレースケール画像と HSI が大きく異なる場合、事前学習による精度向上が限定的になる可能性がある。特に復元対象の HSI の撮像対象が事前にわかっていることは少なく、HSI ごとにデータセットを用意する場合、コストが膨大になってしまう。また、合成によって作成される擬似スペクトルデータはランダムな重みづけによって作成されているため、実際の物質のスペクトルデータに即していない可能性がある。これらの理由から、事前学習で使用するデータセットの質、および量に学習済みモデルの精度が依存してしまう可能性があるため、得られる学習済みモデルの精度が安定しない問題点がある。





## 第4章

# 深層学習モデルと深層展開モデルを組み合わせたグレースケール画像復元

### 4.1 研究背景および関連手法

ハイパースペクトル画像 (HSI: Hyperspectral Image) のような大規模データセットを利用することが困難な条件下における信号復元タスクでは、先験情報をもとに設計した正則化項を用いた最適化ベースの手法を採用することが望ましい場合が多い。正則化項が適切に復元対象の信号をモデリングすることができていれば、観測劣化信号のみから精度良く画像復元が可能である。また、第2.2節に示すように、凸最適化では局所解と大域最適解が一致することから深層学習モデルと比較して初期値への依存性が低くなる。

多くの反復アルゴリズムでは、アルゴリズムの振る舞いを制御するためのハイパーパラメータを持つ。これらのハイパーパラメータは出力に大きく影響する上、通常はヒューリスティックに決定する必要がある。これらの反復アルゴリズムを深層学習技術を用いて拡張することで、ハイパーパラメータを学習用データセットを用いて調整する深層展開 (Deep unfolding, Deep unrolling) と呼ばれる手法がこれまでに提案されている。現在までに、いくつかの深層展開ベースの画像復元手法が提案されている [73–78]。しかし、深層展開の概念を適用した先行研究の多くは、モデルの性能と解釈可能性の向上に焦点を当てたものであり、小規模データセット条件下での汎化性能を考慮したものはごく少数である。

本章では、最適化ベースの手法が持つ学習データなしで精度良く復元が可能である特性

を活用した新たな画像復元手法を提案する。画像復元モデルにおける出力部を深層展開モデルとすることにより、モデル全体の出力を与えられた正則化項に従って復元するように制約する。これに加えて、画像復元モデル全体の表現力の向上と、深層展開モデルの前処理のために、深層学習モデルを連結する。小規模な学習用データセット条件下において、深層学習モデルのみの場合と比較して提案画像復元モデルの汎化性能が優れていることを実験的に示す。

#### 4.1.1 深層展開 (Deep unfolding, Deep unrolling)

深層展開は、反復アルゴリズムの内部パラメータを深層学習技術を用いて学習できるように拡張した手法である。深層展開のモデル図を図 4.1 に示す。通常の反復アルゴリズムでは、計算の開始から終了まで同一か、特定の更新式に基づいて変更されるハイパーパラメータを使用する。定められた収束条件を満たすまで計算を繰り返し、満たした時点で反復アルゴリズムを終了する。これに対し、深層展開によって拡張された反復アルゴリズムでは、反復アルゴリズムの一連の計算を 1 ステップとし、各ステップにおけるハイパーパラメータを学習によって決定する。これにより、性能の向上や収束の高速化などのメリットが得られる。学習するハイパーパラメータは事前に定義する必要があるため、それらの定義に必要な反復アルゴリズムの最大反復回数  $T$  を前もって指定する必要がある。

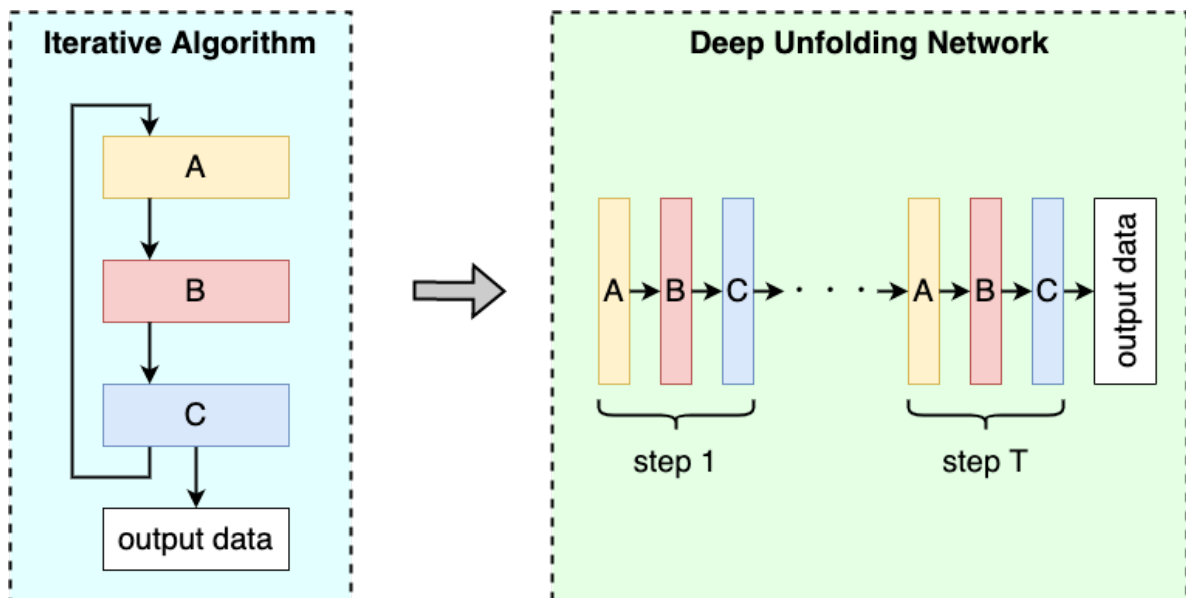


図 4.1: 深層展開による反復アルゴリズムの拡張モデル図

### 4.1.2 深層展開によって拡張された Anisotropic TV

第 2.2 節で説明した Anisotropic TV の式 (2.11), 式 (2.12), 式 (2.13) を, 深層展開を用いて拡張すると以下ようになる.

$$\hat{\mathbf{x}}^{(t+1)} = \operatorname{argmin}_{\hat{\mathbf{x}}} \frac{1}{2} \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|_2^2 + \frac{1}{2\gamma_t} \|\mathbf{z}^{(t)} - \mathbf{D}\hat{\mathbf{x}} - \mathbf{b}^{(t)}\|_2^2 \quad (4.1)$$

$$\mathbf{z}^{(t+1)} = \operatorname{argmin}_{\mathbf{z}} \lambda_t \|\mathbf{z}\|_1 + \frac{1}{2\gamma_t} \|\mathbf{z} - \mathbf{D}\hat{\mathbf{x}}^{(t+1)} - \mathbf{b}^{(t)}\|_2^2 \quad (4.2)$$

$$\mathbf{b}^{(t+1)} = \mathbf{b} + \mathbf{D}\hat{\mathbf{x}}^{(t+1)} - \mathbf{z}^{(t+1)} \quad (4.3)$$

ここで,  $\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_T\}$ ,  $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_T\}$  である. 式 (2.11), 式 (2.12), 式 (2.13) とは異なり, ハイパーパラメータ  $\lambda$  と  $\gamma$  がステップごとに異なる値を使用できるように拡張されている. これにより, 最適化の進行に合わせて正則化の強度やステップ幅を調整することができ, それらを学習データから自動で決定することが期待できる.

## 4.2 提案手法

本手法では, 凸最適化手法を深層展開を用いて拡張したモデル (DUM: Deep Unfolding Model) と前処理のための深層学習モデル (DLM: Deep Learning Model) の 2 つを組み合わせ合わせた画像復元モデルを提案する. DUM は凸最適化の手法をベースとしているため, 与えられた先験情報をもとに画像復元を行う. 具体的には, DUM は, 与えられた事前情報に基づいて, 正則化を満たす出力を生成するように画像復元モデル全体の出力を制約する. さらに, DLM は, DUM が復元しやすくなるような観測画像の変換を学習することで, DUM の画像復元精度の向上に寄与する.

提案手法全体のモデルアーキテクチャを図 4.2 に示す. ここでは, 基本的な畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) と, Anisotropic TV

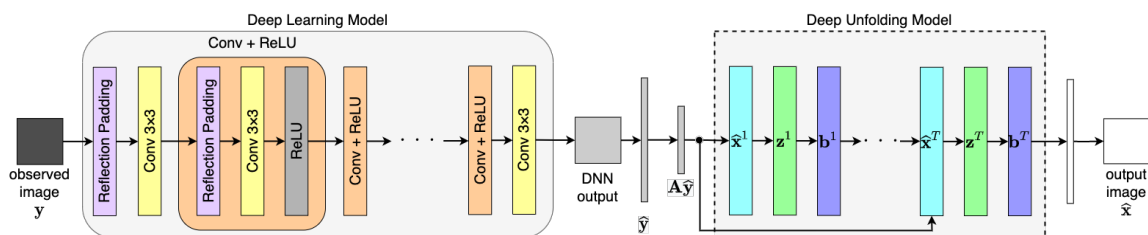


図 4.2: 提案手法 モデルアーキテクチャ

ベースの最適化を深層展開のフレームワークを用いて拡張したモデルをベースとした画像復元モデルについて説明する．提案する画像復元モデルでは，DLM と DUM が直列に連結されている．DLM は観測画像を  $\hat{\mathbf{y}}$  の画像に変換し， $\mathbf{A}\hat{\mathbf{y}}$  が DUM への入力となる．

深層学習モデルによる前処理を組み込んだ反復アルゴリズムは以下のようになっている．復元画像  $\hat{\mathbf{x}}$  を求める式は以下の通りである．

$$\hat{\mathbf{x}}^{(t+1)} = \operatorname{argmin}_{\hat{\mathbf{x}}} \frac{1}{2} \|\mathbf{A}\hat{\mathbf{x}} - f(\mathbf{y}; \Theta)\|_2^2 + \frac{1}{2\gamma_t} \|\mathbf{z}^{(t)} - \mathbf{D}\hat{\mathbf{x}} - \mathbf{b}^{(t)}\|_2^2 \quad (4.4)$$

$$\mathbf{z}^{(t+1)} = \operatorname{argmin}_{\mathbf{z}} \lambda_t \|\mathbf{z}\|_1 + \frac{1}{2\gamma_t} \|\mathbf{z} - \mathbf{D}\hat{\mathbf{x}}^{(t+1)} - \mathbf{b}^{(t)}\|_2^2 \quad (4.5)$$

$$\mathbf{b}^{(t+1)} = \mathbf{b} + \mathbf{D}\hat{\mathbf{x}}^{(t+1)} - \mathbf{z}^{(t+1)} \quad (4.6)$$

損失関数には MSE が用いられる．MSE は DLM と DUM の学習においてエンド・ツー・エンドで適用される．ここで  $f(\mathbf{y}; \Theta)$  は DLM の出力， $\gamma_t$  と  $\lambda_t$  は学習可能なパラメータである．

#### 4.2.1 Deep unfolding モデル (DUM)

DUM のモデルアーキテクチャは図 4.3 のようになっている．DUM では通常の反復ア

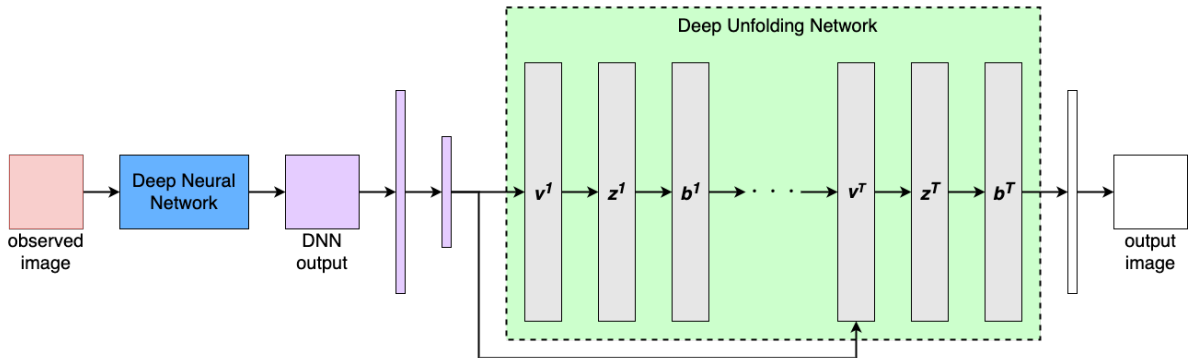


図 4.3: Deep unfolding モデル (DUM) のモデルアーキテクチャ

ルゴリズムのように式 (4.4)，式 (4.5)，式 (4.6) を収束条件を満たすまで繰り返すのではなく，指定された最大反復数である  $T$  回繰り返す．DUM のベースとなっている反復アルゴリズムは，Anisotropic TV を用いた画像復元を行なっているため，画像をなめらかにするように最適化による画像復元を行う．そのため，DUM の出力は，画像をなめらかにするように制約される．

## 4.2.2 Deep learning モデル (DLM)

DLM のモデルアーキテクチャは図 4.4 のようになっている。DLM はリフレクション

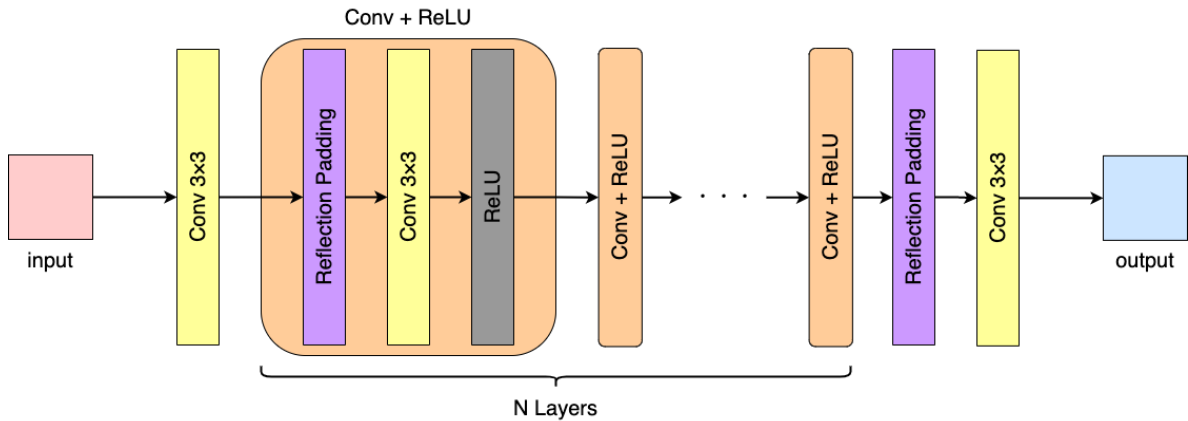


図 4.4: Deep learning モデル (DLM) のモデルアーキテクチャ

パディングを含む、 $3 \times 3$  の畳み込み層と ReLU で構成されるベーシックな CNN を使用する。DLM は、DUM の復元性能を向上させるために、観測劣化画像  $\mathbf{y}$  から  $\hat{\mathbf{y}}$  への変換を学習する。この前処理の目的は、単純に入力画像と教師画像の差を最小化することではなく、DUM で使われる正則化に適した変換を学習することである。この処理の詳細は第 4.3.4 小節で述べる。

## 4.3 実験

### 4.3.1 実験設定

本節では、提案する画像復元モデルが、小規模データセットに適用された場合に、汎化性能を向上させることを実証するためにノイズ除去タスクとピクセル補間タスクを行う。訓練画像の枚数  $N_{\text{train}}$  は 5, 25, 50 枚、テスト画像の枚数  $N_{\text{test}}$  は 50 枚とした。どちらのデータセットも STL10 データセット [79] からランダムにサンプリングして作成した。全ての画像は  $[0, 1]$  に正規化されている。評価指標は Peak Signal-to-Noise Ratio (PSNR), および Structural Similarity Index [80] (SSIM) を使用した。

DLM の中間層のチャンネル数は 64, 層数は 5, DUM における最大ステップ数  $T$  は 25 とした。全てのモデルにおいて、エポック数は 150, 初期学習率は  $1.0 \cdot 10^{-3}$  とした。以下

の実験結果では、DLM と DUM を単独で用いてデータを学習させた場合と、提案する画像復元モデル (Ours) を用いてデータを学習させた場合の結果について述べる。以下、わかりやすくするために、Deep learning モデルと Deep unfolding モデルのどちらかを単独で用いて学習を行った場合を、それぞれ Stand-alone DLM (SA-DLM), Stand-alone DUM (SA-DUM) と呼ぶ。SA-DUM の入力は、観測劣化画像  $\mathbf{y}$  をベクトル化し、行列  $\mathbf{A}$  を乗算したベクトル  $\mathbf{A}\hat{\mathbf{y}}$  とする。

ノイズ除去実験では、標準偏差  $\sigma = 0.15, 0.25$  の白色ガウシアンノイズおよび白色ラプラシアンノイズ除去実験を行った。劣化画像は式 (2.2) に従って与えられる。その後、提案画像復元モデルの性能をさらに検証するため、ピクセル補間実験を行った。劣化画像は式 (2.1) に従って与えられる。行列  $\mathbf{A}$  は各設定に対して固定した。式 (2.1) における  $\mathbf{A}$  は、所定の画素値を 0 とするような 0 と 1 からなるサンプリング行列とした。この行列では無作為に 20% の画素を選んで 0 とするように設定している。また、 $\mathbf{n}$  として標準偏差  $\sigma = 0.2$  のガウシアンノイズを使用している。

#### 4.3.2 ガウシアンノイズ，およびラプラシアンノイズ除去実験

ガウシアンノイズ除去結果を表 4.1，ラプラシアンノイズ除去結果を表 4.2 に示す。表 4.1，および表 4.2 の両方から、学習画像の枚数を 25 枚，50 枚と増やしても Stand-alone DUM の性能に大きな改善は見られないことがわかる。これは、学習可能なパラメータの数が非常に少ないために、モデル自体の表現力が低く、アンダーフィッティングしているためであると考えられる。なお、従来の ADMM で達成可能な性能は、Stand-alone DUM に適用した条件下で達成可能な性能に劣る。さらに、Stand-alone DUM よりも多くのパ

表 4.1: ガウシアンノイズ除去結果

$N_{\text{train}}$		5		25		50	
$\sigma$	Methods	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
0.15	SA-DUM	23.14	$7.397 \cdot 10^{-1}$	23.26	$7.521 \cdot 10^{-1}$	23.16	$7.421 \cdot 10^{-1}$
	SA-DLM	<b>23.42</b>	$7.728 \cdot 10^{-1}$	<b>24.58</b>	<b><math>8.111 \cdot 10^{-1}</math></b>	<b>24.87</b>	$8.157 \cdot 10^{-1}$
	Ours	22.41	<b><math>7.739 \cdot 10^{-1}</math></b>	23.82	$8.111 \cdot 10^{-1}$	24.84	<b><math>8.168 \cdot 10^{-1}</math></b>
0.25	SA-DUM	<b>21.86</b>	<b><math>6.848 \cdot 10^{-1}</math></b>	21.83	$6.861 \cdot 10^{-1}$	21.74	$6.778 \cdot 10^{-1}$
	SA-DLM	20.95	$6.468 \cdot 10^{-1}$	22.12	$7.106 \cdot 10^{-1}$	<b>22.48</b>	$7.190 \cdot 10^{-1}$
	Ours	21.24	$6.644 \cdot 10^{-1}$	<b>22.34</b>	<b><math>7.192 \cdot 10^{-1}</math></b>	22.36	<b><math>7.293 \cdot 10^{-1}</math></b>

表 4.2: ラプラシアンノイズ除去結果

$N_{\text{train}}$		5		25		50	
$\sigma$	Methods	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
0.15	SA-DUM	22.21	$7.006 \cdot 10^{-1}$	22.25	$7.075 \cdot 10^{-1}$	22.30	$7.074 \cdot 10^{-1}$
	SA-DLM	22.14	$7.054 \cdot 10^{-1}$	23.17	$7.565 \cdot 10^{-1}$	23.51	$7.678 \cdot 10^{-1}$
	Ours	<b>22.49</b>	<b><math>7.252 \cdot 10^{-1}</math></b>	<b>23.33</b>	<b><math>7.626 \cdot 10^{-1}</math></b>	<b>23.64</b>	<b><math>7.741 \cdot 10^{-1}</math></b>
0.25	SA-DUM	<b>20.68</b>	<b><math>6.190 \cdot 10^{-1}</math></b>	20.91	$6.390 \cdot 10^{-1}$	20.78	$6.287 \cdot 10^{-1}$
	SA-DLM	19.95	$5.801 \cdot 10^{-1}$	21.14	$6.493 \cdot 10^{-1}$	21.45	$6.670 \cdot 10^{-1}$
	Ours	20.28	$5.998 \cdot 10^{-1}$	<b>21.37</b>	<b><math>6.695 \cdot 10^{-1}</math></b>	<b>21.49</b>	<b><math>6.762 \cdot 10^{-1}</math></b>

ラメータを有する Stand-alone DLM, および提案された画像復元モデルでは, データが増加するに従って, より良い性能を発揮する. しかし, 一般的に劣化が激しい (標準偏差  $\sigma$  の値が大きい) ほど, 過学習の可能性が高くなり汎化性能が低下する傾向がある. 本提案手法のように, モデル全体の出力が事前情報を満たすように制約することで, 提案画像復元モデル (Ours) は, 過学習になりがちな画像ノイズ除去に適用しても高い汎化性能を達成することができる. さらに, これは Stand-alone DLM よりも学習可能なパラメータ数が多い条件下で達成されている. 特に画像の空間的な再構成精度を計測する SSIM は, 両方の実験設定において比較的高い値を有しており, 先験情報による制約が機能していることが見てとれる.

### 4.3.3 ガウシアンノイズ付きピクセル補間の画像復元実験

ガウシアンノイズ付きピクセル補間の実験結果を表 4.3 に示す. ノイズ除去実験結果の

表 4.3: ガウシアンノイズとピクセル補間の混合ノイズにおける実験結果

$N_{\text{train}}$	5		25		50	
Methods	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
SA-DUM	<b>21.91</b>	<b><math>6.840 \cdot 10^{-1}</math></b>	21.99	$6.940 \cdot 10^{-1}$	21.91	$6.857 \cdot 10^{-1}$
SA-DLM	20.54	$6.119 \cdot 10^{-1}$	21.73	$6.890 \cdot 10^{-1}$	21.81	$7.143 \cdot 10^{-1}$
Ours	21.17	$6.594 \cdot 10^{-1}$	<b>22.23</b>	<b><math>7.151 \cdot 10^{-1}</math></b>	<b>22.52</b>	<b><math>7.231 \cdot 10^{-1}</math></b>

表 4.1 および表 4.2 で得られた傾向と同様に、劣化が激しく、学習画像数が 25 枚と 50 枚の場合に、提案画像復元モデル (Ours) は最も良い性能を達成することがわかる。また、DUM は学習画像数が 5 枚の場合に他のモデルを上回ることがわかる。これは、本実験設定ではこれまでのノイズ除去よりも劣化が激しく、データ数が 5 枚と少ない場合には最もロバストな復元手法である Stand-alone DUM が有効であったと考えられる。

#### 4.3.4 Deep learning モデルにおける前処理の効果

ここでは、DUM の前処理を行う Deep learning モデルの効果について検証する。図 4.5 は、無劣化画像、Deep learning モデル単体での出力 (SA-DLM)、提案画像復元手法における Deep learning モデルの出力 (図 4.2 における  $\hat{y}$ )、提案画像復元手法の最終出力 (図 4.2 における  $\hat{x}$ ) の 4 種類の出力を視覚的に比較したものである。

図 4.5 に示すように、提案画像復元モデルにおける DLM は、Deep unfolding モデルで使用される Anisotropic TV 正則化によって元のエッジが過度に平滑化されないように、エッジを強調するような画像を出力する。Deep learning モデルによるこのような前処理は、Deep unfolding モデルで使用される正則化に適応するように学習される。この正則化に基づく最適化された前処理能力は、図の結果画像から観察することができ、Deep learning モデルが単に無劣化画像を近似しているのではないことがわかる。

## 4.4 結論

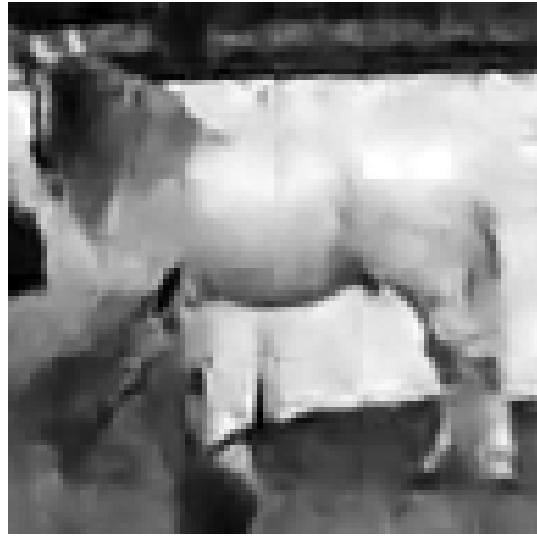
### 4.4.1 本章のまとめ

本章では、小規模データセット条件下における深層学習を用いた画像復元モデルの汎化性能を向上させる手法を提案する。提案画像復元モデルは CNN をベースとした Deep learning モデル (DLM) と反復アルゴリズムを深層展開した Deep unfolding モデル (DUM) から構成される。DUM では、凸最適化を解く反復アルゴリズムを深層展開することで、先験情報を満たすような出力を行うようにモデルが制約されるため、過学習が抑制される。また、DLM は画像復元モデル全体の表現力を高めるとともに、DUM で使用されている先験情報に合わせた画像変換処理を学習する。ノイズ除去タスクおよびピクセル補間タスクにおいて、DUM 単体の出力、DLM 単体の出力、提案画像復元モデルの出力を比較し、DLM 単体の場合と比較して汎化性能が向上することを実験的に示した。





(a) 無劣化画像



(b) SA-DLM



(c) 提案画像復元モデルにおける DLM 出力



(d) 提案画像復元モデルの出力

図 4.5: 表 4.1 における  $N_{\text{train}} = 50$  かつ  $\sigma = 0.15$  の場合の各モデルの結果画像

#### 4.4.2 提案手法の課題

Deep unfolding モデルは、凸最適化における反復アルゴリズムをベースとしている。そのため、計算に時間がかかる処理が反復アルゴリズム内に含まれている場合、その処理が学習におけるボトルネックになる。HSI のように巨大なデータを入力することを考えた

場合、計算時間が膨大になる可能性がある。また、復元精度が Deep unfolding モデル内で使用される正則化に依存している上に、誤差逆伝播によつパラメータ学習のために微分可能な正則化項しか用いることができない。そのため、凸最適化ベースの画像復元手法でよく用いられる特異値分解が使用出来ず、画像復元モデルの設計において大きな制約になってしまう。以上の理由から、本手法を HSI 復元に適用することは困難であると考えられる。

## 第 5 章

# HSI における Separable convolution の構造的優位性

### 5.1 研究背景および関連手法

第 3 章, および第 4 章では, グレースケール画像と擬似スペクトルデータを用いた事前学習に関する手法と, 前処理のための深層学習モデルと凸最適化を解く反復アルゴリズムから設計した深層展開ベースのモデルを組み合わせる手法を提案した. しかし事前学習を用いる手法では, 事前学習で使用するデータセットにノイズ除去モデルの精度が依存してしまう問題点があり, 深層展開を用いた手法では計算量が非常に大きくなってしまいう問題点があった.

本章では, 他ドメインのデータや先験情報などを用いないハイパースペクトル画像 (HSI: Hyperspectral Image) 復元手法を提案する. ゼロショット学習と呼ばれる観測劣化画像のみから学習・復元を行う手法をベースとしている. また同時に Separable convolution の構造的特徴が, ゼロショット HSI 復元タスクにおける HSI の構造的な特徴と合致することにより, 効率的に HSI の情報を捉えることができることを実験的に示す.

#### 5.1.1 自己教師あり学習による RGB 画像復元

近年の深層学習を用いた画像復元手法の発展に伴い, 無劣化画像を必要としない深層学習ベースの RGB 画像復元手法が提案されている [81–88]. 現実的な問題において, 無劣化画像 (またはほぼ劣化がない画像) を得ることが困難な条件下では, 劣化画像と無劣化画像のペアからなる大規模なデータセットを利用することは困難である. 自己教師あり学

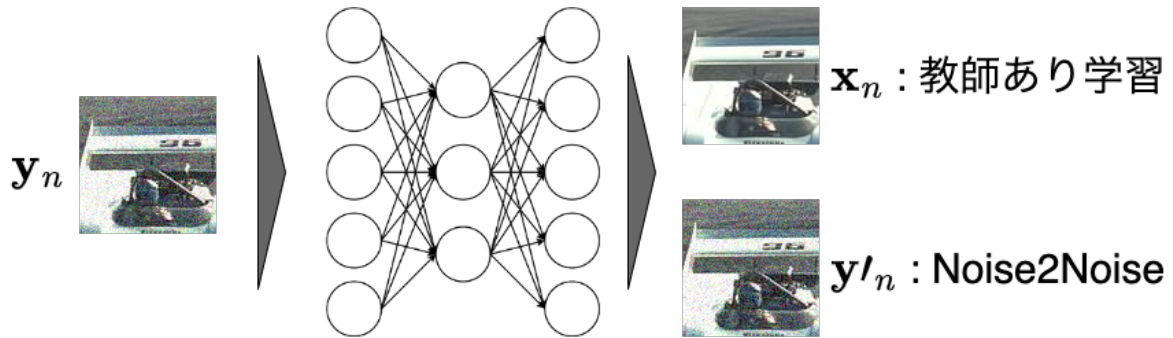


図 5.1: 教師あり学習と Noise2Noise における学習用データペアの違い

習ベースの手法は，このような条件において効力を発揮する．劣化した画像のみから学習ペアを作成するため，無劣化画像を取得することが困難な場合においても画像復元モデルを学習することができる．画像復元における自己教師あり学習手法は，Noise2Noise [81] という 2 枚の劣化画像を用いて学習を行う手法，あるいは Noise2Void [82] という劣化画像の画素値を変化させ，それをもとの画素値に戻すよう学習する手法のいずれかを発展させた手法が多く提案されている．

Noise2Noise [81] では，同じシーンかつ異なるノイズを持つ画像のペアを使用して画像復元モデルを学習する．厳密には自己教師あり学習ではないものの，最初期に提案された無劣化画像を必要としない深層学習ベースのノイズ除去手法である．同一のシーンかつ異なるノイズを持つデータペア  $(y_n, y'_n)$  からなるデータセットを  $D_{n2n} = \{(y_1, y'_1), (y_2, y'_2), \dots, (y_n, y'_n), \dots, (y_N, y'_N)\}$  とすると，画像復元モデルの学習済みパラメータ  $\Theta^*$  は以下の目的関数を最小化することで得られる．

$$\Theta^* = \underset{\Theta}{\operatorname{argmin}} \frac{1}{2N} \sum_{n=1}^N \|f(y_n; \Theta) - y'_n\|_2^2 \quad (5.1)$$

しかし，撮像対象が動く可能性などを考えた場合，現実的な問題において全く同一のシーンを撮影することは困難である．Noise2Noise をベースとする自己教師ありノイズ除去手法では，同一シーンの劣化画像を揃えることの困難性から，1 枚の画像からダウンサンプリングによって作成された 2 枚の劣化画像からなるデータセットを使用して学習する手法が多く提案されている．Noise2Fast [83] では奇数ピクセルと偶数ピクセルに画素を分離し，それぞれサブサンプリングを行うことでダウンサンプリング画像を作成する．その後，ダウンサンプリング画像の画素を縦方向，あるいは横方向にシフトすることで，それぞれのサブサンプリング画像から縦長の画像と横長の画像を生成する．最後にそれぞれのペアで式 (5.1) を用いて学習することで画像復元モデルを得る．Neighbor2Neighbor [85]

では、 $2 \times 2$ の領域ごとに2つの隣接するピクセルをランダムに選択し、サブサンプリングを行うことで2枚のダウンサンプリング画像を作成する。その後、その2枚の画像を用いて式(5.1)に従って学習することを繰り返す。

これに対し Noise2Void [82] では、Noise2Noise が持つ問題点を解決するために、Blind-spot 戦略と呼ばれるダウンサンプリングを使用しない単一の劣化画像から学習ペアを作成する手法を提案している。Noise2Void における Blind-spot 戦略では、劣化画像内のランダムに選択された画素 (Blind-spot) を周囲の画素で置き換えた画像を入力、元の劣化画像を教師画像として学習を行う。Blind-spot の位置にあたる画素同士の誤差のみを学習に使用する。劣化画像からなるデータセットを  $D_{n2v} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$  とすると、画像復元モデルの学習済みパラメータ  $\Theta^*$  は以下の目的関数を最小化することで得られる。

$$\Theta^* = \underset{\Theta}{\operatorname{argmin}} \frac{1}{2N} \sum_{n=1}^N \|M\{f(\check{\mathbf{y}}_n; \Theta) - \mathbf{y}_n\}\|_2^2 \quad (5.2)$$

ここで、 $\check{\mathbf{y}}_n$  は Blind-spot に該当する画素値を周囲の画素値によって置き換えられた画像であり、 $M(\cdot)$  は Blind-spot に該当する画素を抽出する関数である。Noise2Void をベースとする手法では、Blind-Spot の選択方法や、学習における損失関数等を変更した手法が多く提案されている。

### 5.1.2 ゼロショット学習による RGB 画像復元

自己教師あり学習では、主に劣化画像のみからなる大規模データセットを用いて学習する。これに対しゼロショット学習では、学習データと復元対象のデータは同一であり、復元対象のデータのみを学習に使用する、すなわち、学習用データセット  $D_{\text{train}}$  は  $N = 1$  かつ  $D_{\text{train}} = \{\mathbf{y}\}$  であり、復元対象のデータは  $\mathbf{y}$  である。これまでにゼロショット RGB 画像復元手法がいくつか提案されている [89–91]。

Deep image prior [89] は、図 5.2 に示すとおり、特定の分布からサンプリングされたノイズ  $\mathbf{n}_{\text{in}}$  を入力、劣化画像  $\mathbf{y}$  そのものを教師画像として学習を行う。Deep image prior の損失関数は以下の式で表される。

$$L(\mathbf{y}, \mathbf{n}_{\text{in}}) = \|f(\mathbf{n}_{\text{in}}; \Theta) - \mathbf{y}\|_2^2 \quad (5.3)$$

畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) は滑らかで構造的なデータを優先して学習するバイアスを持つため、CNN をベースとする画像復元モデルは学習の過程において劣化画像より先に無劣化画像に近い画像を出力する。その

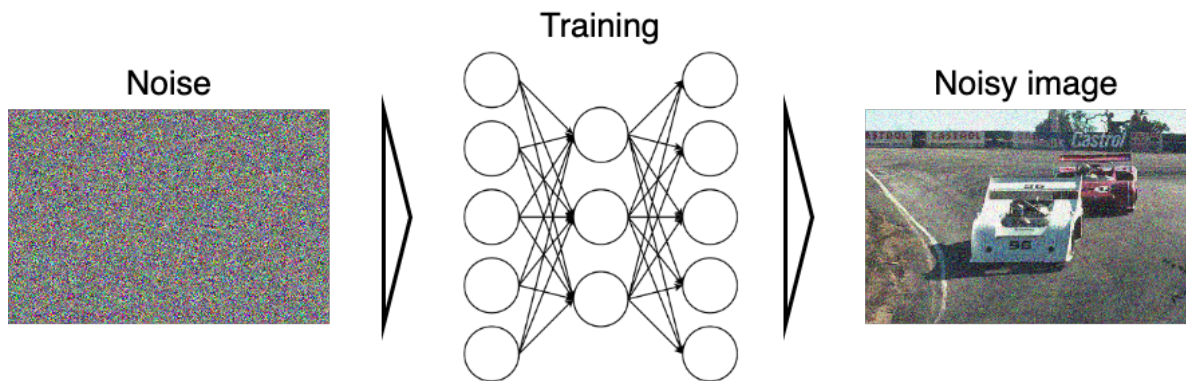


図 5.2: Deep image prior におけるデータペア

後、学習が進むにつれて、 $f(\cdot; \Theta)$  は観測劣化画像  $\mathbf{y}$  そのものを出力するように学習していく。ノイズが復元され始める前に学習を早期停止 (Early-stopping) することで、復元画像  $\hat{\mathbf{x}}$  を得ることができる。しかし、早期停止の適切なタイミングを事前に知ることは困難であるという問題点がある。バイアスについては CNN に限らず、ニューラルネットワークにおいて低周波成分を優勢して学習するバイアスがある [92, 93] ことが知られており、Neural representation learning でも MLP において同様の傾向が見られる [94]。

また、Self2Self [90] は Noise2Void をベースとしたゼロショット RGB 画像復元手法である。画像を入力時、および推論後に Dropout を用いてランダムに画素値を欠損させ、復元した画像の平均を取ることで、ゼロショット画像でも過学習を抑制することができる。しかし、平均を取るために複数枚の欠損画像に対して復元処理を行う必要があるため、計算コストが非常に高い問題点がある。

近年提案された Zero-shot Noise2Noise [91] では、Noise2Noise をベースとしたゼロショット RGB ノイズ除去手法を提案している。従来の自己教師あり学習手法、ゼロショット学習手法と比較して同等の精度であるにもかかわらず、従来手法と比較してモデルサイズが小さく、かつ高速な学習を行うことができる。計算コストが小さいため、GPU を使わない CPU のみでの学習も可能である。

## 5.2 提案手法

本章では、早期停止を必要としない新たなゼロショット HSI 復元手法を提案する。

### 5.2.1 HSI の構造的特徴について

HSI は空間的構造の情報とスペクトル情報から構成されている。そのため、HSI は他の 3 次元データとは異なり動くオブジェクトが存在せず、輝度が異なるグレースケール画像が連続するようなデータである。HSI に対してタッカー分解を行い、特異値をプロットした例を図 5.3 に示す。

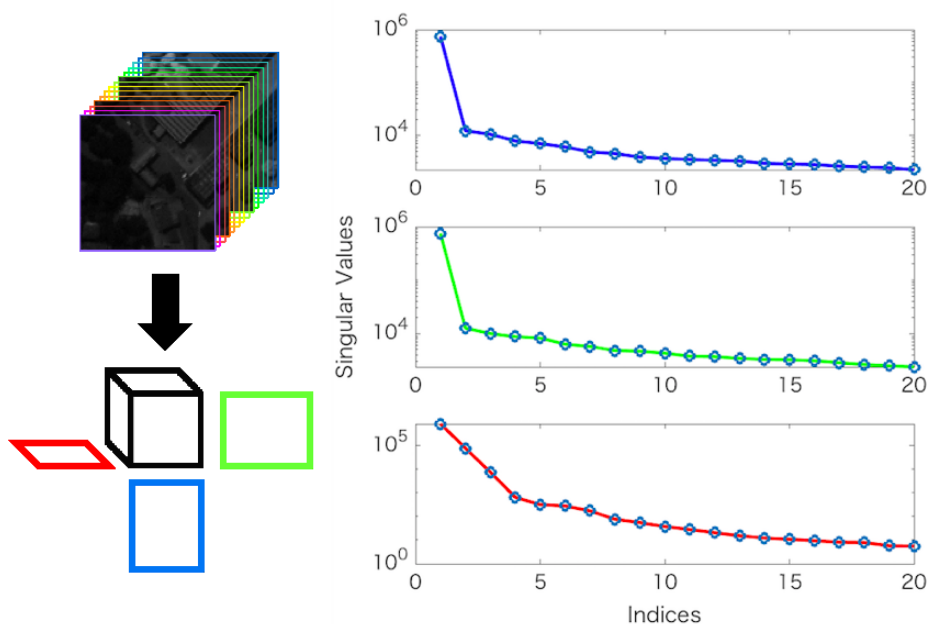


図 5.3: HSI に対するタッカー分解の特異値プロット

図 5.3 から、全てのプロットにおいて特異値が急速な減衰傾向にあり、青と緑で示されている空間方向のプロットと比較して赤色で示されているスペクトル方向の低ランク性が顕著であることがわかる。ハイパースペクトル画素で測定されたスペクトルは、エンドメンバーと呼ばれるごく一部の物質のスペクトル線形和で近似できることがよく知られており、空撮シーンでは、エンドメンバーは身近な巨視的物質に対応する。一般に近傍のエンドメンバーは相関が高い傾向があり、これが HSI の、特にスペクトル方向の低ランク性につながる。

加えて、HSI における様々な軸方向に沿った差分のヒストグラムを図 5.4 に示す。図 5.4 では、4 つのプロット (緑、ピンク、青、赤) がそれぞれ  $x$  軸、 $y$  軸、 $z$  軸と、三次元空間を  $y$  軸と  $z$  軸に沿って対角線方向で横切るような平面である  $y-z$  平面に対する隣接画素の差分のヒストグラムを示している。各差分は、各方向に平面に微分可能なフィルタ

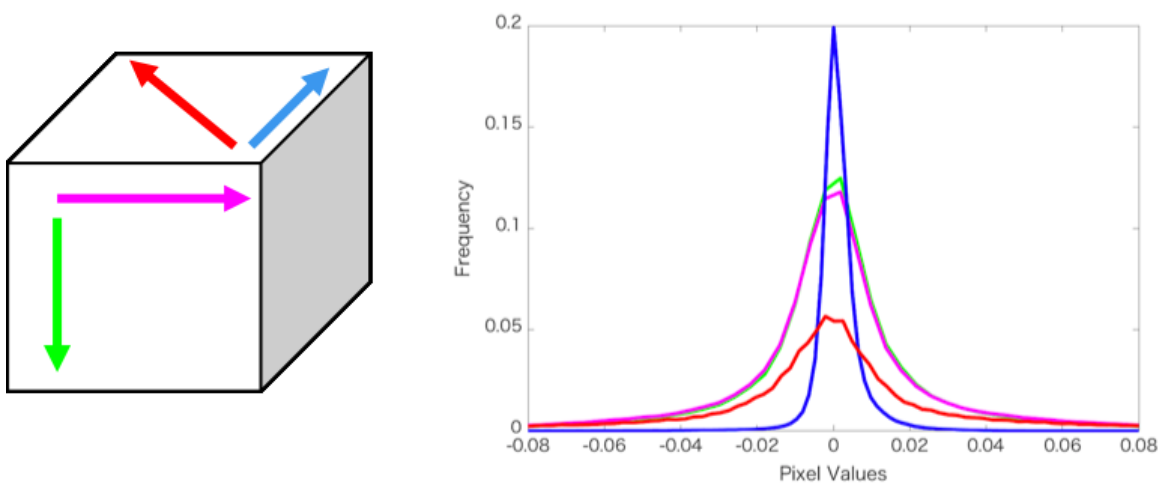


図 5.4: HSI の各軸方向差分のヒストグラム

を適用することで計算されている。ゼロ付近のピークが高いほど特定の軸における差分の値が小さく、近接ピクセル同士の相関が高いことを示す。図 5.4 から、赤い線で示された対角相関が他の 3 方向の相関より低いことは明らかである。これらの特性は、ほとんどの HSI で成立する。これらの方向依存性により、HSI を空間方向とスペクトル方向で別々に扱うことができる。

## 5.2.2 Separable convolution の構造的特徴について

第 3.1.2 小節で述べたように、Separable convolution は空間方向の情報の畳み込みとスペクトル方向の情報の畳み込みに、それぞれ異なるカーネルを用いている。そのため、通常の密な畳み込みとは異なり、直接的に  $x-z$  軸や  $y-z$  軸などの対角方向の相関を利用することができない。これは、パラメータ数を削減できることに対してモデルアーキテクチャによってもたらされる制約のうちの一つであり、このような制約はモデルが十分な性能を発揮できない原因となりうる。しかし、第 5.2.1 小節に示すように、HSI はその他の 3 次元データと比較して空間方向とスペクトル方向における特性が大きく異なる上に、対角方向の情報の相関性が低い。この HSI が持つ構造的特性により、Separable convolution が持つ構造的に対角方向の情報を利用することが困難である欠点による精度低下を抑制することができる。これは HSI を用いた深層学習モデルにおいて、より高精度かつロバストなモデルの構築を実現する。



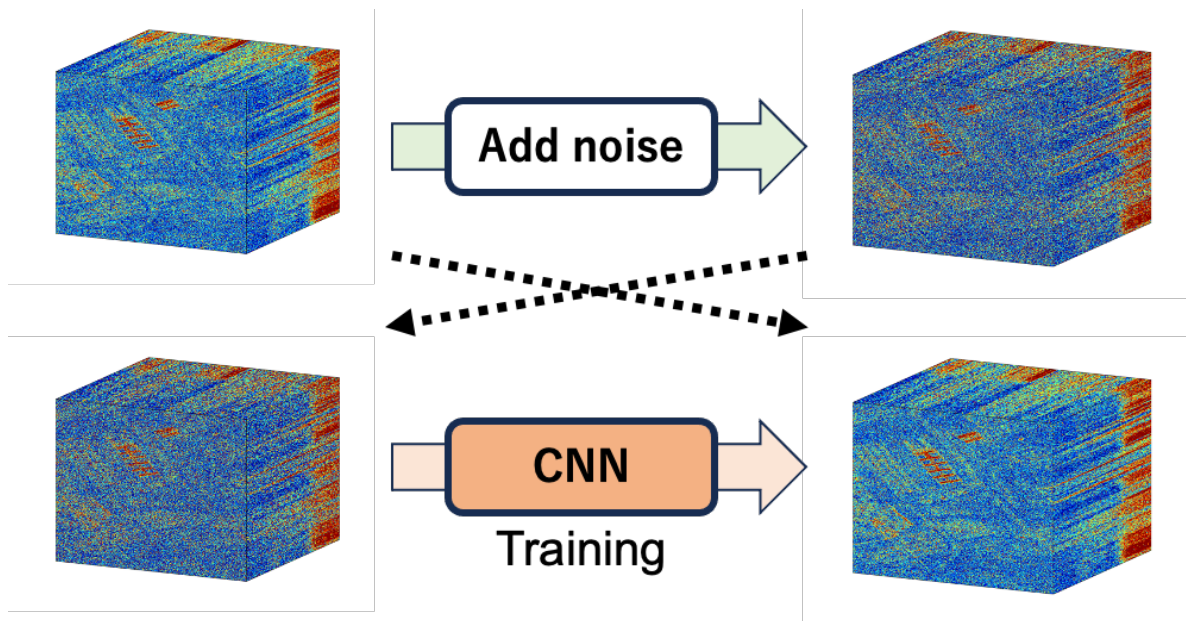


図 5.5: 提案手法における学習用データペア作成方法

### 5.2.3 提案手法における学習方法

本提案手法における学習に使用するデータペアの作成方法を図 5.5 に示す。提案したノイズ除去方法は、HSI が雑音によって劣化していても豊富な情報を持っているという事実に基づいている。観測劣化 HSI である  $\mathbf{y}$  からノイズのない HSI である  $\mathbf{x}$  を復元することができる深層学習モデルは、 $\mathbf{y} + \mathbf{n}$  から  $\mathbf{y}$  を復元する能力を持っていると仮定する。

観測劣化 HSI の劣化過程におけるノイズが加法性白色ガウシアンノイズに従う場合の例を以下に示す。上述した仮定に基づき、観測劣化 HSI からノイズの標準偏差  $\sigma'$  を推定し、標準偏差  $\tilde{\sigma} = (1 + \alpha)\sigma'$ , ( $\alpha \ll 1$ ) のノイズを  $\mathbf{y}$  に加える。これを式で表すと以下のようなになる。

$$\tilde{\mathbf{y}} = \mathbf{y} + \tilde{\mathbf{n}} \quad (5.4)$$

ここで、 $\tilde{\mathbf{n}}$  は平均 0, 標準偏差  $\tilde{\sigma}$  のガウシアンノイズである。推定した標準偏差の値の誤差を考慮し、 $\alpha$  の値は指定した範囲内でランダムにサンプリングされる。これにより、二重劣化 HSI である  $\tilde{\mathbf{y}}$  と観測劣化 HSI である  $\mathbf{y}$  からなる学習用データセット  $\{\tilde{\mathbf{y}}, \mathbf{y}\}$  が作成される。本提案手法では二重劣化 HSI を作成するために、観測劣化 HSI がどのようなノイズによって劣化しているか、ノイズの標準偏差が推定できない場合はその標準偏差に関する情報など、観測劣化 HSI の劣化過程に関する情報が必要となる。

学習用データセットを使用して深層学習モデルのパラメータを最適化し、 $l_2$  ノルムを使用した単純な損失関数を最小化する。

$$\min_{\Theta} \|f(\tilde{\mathbf{y}}, \Theta) - \mathbf{y}\|_2^2 \quad (5.5)$$

この手法によって学習された深層学習モデルは十分な性能を発揮するが、数百エポックごとに観測劣化 HSI を深層学習モデルを用いて復元した HSI に置き換えることで、さらに性能が向上する可能性がある。本提案手法では、Deep image prior における観測劣化 HSI をそのまま出力してしまう問題点が解決されている。加えて、二重劣化 HSI は観測劣化 HSI に対して大量に作成することができる上に、タスクもノイズ除去に限らない。

## 5.2.4 モデルアーキテクチャ

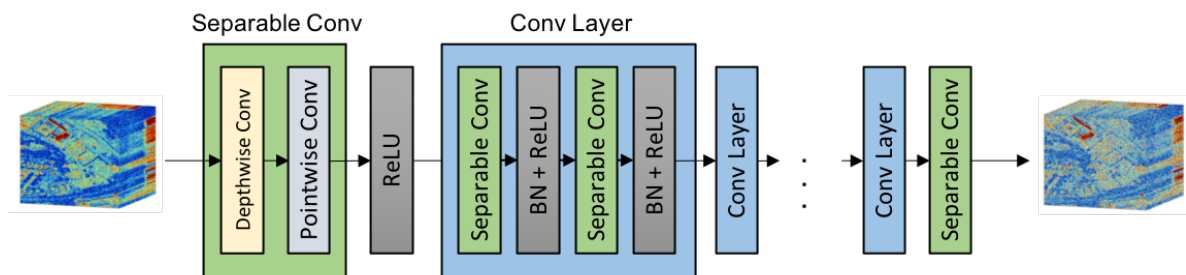


図 5.6: モデルアーキテクチャ モデル図

密な係数を持つ通常の畳み込み層を使用した CNN は、同じ数の層を持つ Separable convolution を用いた CNN よりも画像の特徴を捉える能力に優れているが、HSI のように多くのバンドを持つ 3 次元キューブ状のデータへの訓練は困難である。なぜなら、スペクトル方向の特徴を捉えるために、それぞれの層により多くのカーネルが使用されるため、カーネルの最適化が非常に難しくなるからである。第 5.2.1 小節で述べた HSI の特性を考慮すると、空間的・スペクトルの平面における対角線上の相関を無視しても問題ないと考えられる。そのため、本提案手法では全ての畳み込み層において Separable convolution を採用する。

本提案手法におけるモデルアーキテクチャは図 5.6 のようになっている。モデルアーキテクチャは、第 3.2.1 小節で紹介されている Separable convolution を用いた畳み込み層、Batch normalization, 活性化関数 ReLU からなる DnCNN ベースの CNN に類似している。Skip connection が削除され、中間層での処理が Batch normalization, ReLU, Separable convolution の順番から、Separable convolution, Batch normalization, ReLU

の順に変更されている。Separable convolution を用いることで、モデルの表現力を保持しつつ、過学習のリスクを抑えている。

## 5.3 実験

### 5.3.1 実験設定

本実験では、新たに提案したゼロショット HSI 復元手法の有効性、ゼロショット HSI 復元における Separable convolution の有効性、画像復元以外の HSI 関連タスクにおける Separable convolution の有効性について検証する。新たに提案したゼロショット HSI 復元の有効性は、加法性白色ガウシアンノイズ除去実験において比較を行った。劣化 HSI は式 (2.2) に従って与えられる。比較手法としては、3次元データに対する一般的なガウシアンノイズ除去手法である BM4D [72]、高精度な最適化ベースのガウシアンノイズ除去手法である FastHyDe [38]、ゼロショット RGB 画像復元手法である Deep image prior [89]、および自己教師あり画像復元手法である Noise2Void [82] を使用した。Deep image prior、および Noise2Void は HSI 復元で使用できるように拡張して使用した。ガウシアンノイズ除去実験におけるデータは Frisco, Stanford, PaviaU, IndianPine, Washington DC を [0, 1] に正規化して使用した。評価指標には Peak Signal-to-Noise Ratio (PSNR) を用いる。

また、ゼロショット HSI 復元における Separable convolution の有効性について検証を行うために、ノイズ除去（加法性白色ガウシアンノイズ除去、加法性白色ラプラシア

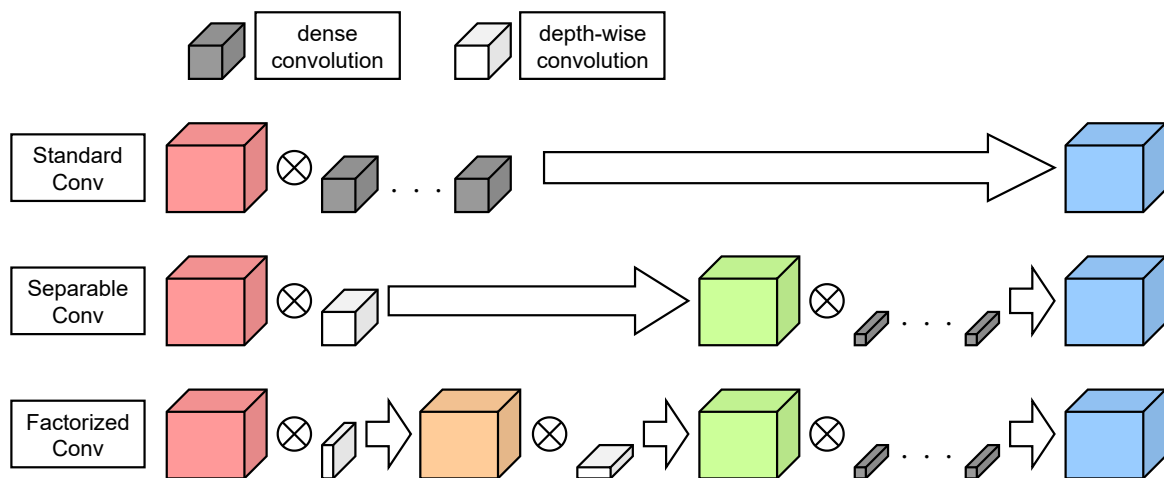


図 5.7: 実験で使用する畳み込み手法による処理の違い

表 5.1: 各データにおけるモデルの学習可能パラメータ数

畳み込み手法	Frisco, Stanford	PaviaC	PaviaU
Standard	240,094	120,732	121,273
Separable	237,688	116,892	119,068
Factorized	234,728	114,852	117,008

ンノイズ除去), およびピクセル補間の実験を行った. 比較においては, 本章の提案手法のモデルにおいて使用されている畳み込み層を, 図 5.7 に記載している通常の畳み込み (Standard), Separable convolution (Separable), Separable convolution の空間方向の畳み込みをさらに縦方向と横方向に分離した Factorized convolution (Factorized) に変更し, 復元精度の比較を行う. この際, パラメータ数がほとんど同一になるようにチャンネル数を変更することで調整を行なった. 各データに対するモデルパラメータは表 5.1 に記載されている. 畳み込み精度の検証におけるデータセットは Frisco, Stanford, PaviaC, PaviaU であり,  $[0, 1]$  に正規化して使用した. PaviaC は画像が大きいため, 画像の一部を切り抜いて使用している. 評価指標は, HSI 全体の復元精度を評価する PSNR と, 空間方向の復元精度を評価する Structural Similarity Index [80] (SSIM) を用いた.

これに加えて, 画像復元以外のタスクにおける Separable convolution の有効性を検証するために, クラス分類実験を行なった. クラス分類実験では, 近年提案された HSI に対するクラス分類手法である MFT [12] をベースラインとして使用した. MFT では, HSI から作成されたパッチに対し固定長のベクトルへの埋め込みを行う. 本実験では, 埋め込み処理におけるモデルアーキテクチャを, Batch normalization, ReLU, Separable convolution からなるモジュールを連結した構造に変更し, クラス分類精度を比較する. 実験に使用するソースコード, およびデータセットは, MFT の著者らによって提供されているソースコードとパッチ分割済み Trento データセットを使用した\*1. 評価指標は, MFT の論文で使用されている Overall Accuracy (OA), Average Accuracy (AA), Statistical Kappa ( $\kappa$ ) を使用した.

\*1 <https://github.com/AnkurDeria/MFT>

表 5.2: ガウシアンノイズ除去実験結果

Methods	$\sigma$	PaviaU	Frisco	Stanford	IndianPine	Washington
BM4D [72]	0.05	38.53	40.41	40.74	38.01	39.59
	0.10	34.94	36.74	37.04	35.26	35.58
	0.15	32.71	34.75	35.04	33.69	33.41
	0.20	31.15	33.39	33.67	32.55	31.98
	0.25	30.00	32.33	32.65	31.65	30.89
FastHyDe [38]	0.05	<b>40.23</b>	<b>45.78</b>	<b>46.08</b>	38.96	<b>44.86</b>
	0.10	37.27	41.78	42.19	37.29	40.71
	0.15	35.35	39.57	40.20	36.04	38.38
	0.20	34.11	38.20	38.76	35.17	36.80
	0.25	33.05	37.13	37.39	34.48	35.71
DIP [89]	0.05	34.45	37.72	37.18	36.12	33.84
	0.10	32.85	33.82	32.43	31.65	30.56
	0.15	30.52	31.08	28.70	28.42	27.23
	0.20	27.91	27.61	25.65	25.99	24.71
	0.25	25.62	25.45	24.08	24.12	22.63
N2V [82]	0.05	31.32	35.57	37.20	33.42	35.30
	0.10	30.64	35.35	36.14	34.00	34.57
	0.15	30.36	35.26	35.57	33.81	33.78
	0.20	30.10	34.56	34.64	33.34	33.04
	0.25	29.71	34.06	34.11	32.90	32.73
Ours	0.05	38.66	42.62	45.78	<b>39.74</b>	44.09
	0.10	<b>37.31</b>	<b>41.81</b>	<b>42.48</b>	<b>37.66</b>	<b>41.43</b>
	0.15	<b>35.95</b>	<b>39.76</b>	<b>40.35</b>	<b>36.48</b>	<b>39.52</b>
	0.20	<b>34.81</b>	<b>38.72</b>	<b>39.02</b>	<b>35.55</b>	<b>37.85</b>
	0.25	<b>33.73</b>	<b>37.58</b>	<b>37.90</b>	<b>34.75</b>	<b>36.63</b>

### 5.3.2 二重劣化 HSI を用いたゼロショット HSI ノイズ除去におけるガウシアンノイズ除去実験

本節では標準偏差  $\sigma = 0.05, 0.10, 0.15, 0.20, 0.25$  における白色ガウシアンノイズ除去の実験結果について述べる。実験結果を表 5.2 に示す。

表 5.2 に示す通り、 $\sigma = 0.05$  の場合において、FastHyDe に劣っているものの、それ以外のほとんどの条件において、提案手法が従来手法より優れていることが見て取れる。特に、その他のゼロショット学習ベースの画像復元手法と比較して大きく上回っている。深層学習モデルにおける膨大なパラメータ数による高い表現力が、激しい劣化における困難な画像復元において有効に機能していることから、劣化が激しい場合において FastHyDe を上回る精度を得られていると考えられる。

### 5.3.3 ノイズ除去タスクにおける Separable convolution の有効性

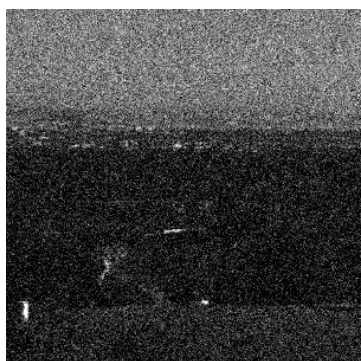
本節では提案手法における畳み込み層を変更した場合における、標準偏差  $\sigma = 0.1$  および  $\sigma = 0.15$  のガウシアンノイズ除去、ラプラシアンノイズ除去の実験結果について述べる。劣化 HSI は式 (2.1) に従って与えられる。ガウシアンノイズ除去実験の結果を表 5.3, 表 5.4, 図 5.8, ラプラシアンノイズ除去実験の結果を表 5.5, 表 5.6, 図 5.9 に示す。

表 5.3: 畳み込み手法変更におけるガウシアンノイズ除去結果 (PSNR)

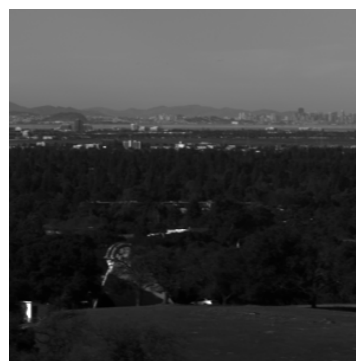
$\sigma$	畳み込み手法	Frisco	Stanford	PaviaC	PaviaU
0.1	Standard	40.38	38.94	37.05	35.28
	Separable	<b>41.81</b>	<b>42.48</b>	<b>37.52</b>	<b>37.31</b>
	Factorized	41.47	41.69	36.66	36.52
0.15	Standard	37.56	34.39	35.45	35.56
	Separable	<b>39.88</b>	<b>40.28</b>	<b>36.04</b>	<b>35.95</b>
	Factorized	39.67	36.35	35.92	35.87

表 5.4: 畳み込み手法変更におけるガウシアンノイズ除去結果 (SSIM)

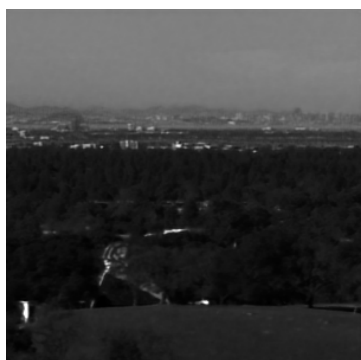
$\sigma$	畳み込み手法	Frisco	Stanford	PaviaC	PaviaU
0.1	Standard	$9.733 \cdot 10^{-1}$	$9.571 \cdot 10^{-1}$	$9.491 \cdot 10^{-1}$	$9.373 \cdot 10^{-1}$
	Separable	<b><math>9.896 \cdot 10^{-1}</math></b>	<b><math>9.905 \cdot 10^{-1}</math></b>	<b><math>9.772 \cdot 10^{-1}</math></b>	<b><math>9.773 \cdot 10^{-1}</math></b>
	Factorized	$9.760 \cdot 10^{-1}$	$9.761 \cdot 10^{-1}$	$9.360 \cdot 10^{-1}$	$9.350 \cdot 10^{-1}$
0.15	Standard	$9.531 \cdot 10^{-1}$	$9.270 \cdot 10^{-1}$	$9.322 \cdot 10^{-1}$	$9.337 \cdot 10^{-1}$
	Separable	<b><math>9.860 \cdot 10^{-1}</math></b>	<b><math>9.861 \cdot 10^{-1}</math></b>	<b><math>9.694 \cdot 10^{-1}</math></b>	<b><math>9.718 \cdot 10^{-1}</math></b>
	Factorized	$9.668 \cdot 10^{-1}$	$9.478 \cdot 10^{-1}$	$9.335 \cdot 10^{-1}$	$9.326 \cdot 10^{-1}$



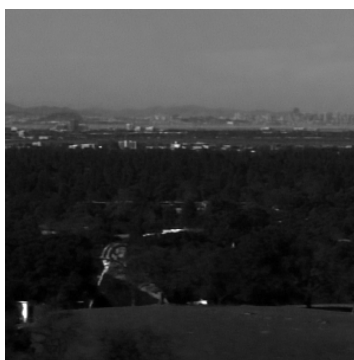
(a) ノイジー画像



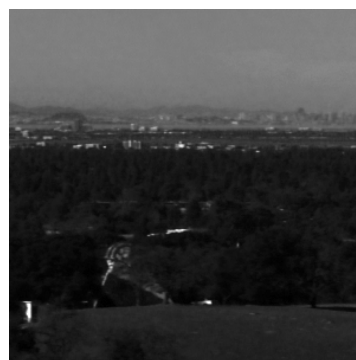
(b) 無劣化画像



(c) Standard



(d) Separable



(e) Factorized

図 5.8: ガウシアンノイズ除去における畳み込み手法による復元結果の違い

表 5.5: 畳み込み手法変更におけるラプラシアンノイズ除去結果 (PSNR)

$\sigma$	畳み込み手法	Frisco	Stanford	PaviaC	PaviaU
0.1	Standard	38.93	39.63	35.35	34.90
	Separable	<b>40.61</b>	40.12	<b>36.55</b>	<b>36.29</b>
	Factorized	40.22	<b>40.27</b>	35.97	35.95
0.15	Standard	37.87	37.43	<b>34.08</b>	<b>33.89</b>
	Separable	<b>38.61</b>	38.40	33.31	33.82
	Factorized	38.58	<b>38.91</b>	33.55	33.67

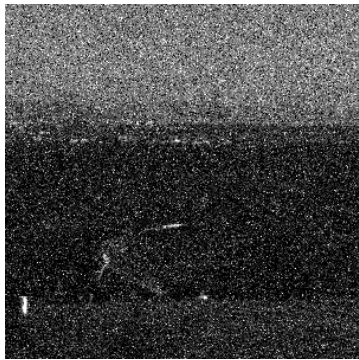
表 5.6: 畳み込み手法変更におけるラプラシアンノイズ除去結果 (SSIM)

$\sigma$	畳み込み手法	Frisco	Stanford	PaviaC	PaviaU
0.1	Standard	$9.627 \cdot 10^{-1}$	$9.681 \cdot 10^{-1}$	$9.339 \cdot 10^{-1}$	$9.300 \cdot 10^{-1}$
	Separable	<b><math>9.881 \cdot 10^{-1}</math></b>	<b><math>9.841 \cdot 10^{-1}</math></b>	<b><math>9.726 \cdot 10^{-1}</math></b>	<b><math>9.730 \cdot 10^{-1}</math></b>
	Factorized	$9.709 \cdot 10^{-1}$	$9.689 \cdot 10^{-1}$	$9.292 \cdot 10^{-1}$	$9.343 \cdot 10^{-1}$
0.15	Standard	$9.528 \cdot 10^{-1}$	$9.462 \cdot 10^{-1}$	$9.074 \cdot 10^{-1}$	$9.057 \cdot 10^{-1}$
	Separable	<b><math>9.806 \cdot 10^{-1}</math></b>	<b><math>9.790 \cdot 10^{-1}</math></b>	<b><math>9.363 \cdot 10^{-1}</math></b>	<b><math>9.497 \cdot 10^{-1}</math></b>
	Factorized	$9.538 \cdot 10^{-1}$	$9.596 \cdot 10^{-1}$	$8.739 \cdot 10^{-1}$	$8.820 \cdot 10^{-1}$

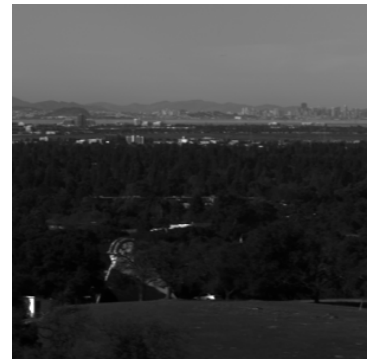
表 5.3, 表 5.4, 表 5.5, 表 5.6 から, ノイズ除去におけるほとんどの実験設定において, Separable convolution がその他の畳み込みと比較して優れていることがわかる. 表 5.1 にあるように学習可能パラメータ数は通常の畳み込み (Standard) が最も多いにも関わらず, 全体の傾向として Separable convolution, Factorized convolution が通常の畳み込みより良い復元精度を示している. このことから, 空間方向とスペクトル方向を分離する畳み込み方法は, ゼロショット HSI ノイズ除去において有効であると考えられる.

また, Factorized convolution は他の畳み込みと比較して, PSNR の復元精度の差よりも SSIM の復元精度の差が大幅に大きくなっている. これは, 表 5.3, 表 5.4 における PaviaC, PaviaU, 表 5.5, 表 5.6 における Frisco, Stanford, PaviaC で顕著である. 表 5.3, および表 5.4 では,  $\sigma = 0.1$  の PaviaU において, PSNR において通常の畳み込みを上回っているにも関わらず SSIM では下回っており,  $\sigma = 0.15$  の PaviaC,





(a) ノイジー画像



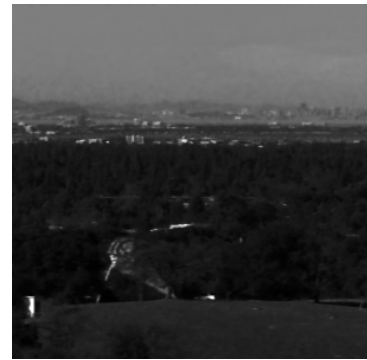
(b) 無劣化画像



(c) Standard



(d) Separable



(e) Factorized

図 5.9: ラプラシアンノイズ除去における畳み込み手法による復元結果の違い

PaviaU では PSNR においては Separable convolution に近い数値であるものの、SSIM では大きく差が開いている。また、表 5.5, 表 5.6 では、 $\sigma = 0.1$  および  $\sigma = 0.15$  の Stanford において、PSNR では Separable convolution を上回っているにも関わらず SSIM は Separable convolution の方が良く、 $\sigma = 0.15$  の Frisco では PSNR の値が Separable convolution に近いにも関わらず、通常の畳み込みに近い SSIM の値となっている。Factorized convolution は空間方向の畳み込みを縦方向と横方向に分離する構造であるため、 $x$  軸と  $y$  軸からなる  $x - y$  平面における対角方向の相関を利用することが困難になる。これは空間的構造の対角方向を利用した補間を困難にするため、SSIM が低下したと考えられる。

図 5.8, および図 5.9 には、標準偏差  $\sigma = 0.1$  の場合のそれぞれのノイズに対する復元画像を表示している。可視化のために、Frisco データの 40 バンド目を抽出し、画素値の補正を行っている。図中に表示されている画像からわかるように、どの畳み込み手法も優

れた復元精度を示しているものの、特に背景の部分において、Separable convolution が優れた復元精度を有することがわかる。

### 5.3.4 ピクセル補間タスクにおける Separable convolution の有効性

本節ではピクセル補間の実験結果について述べる。観測劣化 HSI は以下の式で与えられる。

$$\mathbf{y} = R(\mathbf{x}, r) \quad (5.6)$$

ここで、 $R(\cdot, r)$  は  $[0, 1]$  の欠損割合  $r$  に従ってピクセルをランダムに欠損させる関数であり、本実験では  $r = 0.25$  とする。本提案手法におけるピクセル補間タスクでは、無劣化 HSI からランダムに  $r$  割のピクセルを欠損した HSI を観測劣化 HSI とし、観測劣化 HSI からさらに  $r$  割のピクセルを欠損させた HSI を二重劣化 HSI として使用する。その後、第 5.2.3 小節で説明したように深層学習モデルを学習し、画像復元を行う。

実験結果を表 5.7, 表 5.8, 図 5.10 に示す。

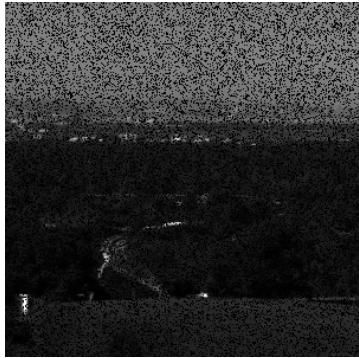
表 5.7, 表 5.8 から、PaviaC を除いたすべてのデータにおいて、Separable convolution を用いたモデルが優れた復元精度を有していることがわかる。また、図 5.10 には、ピクセル補間タスクにおける復元画像を表示している。ノイズ除去実験の際と同様に可視化のために、Frisco データの 40 バンド目を抽出し、画素値の補正を行っている。図中の画像

表 5.7: 畳み込み手法変更におけるピクセル補間結果 (PSNR)

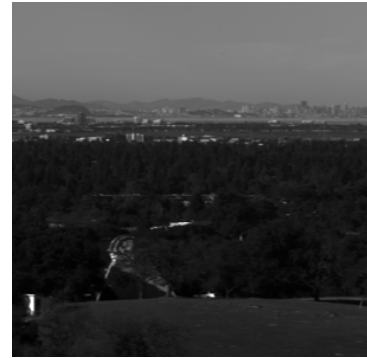
r	畳み込み手法	Frisco	Stanford	PaviaC	PaviaU
0.25	Standard	29.15	36.28	<b>36.61</b>	34.91
	Separable	<b>39.78</b>	<b>41.84</b>	36.05	<b>36.62</b>
	Factorized	39.30	41.00	34.71	36.07

表 5.8: 畳み込み手法変更におけるピクセル補間結果 (SSIM)

r	畳み込み手法	Frisco	Stanford	PaviaC	PaviaU
0.25	Standard	$9.529 \cdot 10^{-1}$	$9.591 \cdot 10^{-1}$	<b><math>9.608 \cdot 10^{-1}</math></b>	$9.512 \cdot 10^{-1}$
	Separable	<b><math>9.855 \cdot 10^{-1}</math></b>	<b><math>9.886 \cdot 10^{-1}</math></b>	$9.584 \cdot 10^{-1}$	<b><math>9.580 \cdot 10^{-1}</math></b>
	Factorized	$9.810 \cdot 10^{-1}$	$9.840 \cdot 10^{-1}$	$9.442 \cdot 10^{-1}$	$9.526 \cdot 10^{-1}$



(a) ノイジー画像



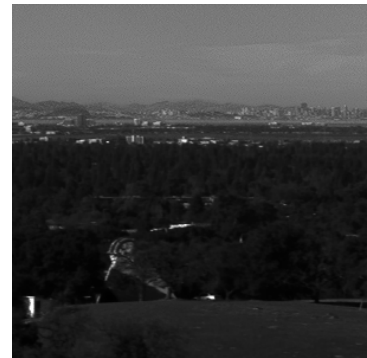
(b) 無劣化画像



(c) Standard



(d) Separable



(e) Factorized

図 5.10: ピクセル補間タスクにおける畳み込み手法による復元結果の違い

から、通常の畳み込み (Standard) では画像中の建物などがぼやけてしまっているのに対し、Separable convolution, および Factorized convolution を用いた場合の結果画像は優れた復元結果を示している。加えて、Separable convolution と Factorized convolution の結果画像の背景部分を比較してみると、Factorized convolution はアーチファクトが発生していることがわかる。そのため、視覚的評価において、Separable convolution における復元結果画像が最も優れていることがわかる。

本小節と第 5.3.3 小節の結果から、空間方向とスペクトル方向の畳み込みを分離し、なおかつ空間方向を正方形のカーネルを用いて一括で畳み込みを行う Separable convolution を用いたモデルアーキテクチャが、ゼロショット HSI 復元において優れていることが示された。

表 5.9: クラス分類モデルにおける学習可能パラメータ数

Methods	Trainable params
MFT	272,943
Separable base	265,663
Separable deep	269,695
Separable wide	277,633

表 5.10: クラス分類結果

Methods	Overall Acc. (OA)	Average Acc. (AA)	KAPPA
MFT	97.87±00.36	94.29±03.93	97.15±00.48
Separable base	<b>98.00±00.38</b>	<u>95.56±01.24</u>	<b>97.33±00.52</b>
Separable deep	97.33±00.29	93.73±01.37	96.43±00.40
Separable wide	<u>97.98±00.56</u>	<b>95.78±01.49</b>	<u>97.30±00.75</u>

### 5.3.5 クラス分類タスクにおける Separable convolution の有効性

本小節では、HSI に対するクラス分類タスクにおける Separable convolution の有効性について述べる。本実験ではオリジナルの MFT (MFT) に対して、埋め込み処理における畳み込み構造を中間層のチャンネル数と層数を変化させた 3 種類のモデルアーキテクチャに変更する。入力 HSI のバンド数  $B$  に対し、中間層チャンネル数  $2B$  かつ層数を 4 層としたモデル (Separable base)、中間層チャンネル数  $3B$  かつ層数を 3 層としたモデル (Separable wide)、中間層チャンネル数  $B$  かつ層数を 12 層としたモデル (Separable deep) を比較として使用した。各モデルにおける中間層チャンネル数、および層数はオリジナル MFT とパラメータ数が近くなるように設定されている。学習パラメータ数は表 5.9 に記載する。実験結果を表 5.10 に示す。

表 5.10 では、すべての実験設定において 5 回学習、推論し、各評価指標の平均値 ± 標準標準偏差で表記している。また、表における太字はその実験設定において最も良い精度を示し、下線は 2 番目に良い復元精度を示す。表 5.10 から、オリジナルの MFT のパラメータ数より少ないにも関わらず、Separable convolution による埋め込みを用いたモデルの精度が優れていることがわかる。OA、および KAPPA では Separable base モデ

ルが最も良く、AA では Separable wide モデルに及ばないものの 2 番目に良い精度であり、MFT を上回る精度を示している。よって、3 次元畳み込みと HetConv [95] からなる MFT で使用されている埋め込みと比較して、Separable convolution を用いた埋め込みが有効であることが示された。この結果から、Separable convolution は HSI が持つデータの特徴をより良く捉えることができ、画像復元に限らず様々なタスクで利用することができる可能性があることがわかる。

## 5.4 結論

### 5.4.1 本章のまとめ

本研究では観測劣化 HSI の劣化過程に関する情報を利用するゼロショット HSI 復元手法を提案する。観測劣化 HSI をさらに劣化させた二重劣化 HSI を作成し、二重劣化画像を入力、観測劣化 HSI を教師画像として学習する。白色ガウシアンノイズ除去実験において、従来の最適化ベース手法並びにゼロショット学習ベースの画像復元手法と比較して優れた復元性能を示した。また、ゼロショット HSI 復元における Separable convolution の有効性についても同時に検証した。HSI は空間構造に関する情報とスペクトルに関する情報によって構成される三次元データであり、スペクトル方向を含む対角方向の相関が低いという特性がある。Separable convolution は空間方向の畳み込みとスペクトル方向の畳み込みを構造上分離しており、この畳み込みの構造がゼロショット HSI 復元において有効に機能する。白色ガウシアンノイズ除去、および白色ラプラシアンノイズ除去実験において、ほぼ同等のパラメータ数を持つ通常の畳み込みを用いたモデル、Separable convolution を用いたモデル、Factorized convolution を用いたモデルを比較し、Separable convolution を用いたモデルが優れた復元精度を有することを実験的に示した。

### 5.4.2 提案手法の課題

本提案手法では、二重劣化 HSI を作成する際に、観測劣化 HSI の劣化過程の情報を使用する。しかし、通常劣化過程を知ることはできないため、現実の劣化 HSI 復元問題において本手法を適用することが困難であると考えられる。劣化過程として何らかの分布を仮定し使用することも考えられるが、劣化の強度を推定する必要があるため、依然として実利用上の制約がある。



## 第6章

# 3D パターン化マスクを使用したゼロショット HSI ノイズ除去

### 6.1 研究背景および関連手法

第3章, および第5章の手法では観測劣化ハイパースペクトル画像 (HSI: Hyperspectral Image) 以外のデータや観測劣化 HSI に関する情報を学習に利用していた. 第3章では, 観測劣化 HSI に近いドメインのグレースケール画像と擬似スペクトルデータ, 第5章では, 観測劣化 HSI における劣化過程の情報を必要とする. しかしこれらの手法では, 学習に使用する追加データが観測劣化 HSI と大きく異なる場合, ノイズ除去モデルの精度が低下してしまう問題点がある.

本章では, 深層学習ベースのゼロショット HSI ノイズ除去手法を新たに提案する. 従来手法では, 学習を早期停止するタイミングや, 第5章で使用されている二重劣化画像作成のための劣化過程など, 画像復元精度を大きく左右するハイパーパラメータを事前に決定する必要があった. 本提案手法では, Blind-spot 戦略をもとにした学習ペア作成方法を用いることで, それらの必要性を回避している. 本手法の利点は, 学習画像や無劣化画像を必要とせず, ノイズ分布情報も必要としない点である. 本章では, 3D パターン化マスク, およびランダム 3D パターン化マスクの2種類のマスクを提案する. マスクによって特定のパターンに従って画素値を欠損させた後, 復元するように学習する. 3次元的なマスクを使用することにより, 空間方向とスペクトル方向の両方から情報を取り込み, 復元精度を向上させる. 加えて, ランダム 3D パターン化マスクでは, Blind-spot の組み合わせをランダム化することで, 学習のロバスト性を向上させる. 本手法が, 特定の種類のノイズにおいて従来手法よりも高い復元精度を達成し, 特にランダム 3D パターン化マスク

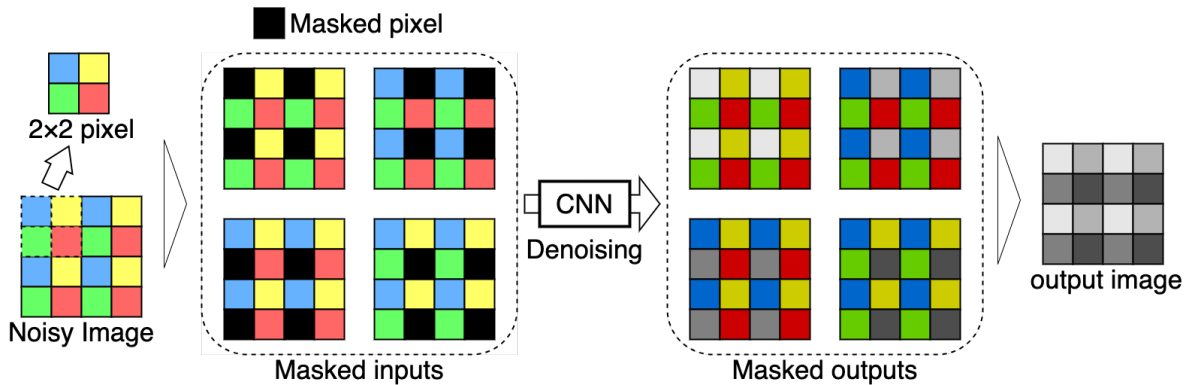


図 6.1: Blind2Unblind における Blind-spot 戦略

においてゼロショット学習時の過学習を効果的に抑制することを実証する。

### 6.1.1 Blind2Unblind

Blind2Unblind [86] は、近年提案された自己教師ありノイズ除去手法であり、Blind-spot 戦略をベースとしている。劣化画像のみからなる大規模なデータセット  $D_{B2U} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n, \dots, \mathbf{y}_N\}$  を用いて深層学習モデルの学習を行う。Blind2Unblind 内で使用されている Blind-spot 戦略を、図 6.1 に示す。この手法では、Blind-spot を  $2 \times 2$  の領域に規則的に配置し、異なる Blind-spot を持つ 4 つのマスク済み画像をモデルに与え、各入力に対応する復元画像を得る。そして、得られた復元画像から、それぞれの Blind-spot に対応する画素のみを抽出して合成することで出力を得ることができる。学習過程において損失関数は、上述した Blind-spot に対応する画素値をつなぎ合わせて得られる出力と、劣化画像を深層学習モデルに直接入力することで得られる出力との重み付き和であり、観測劣化画像から得られる出力の重みを徐々に大きくする。学習後は、復元対象の劣化画像を直接モデルに入力することで復元を行う。この手法は、大量の劣化画像に依存しているため、同じ条件（例えば、バンド数や撮像対象物）で大量のデータを用意するコストが法外に高い HSI に直接適用するのは困難である。

### 6.1.2 ゼロショット学習による HSI 復元

ゼロショット RGB 画像復元手法をベースとしたゼロショット HSI 復元手法もこれまでに提案されている [96]。Deep hs prior [96] は、前章で述べた Deep image prior を HSI に対して適用できるよう拡張したゼロショット HSI 復元手法である。Deep hs prior にお



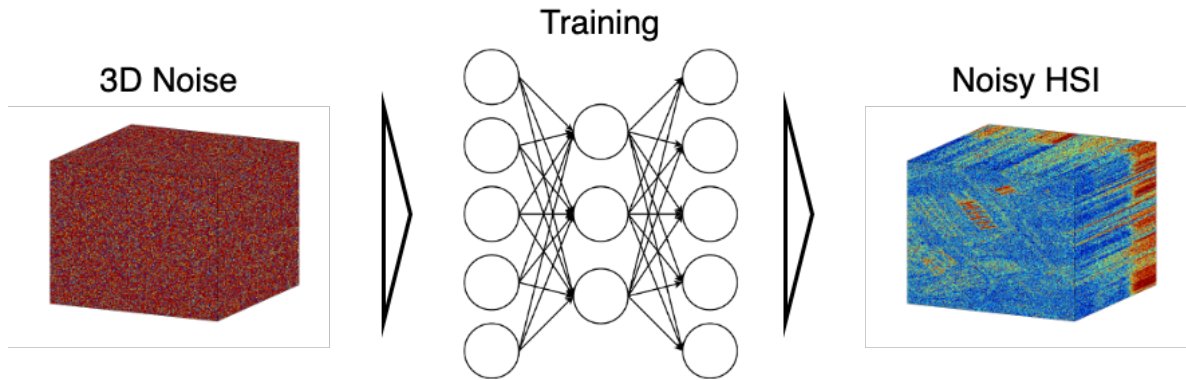


図 6.2: Deep hs prior におけるデータペア

けるデータペアを図 6.2 に示す. 図 6.2 に示すように, 特定の分布からサンプリングされたノイズを入力とし, 観測劣化 HSI を教師画像とする. Deep hs prior も式 (5.3) の損失関数によって学習される. Deep image prior と同様の学習方法を用いているため, 学習が進みすぎると観測劣化 HSI をそのまま出力してしまうという欠点が依然として残っている. こちらの手法も Deep image prior と同様に学習の早期停止が必要であるため, 前もって適切なエポック数を指定する必要があるが, それらは困難である.

第 5 章で提案した手法は, 観測劣化 HSI における劣化過程の情報を利用することでこの問題に対処している. 観測劣化 HSI を教師画像とし, 観測劣化 HSI の劣化過程に沿ってさらに劣化させた二重劣化 HSI を入力として, 教師あり学習手法のようにモデルを学習する. 1 つの観測劣化 HSI に対してノイズを変更することで多数の二重劣化 HSI を生成できるため, ゼロショット学習の過程で大規模な学習データセットを作成し, 学習させることができる. しかし, 観測劣化 HSI に関する劣化過程の情報 (ノイズ分布など) を正確に推定できない場合, ノイズ除去性能が低下する可能性がある.

## 6.2 3D パターン化マスクを用いたゼロショット HSI ノイズ除去

本節では, 3D パターン化マスクを用いた Blind-spot 戦略, モデルアーキテクチャ, 学習時に使用する損失関数について述べる.

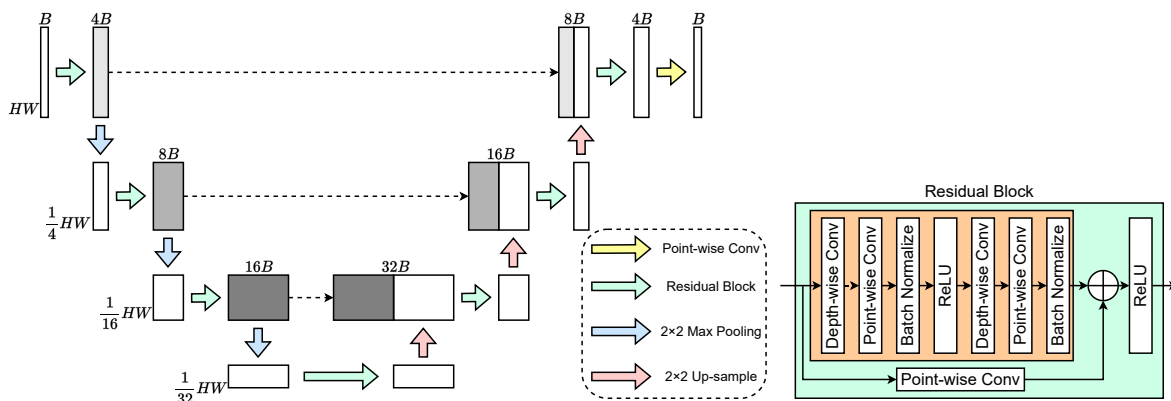


図 6.3: 3D パターン化マスクを用いたゼロショット HSI ノイズ除去におけるモデルアーキテクチャ

### 6.2.1 3D パターン化マスクを用いた学習におけるモデルアーキテクチャ

3D パターン化マスクを用いたゼロショット HSI ノイズ除去におけるモデルアーキテクチャを図 6.3 に示す。図 6.3 において、 $H$ 、 $W$ 、 $B$  は入力  $\mathbf{x} \in \mathbb{R}^{H \times W \times B}$  の縦、横、バンド数の大きさを表す。図中に記載されている添え字は、モデルの中間特徴の大きさの変化を示す。モデルの各処理は図の中央に記述されている。2×2 Max pooling, または 2×2 Up sample は中間出力の空間サイズを変更し、Residual block 内における Point-wise convolution は、チャンネル数を  $C_{in}$  から  $C_{out}$  に変更する。

基本構造は U-Net [97] をベースとし、Separable convolution, および Skip connection を採用している。Separable convolution は、モデル作成における必要パラメータ数削減と、ゼロショット HSI 復元において有効に機能することが第 5 章にて示されている。また、Deep hs prior の研究において、U-Net ベースのゼロショット HSI 復元では 3 次元畳み込みと比較して 2 次元畳み込みの方が精度の面で優れていることが示されているため、本モデルでも採用している。

本モデルでは入力される観測劣化 HSI のサイズに合わせて深層学習モデルの中間チャンネル数を設定する。U-Net では処理の途中において中間層のチャンネル数が変化するため、特徴マップ同士の和を計算することができない。そのため、Skip connection において Point-wise convolution を用いてチャンネル数の調整を行っている。

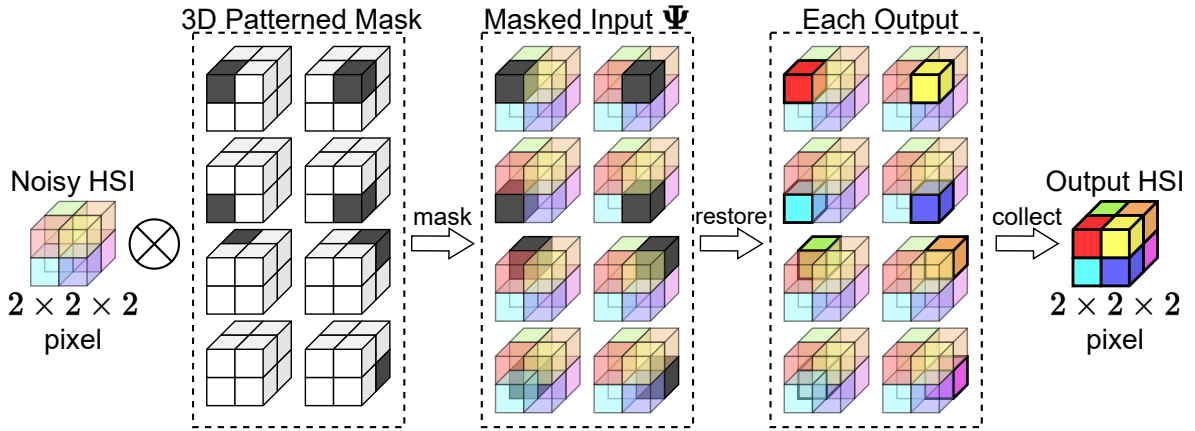


図 6.4: 3D パターン化マスクを用いたゼロショット HSI ノイズ除去における Blind-spot 戦略

## 6.2.2 3D パターン化マスクを用いた Blind-spot 戦略

提案手法における 3 次元パターン化マスクを用いた Blind-spot 戦略は図 6.4 によってなっている。提案手法の Blind-spot 戦略は Blind2Unblind の Blind-spot 戦略を 3 次元方向に拡張したようなものになっている。観測劣化 HSI の  $2 \times 2 \times 2$  の領域ごとに異なる Blind-spot をもつマスクを適用することで、8 つのマスク済み HSI を入手する。それぞれのマスク済み HSI を深層学習モデルに入力し、復元されたピクセルを結合することで、出力を得る。マスクを 3 次元化することにより、空間方向の情報だけでなく、スペクトル方向の情報を利用して HSI 復元を行うよう学習することが可能になる。

## 6.2.3 損失関数

本手法では、Blind2Unblind [86] と同様に、マスク済み HSI と観測劣化 HSI を用いて、以下の損失関数を最小化することを目的とする。

$$L(\mathbf{y}) = \|c\{f(\Psi; \Theta)\} + \hat{f}(\mathbf{y}; \Theta) - (\alpha + 1)\mathbf{y}\|_2^2 + \beta\|c\{f(\Psi; \Theta)\} - \mathbf{y}\|_2^2 \quad (6.1)$$

ここで、 $\Psi$  は複数のマスク済み HSI、 $c(\cdot)$  はモデル出力から Blind-spot の画素を抽出して結合する関数、 $\hat{f}(\cdot; \Theta)$  は誤差逆伝播なしでモデル出力を得る関数、 $\alpha$  と  $\beta$  はハイパーパラメータである。右項  $\beta\|c\{f(\Psi; \Theta)\} - \mathbf{y}\|_2^2$  は学習を安定化させるための正則化項であり、 $\beta$  はこの正則化の強さを決める。本論文では、 $\alpha$  と  $\beta$  のチューニングは行わず、デ

フォルト値の  $\alpha = \beta = 1$  を採用する。復元画像は、Blind2Unblind とは異なり、マスク済み HSI と観測劣化 HSI からの復元出力の重み付き和から得られる。また、提案手法は学習中にハイパーパラメータ  $\alpha$  を変更しない。式 (6.1) において、マスク済み HSI から得られる出力と、観測劣化 HSI を直接モデルに入力することで得られる出力が、それぞれ観測劣化 HSI である  $\mathbf{y}$  に近くなるようにモデルを学習する。

## 6.3 3D パターン化マスクを用いたゼロショットノイズ除去に関する実験

### 6.3.1 実験設定

本節では、本提案手法の有効性を示すために、3D パターン化マスクを用いたゼロショット HSI ノイズ除去実験を行う。実験では、式 (2.2) のノイズとして、標準偏差  $\sigma = 0.1, 0.15$  の白色ガウシアンノイズを用いた。また、提案する Blind-spot 戦略の有効性について検証するため、関連する手法である Blind2Unblind の Blind-spot 戦略とも比較を行った。ノイズ除去に関する実験結果は第 6.3.2 小節、提案する Blind-spot 戦略の有効性については第 6.3.3 小節に記載する。

提案モデルの学習では、入力 HSI からパッチを繰り返し切り出し、Data augmentation としてランダムな左右反転と、0, 90, 180, 240 度の回転を適用する。バッチサイズは 8、初期学習率は  $1.0 \cdot 10^{-3}$ 、エポック数は 100 とし、20 エポックごとに半減する。学習時には観測劣化ハイパースペクトル画像  $\mathbf{y} \in \mathbb{R}^{H \times W \times B}$  から  $40 \times 40 \times B$  のパッチを繰り返し抽出し使用した。従来手法と同様に、オプティマイザには Adam [71] を用いた。

テスト画像には、HSI ノイズ除去実験でよく用いられる Frisco, Stanford, PaviaC, PaviaU を  $[0, 1]$  に正規化して用いる。PaviaC の画像はサイズが大きいため、画像の一部を切り取って実験に使用した。それぞれの画像のサイズは表 6.1 に記載されている。比較

表 6.1: 実験に使用する HSI のサイズ

Images	Sizes
Frisco	$300 \times 300 \times 148$
Stanford	$290 \times 300 \times 148$
PaviaC	$600 \times 400 \times 102$
PaviaU	$610 \times 340 \times 103$

手法としては以下の4つのゼロショットベースの手法を用いる。最適化ベースの手法としては、一般的な3Dデータに対するガウシアンノイズ除去手法であるBM4D [72]と、高精度なガウシアンノイズ除去手法であるFastHyDe [38]を用いた。深層学習ベースの手法としては、Deep image priorに基づくHSI復元手法であるDeep hs prior [96]と、第5章で提案した手法（第5章提案手法）を採用する。BM4D, FastHyDe, Deep hs priorはそれぞれ著者らによって提供されているコードを一部修正し、ハイパーパラメータをチューニングして使用した。

評価指標としては、画像全体の誤差を評価するためにPeak Signal-to-Noise Ratio (PSNR)とERGAS [98] (Relative Dimensionless Global Error in Synthesis)、画像の空間構造の類似度を評価するためにStructural Similarity Index [80] (SSIM)、スペクトル方向の類似度を評価するためにSpectral Angle Mapper [99] (SAM)を用いた。

### 6.3.2 ガウシアンノイズ除去実験

本節では白色ガウシアンノイズ除去実験について述べる。実験結果を表6.2、および表6.3に示す。それぞれの表の評価指標の横に記載されている矢印は、評価指標の値が高い（あるいは低い）ほど復元精度が高いことを表している。表中の太字の値はその実験設定で最も良い復元精度を表し、下線の値は2番目に良い復元精度を表す。

表6.2の結果から、標準偏差 $\sigma = 0.1$ において、3Dパターン化マスクを用いた本提案手法は他のゼロショットベースの手法よりも復元精度が高いことがわかる。また、表6.3の結果から、標準偏差 $\sigma = 0.15$ において劣化過程に関する情報を利用している従来手法とほぼ同等の復元精度が得られることがわかる。SSIMとSAMの両方が従来手法より優れていることから、提案手法が空間方向とスペクトル方向のどちらかの誤差を過度に減らすように学習しているのではなく、バランスの良いノイズ除去結果を達成していることを示している。

また、提案手法で用いられているSeparable convolutionを通常の2次元畳み込みに変更し、Friscoデータで学習を行った。この場合、パラメータのサイズは約227.58Mから約1.47Gに増加したが、過学習によってPSNRは2.79dB減少した。これはSeparable convolutionがモデルの表現力を低下させることなく、過学習の抑制に貢献していることを示している。

表 6.2: 標準偏差  $\sigma = 0.1$  における 3D パターン化マスクを用いたガウシアンノイズ除去結果

Images	Methods	PSNR $\uparrow$	SSIM $\uparrow$	SAM $\downarrow$	ERGAS $\downarrow$
Frisco	BM4D	37.52	$9.728 \cdot 10^{-1}$	$7.251 \cdot 10^{-2}$	44.15
	FastHyDe	41.80	$9.893 \cdot 10^{-1}$	$4.682 \cdot 10^{-2}$	27.64
	Deep hs prior	39.17	$9.812 \cdot 10^{-1}$	$6.149 \cdot 10^{-2}$	36.24
	第 5 章提案手法	<u>41.81</u>	<u><math>9.896 \cdot 10^{-1}</math></u>	<u><math>4.484 \cdot 10^{-2}</math></u>	<u>26.88</u>
	Ours	<b>42.07</b>	<b><math>9.901 \cdot 10^{-1}</math></b>	<b><math>4.327 \cdot 10^{-2}</math></b>	<b>26.13</b>
Stanford	BM4D	37.84	$9.722 \cdot 10^{-1}$	$1.238 \cdot 10^{-1}$	60.41
	FastHyDe	42.22	$9.892 \cdot 10^{-1}$	$8.493 \cdot 10^{-2}$	36.27
	Deep hs prior	39.59	$9.814 \cdot 10^{-1}$	$1.119 \cdot 10^{-1}$	48.73
	第 5 章提案手法	<u>42.48</u>	<b><math>9.905 \cdot 10^{-1}</math></b>	<u><math>7.656 \cdot 10^{-2}</math></u>	<u>35.07</u>
	Ours	<b>42.65</b>	<b><math>9.905 \cdot 10^{-1}</math></b>	<b><math>7.533 \cdot 10^{-2}</math></b>	<b>34.81</b>
PaviaC	BM4D	35.56	$9.637 \cdot 10^{-1}$	$1.044 \cdot 10^{-1}$	46.08
	FastHyDe	37.33	<u><math>9.773 \cdot 10^{-1}</math></u>	$8.460 \cdot 10^{-2}$	37.53
	Deep hs prior	35.58	$9.669 \cdot 10^{-1}$	$1.025 \cdot 10^{-1}$	45.80
	第 5 章提案手法	<u>37.52</u>	$9.772 \cdot 10^{-1}$	<u><math>8.296 \cdot 10^{-2}</math></u>	<u>36.77</u>
	Ours	<b>37.85</b>	<b><math>9.790 \cdot 10^{-1}</math></b>	<b><math>7.973 \cdot 10^{-2}</math></b>	<b>35.37</b>
PaviaU	BM4D	35.67	$9.671 \cdot 10^{-1}$	$9.053 \cdot 10^{-2}$	37.99
	FastHyDe	37.28	<u><math>9.783 \cdot 10^{-1}</math></u>	$7.451 \cdot 10^{-2}$	31.53
	Deep hs prior	35.69	$9.693 \cdot 10^{-1}$	$8.887 \cdot 10^{-2}$	37.75
	第 5 章提案手法	<u>37.31</u>	$9.773 \cdot 10^{-1}$	<u><math>7.447 \cdot 10^{-2}</math></u>	<u>31.50</u>
	Ours	<b>37.72</b>	<b><math>9.794 \cdot 10^{-1}</math></b>	<b><math>7.085 \cdot 10^{-2}</math></b>	<b>30.04</b>

### 6.3.3 3D パターン化マスクの有効性に関する検証

本節では、提案手法における Blind-spot 戦略を Blind2Unblind で提案されている Blind-spot 戦略に置き換えた場合の結果と比較することで、提案した 3D パターン化マスクの有効性を検証する。第 6.3.2 小節の実験で使用した 4 つのデータに対する評価指標の平均値を表 6.4 に示す。

表 6.3: 標準偏差  $\sigma = 0.15$  における 3D パターン化マスクを用いたガウシアンノイズ除去  
結果

Images	Methods	PSNR $\uparrow$	SSIM $\uparrow$	SAM $\downarrow$	ERGAS $\downarrow$
Frisco	BM4D	35.52	$9.592 \cdot 10^{-1}$	$9.036 \cdot 10^{-2}$	55.26
	FastHyDe	39.64	$9.834 \cdot 10^{-1}$	$5.927 \cdot 10^{-2}$	35.23
	Deep hs prior	36.32	$9.607 \cdot 10^{-1}$	$8.772 \cdot 10^{-2}$	49.68
	第 5 章提案手法	<u>39.88</u>	<b><math>9.860 \cdot 10^{-1}</math></b>	<b><math>5.354 \cdot 10^{-2}</math></b>	<b>33.02</b>
	Ours	<b>39.95</b>	<u><math>9.843 \cdot 10^{-1}</math></u>	<u><math>5.489 \cdot 10^{-2}</math></u>	<u>33.97</u>
Stanford	BM4D	35.81	$9.573 \cdot 10^{-1}$	$1.544 \cdot 10^{-1}$	76.04
	FastHyDe	<u>40.29</u>	$9.834 \cdot 10^{-1}$	$1.036 \cdot 10^{-1}$	45.30
	Deep hs prior	37.46	$9.667 \cdot 10^{-1}$	$1.539 \cdot 10^{-1}$	62.55
	第 5 章提案手法	40.28	<b><math>9.861 \cdot 10^{-1}</math></b>	<u><math>9.955 \cdot 10^{-2}</math></u>	<b>44.40</b>
	Ours	<b>40.42</b>	<u><math>9.853 \cdot 10^{-1}</math></u>	<b><math>9.462 \cdot 10^{-2}</math></b>	<u>44.50</u>
PaviaC	BM4D	33.46	$9.435 \cdot 10^{-1}$	$1.324 \cdot 10^{-1}$	58.69
	FastHyDe	35.50	$9.665 \cdot 10^{-1}$	$1.040 \cdot 10^{-1}$	46.31
	Deep hs prior	33.93	$9.504 \cdot 10^{-1}$	$1.242 \cdot 10^{-1}$	55.42
	第 5 章提案手法	<b>36.04</b>	<b><math>9.694 \cdot 10^{-1}</math></b>	<b><math>9.563 \cdot 10^{-2}</math></b>	<b>43.50</b>
	Ours	<u>35.97</u>	<u><math>9.687 \cdot 10^{-1}</math></u>	<u><math>9.837 \cdot 10^{-2}</math></u>	<u>43.90</u>
PaviaU	BM4D	33.62	$9.499 \cdot 10^{-1}$	$1.142 \cdot 10^{-1}$	48.10
	FastHyDe	35.34	$9.680 \cdot 10^{-1}$	$9.259 \cdot 10^{-2}$	39.39
	Deep hs prior	33.80	$9.522 \cdot 10^{-1}$	$1.107 \cdot 10^{-1}$	46.95
	第 5 章提案手法	<b>35.95</b>	<b><math>9.718 \cdot 10^{-1}</math></b>	<b><math>8.412 \cdot 10^{-2}</math></b>	<b>36.77</b>
	Ours	<u>35.78</u>	<u><math>9.683 \cdot 10^{-1}</math></u>	<u><math>8.851 \cdot 10^{-2}</math></u>	<u>37.51</u>

表 6.4 から、標準偏差  $\sigma = 0.1, 0.15$  の両方において 3D パターン化マスクによる Blind-spot 戦略が優れていることが見て取れる。標準偏差  $\sigma = 0.1$  において、Blind2Unblind の Blind-spot 戦略を用いた場合の Frisco の PSNR は 41.02 dB であり、提案手法の PSNR より約 1.0dB 低い。さらに、Blind2Unblind の Blind-Spot 戦略を用いた Stanford の PSNR は 31.14 dB であり、過学習により提案手法の PSNR より約 11.5 dB 低い。この結果は、提案するゼロショット HSI 復元における Blind-Spot 戦略は、Blind2Unblind と比較して、HSI のスペクトル情報を利用することで復元精度を向上させ、入力するデータ

表 6.4: 3D パターン化マスクによる Blind-spot 戦略の有効性に関する検証

$\sigma$	Blind-spot 戦略	PSNR	SSIM	SAM	ERGAS
0.1	Blind2Unblind [86]	36.01	$9.471 \cdot 10^{-1}$	$1.402 \cdot 10^{-1}$	61.22
	Ours	<b>40.07</b>	<b><math>9.847 \cdot 10^{-1}</math></b>	<b><math>6.729 \cdot 10^{-2}</math></b>	<b>31.58</b>
0.15	Blind2Unblind [86]	33.19	$8.787 \cdot 10^{-1}$	$2.066 \cdot 10^{-1}$	100.6
	Ours	<b>37.99</b>	<b><math>9.764 \cdot 10^{-1}</math></b>	<b><math>8.446 \cdot 10^{-2}</math></b>	<b>40.14</b>

数を増やすことで過学習を抑制することを示している。

## 6.4 ランダム 3D パターン化マスクを用いたゼロショット HSI 復元

本節ではゼロショット学習におけるロバスト性を高めるために、第 6.2 節の手法を拡張する。3D パターン化マスクにおける Blind-spot の選択をランダム化することによって、データを増強し、よりロバストな学習を可能にする。

### 6.4.1 ランダム 3D パターン化マスクを用いた学習におけるモデルアーキテクチャ

ランダム 3D パターン化マスクにおけるモデルアーキテクチャは、図 6.5 のようになっている。このモデルアーキテクチャは第 6.2.1 小節におけるモデルを一部変更したものである。提案したモデルアーキテクチャでは HSI のバンド数によってチャンネル数を変更せず、すべての HSI に対して固定した。これは、入力 HSI の保有バンド数が大きい場合において、モデルの中間層のチャンネル数が膨大になることを防ぐためである。また、同時に Skip connection を削除している。これは、Point-wise convolution によるパラメータ数増加を抑えると同時に、モデルが観測劣化 HSI に近い画像を出力しやすくなることを防ぐためである。

### 6.4.2 ランダム 3D パターン化マスクを用いた Blind-spot 戦略

ランダム 3D パターン化マスクのイメージ図を図 6.6 に示す。第 6.2.2 小節における図 6.4 では、 $2 \times 2 \times 2$  の領域ごとに異なる Blind-spot を持つ 8 種類のマスクを使用してい



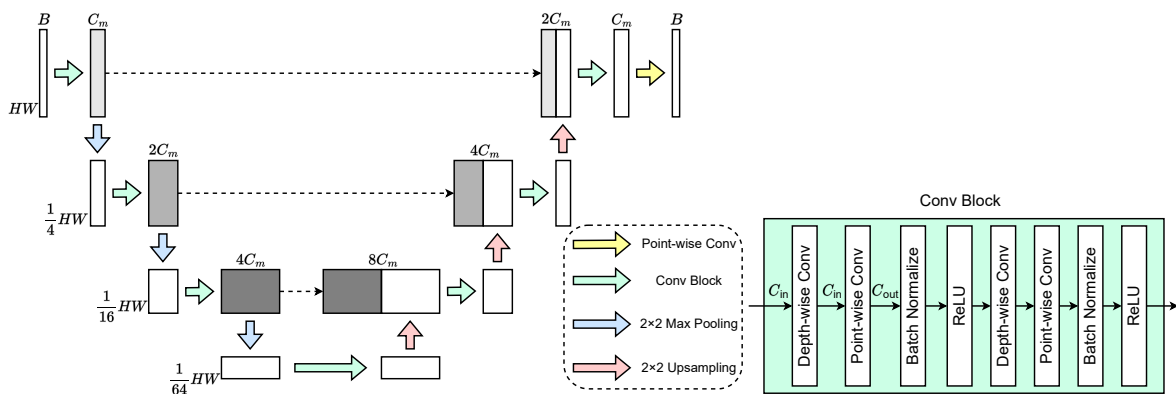


図 6.5: ランダム 3D パターン化マスクを用いたゼロショット HSI ノイズ除去におけるモデルアーキテクチャ

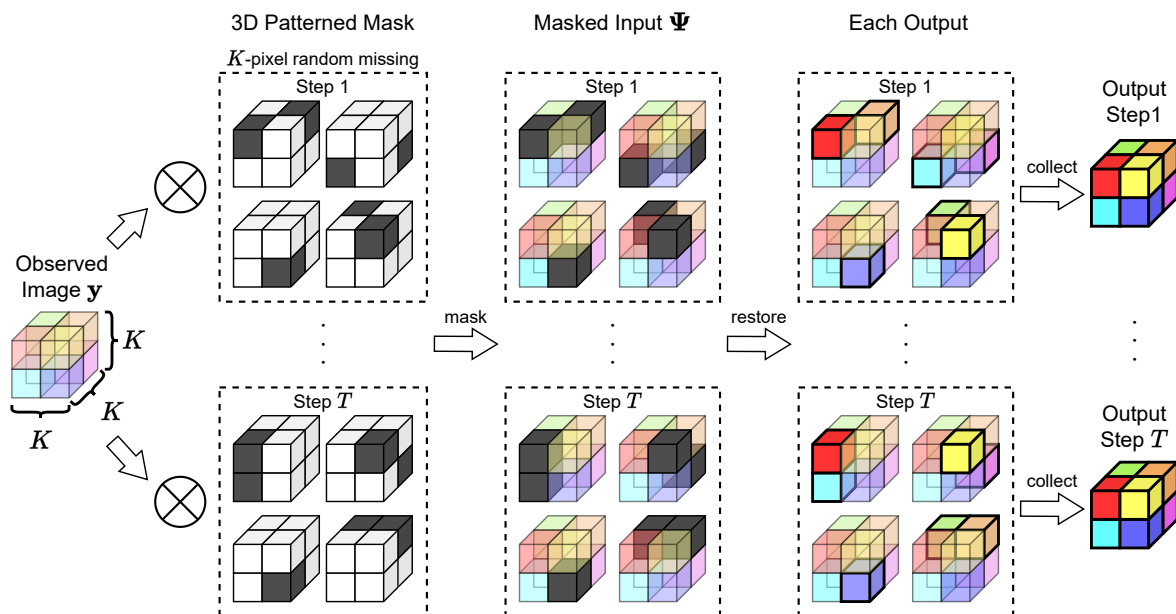


図 6.6: 3D パターン化マスクを用いたゼロショット HSI ノイズ除去における Blind-spot 戦略

た。これに対し，ランダム 3D パターン化マスクでは，入力 HSI 中の  $K \times K \times K$  の領域から  $K$  箇所の Blind-spot の組み合わせをランダムに選択する．例えば  $K = 2$  の場合，1 回の学習で 4 つの 3D パターン化マスクが使用され，各マスクには 2 箇所の Blind-spot が含まれる．Blind-spot をランダムに選択することで，学習データ量を増やし，過学習を抑制する．3D パターン化マスクの時と同様に，Blind-spot の位置を選択する際，他の

マスクと重複しないようにする必要がある。さらに、ランダム 3D パターン化マスクでは Blind-spot の組み合わせが学習ステップごとにランダムに変更される。

得られた各 3D パターン化マスクは、入力サイズに合わせて複製され、観測劣化 HSI と要素ごとに乗算し、マスク済み HSI が作成される。その後、3D パターン化マスクと同様の手順で出力を得る。K の値を大きくすると、復元精度は向上するが、マスク済み HSI の数が増えるため、計算コストが増大する。マスクを 3D パターン化マスクから変更することの有効性については第 6.5.3 小節で後述する。観測劣化 HSI の劣化が激しい場合、モデルが過学習する可能性が高くなる。この問題に対処するため、 $l_2$  重み減衰 (Weight-decay) をモデルパラメータに適用し、過学習を抑制する。

## 6.5 ランダム 3D パターン化マスクに関する実験

### 6.5.1 実験設定

本節では、いくつかのノイズ除去実験を通して、提案手法の有効性を実証する。さらに、提案するモデルアーキテクチャと Blind-spot 戦略の有効性を検証するために、アブレーション研究を行った。モデルアーキテクチャのアブレーション研究では、ランダム 3D パターン化マスクにおける通常の畳み込みと Separable convolution の復元精度の違いを調べた。Blind-spot 戦略のアブレーション研究では、3次元に拡張した Blind2Unblind [86] の Blind-spot 戦略、3D パターン化マスクによる Blind-spot 戦略、ランダム 3D パターン化マスクによる Blind-spot 戦略を比較する。ノイズ除去実験は第 6.5.2 小節、アブレーション研究は第 6.5.3 小節で述べる。

学習時のバッチサイズは 24 とした。提案モデルの中間チャンネル数  $C_m$  は 200 とした。初期学習率は  $1.0 \cdot 10^{-3}$ 、エポック数は 200 とした。学習率は 40 エポック毎に半減させた。K は  $K = 3$  に設定し、それぞれ実験を行った。さらに、ハイパーパラメータの  $\alpha$  と  $\beta$  の両方を 1 に設定した。その他の設定はすべて第 6.3.1 小節と同様である。

### 6.5.2 ノイズ除去実験

ここでは、様々な種類のノイズを用いたノイズ除去実験の結果を示す。ノイズ除去実験には、平均値 0、標準偏差  $\sigma$  のガウシアンノイズ、平均値 0、標準偏差  $\sigma$  のラプラシアンノイズ、画素値に依存するポアソンノイズ、および現実世界の劣化 HSI を用いた。標準偏差  $\sigma$  は 0.1 と 0.15 とした。ガウシアンノイズ、およびラプラシアンノイズの劣化過程は式 (2.2) に従う。

表 6.5: 標準偏差  $\sigma = 0.1$  におけるランダム化 3D パターン化マスクを用いたガウシアンノイズ除去実験結果

Images	Methods	PSNR $\uparrow$	SSIM $\uparrow$	SAM $\downarrow$	ERGAS $\downarrow$
Frisco	BM4D	37.52	$9.728 \cdot 10^{-1}$	$7.251 \cdot 10^{-2}$	44.15
	FastHyDe	41.80	$9.893 \cdot 10^{-1}$	$4.682 \cdot 10^{-2}$	27.64
	Deep hs prior	39.17	$9.812 \cdot 10^{-1}$	$6.149 \cdot 10^{-2}$	36.24
	第 5 章提案手法	<u>41.81</u>	<u><math>9.896 \cdot 10^{-1}</math></u>	<u><math>4.484 \cdot 10^{-2}</math></u>	<u>26.88</u>
	Ours	<b>42.25</b>	<b><math>9.903 \cdot 10^{-1}</math></b>	<b><math>3.486 \cdot 10^{-2}</math></b>	<b>20.90</b>
Stanford	BM4D	37.84	$9.722 \cdot 10^{-1}$	$1.238 \cdot 10^{-1}$	60.41
	FastHyDe	42.22	$9.892 \cdot 10^{-1}$	$8.493 \cdot 10^{-2}$	<u>36.27</u>
	Deep hs prior	39.59	$9.814 \cdot 10^{-1}$	$1.119 \cdot 10^{-1}$	48.73
	第 5 章提案手法	<u>42.48</u>	<b><math>9.905 \cdot 10^{-1}</math></b>	<u><math>7.656 \cdot 10^{-2}</math></u>	<b>35.07</b>
	Ours	<b>42.59</b>	<u><math>9.901 \cdot 10^{-1}</math></u>	<b><math>4.704 \cdot 10^{-2}</math></b>	40.75
PaviaC	BM4D	35.56	$9.637 \cdot 10^{-1}$	$1.044 \cdot 10^{-1}$	46.08
	FastHyDe	37.33	<u><math>9.773 \cdot 10^{-1}</math></u>	$8.460 \cdot 10^{-2}$	<u>37.53</u>
	Deep hs prior	35.58	$9.669 \cdot 10^{-1}$	$1.025 \cdot 10^{-1}$	45.80
	第 5 章提案手法	<u>37.52</u>	$9.772 \cdot 10^{-1}$	<u><math>8.296 \cdot 10^{-2}</math></u>	<b>36.77</b>
	Ours	<b>37.72</b>	<b><math>9.781 \cdot 10^{-1}</math></b>	<b><math>8.213 \cdot 10^{-2}</math></b>	41.70
PaviaU	BM4D	35.67	$9.671 \cdot 10^{-1}$	$9.053 \cdot 10^{-2}$	37.99
	FastHyDe	37.28	<u><math>9.783 \cdot 10^{-1}</math></u>	$7.451 \cdot 10^{-2}$	<u>31.53</u>
	Deep hs prior	35.69	$9.693 \cdot 10^{-1}$	$8.887 \cdot 10^{-2}$	37.75
	第 5 章提案手法	<u>37.31</u>	$9.773 \cdot 10^{-1}$	<u><math>7.447 \cdot 10^{-2}</math></u>	<b>31.50</b>
	Ours	<b>37.73</b>	<b><math>9.786 \cdot 10^{-1}</math></b>	<b><math>6.012 \cdot 10^{-1}</math></b>	35.56

ガウシアンノイズ除去実験の結果を表 6.5, および表 6.6 に示す. 表 6.5, 表 6.6 の結果から, 提案手法は 3D パターン化マスクの場合と同様に, 他のゼロショット法と比べて同等以上の復元精度を持つことがわかる.

表 6.5 から, 標準偏差 0.1 において提案手法は Frisco では全ての指標で, PaviaC と PaviaU では 3 つの指標で, Stanford では 2 つの指標で従来手法を上回った. Stanford のデータでは, 提案手法は SSIM において 2 番目に優れた復元精度を示している.

また, 表 6.6 から, 標準偏差は 0.15 において提案手法は Stanford データを除き, 従来

表 6.6: 標準偏差  $\sigma = 0.15$  におけるランダム化 3D パターン化マスクを用いたガウシアンノイズ除去実験結果

Images	Methods	PSNR $\uparrow$	SSIM $\uparrow$	SAM $\downarrow$	ERGAS $\downarrow$
Frisco	BM4D	35.52	$9.592 \cdot 10^{-1}$	$9.036 \cdot 10^{-2}$	55.26
	FastHyDe	39.64	$9.834 \cdot 10^{-1}$	$5.927 \cdot 10^{-2}$	35.23
	Deep hs prior	36.32	$9.607 \cdot 10^{-1}$	$8.772 \cdot 10^{-2}$	49.68
	第 5 章提案手法	<u>39.88</u>	<b><math>9.860 \cdot 10^{-1}</math></b>	<u><math>5.354 \cdot 10^{-2}</math></u>	<u>33.02</u>
	Ours	<b>40.28</b>	<u><math>9.852 \cdot 10^{-1}</math></u>	<b><math>4.103 \cdot 10^{-2}</math></b>	<b>25.64</b>
Stanford	BM4D	35.81	$9.573 \cdot 10^{-1}$	$1.544 \cdot 10^{-1}$	76.04
	FastHyDe	<b>40.29</b>	<u><math>9.834 \cdot 10^{-1}</math></u>	$1.036 \cdot 10^{-1}$	<u>45.30</u>
	Deep hs prior	37.46	$9.667 \cdot 10^{-1}$	$1.539 \cdot 10^{-1}$	62.55
	第 5 章提案手法	<u>40.28</u>	<b><math>9.861 \cdot 10^{-1}</math></b>	<u><math>9.955 \cdot 10^{-2}</math></u>	<b>44.40</b>
	Ours	40.18	$9.831 \cdot 10^{-1}$	<b><math>5.985 \cdot 10^{-2}</math></b>	53.81
PaviaC	BM4D	33.46	$9.435 \cdot 10^{-1}$	$1.324 \cdot 10^{-1}$	58.69
	FastHyDe	35.50	$9.665 \cdot 10^{-1}$	$1.040 \cdot 10^{-1}$	<u>46.31</u>
	Deep hs prior	33.93	$9.504 \cdot 10^{-1}$	$1.242 \cdot 10^{-1}$	55.42
	第 5 章提案手法	<u>36.04</u>	<u><math>9.694 \cdot 10^{-1}</math></u>	<u><math>9.563 \cdot 10^{-2}</math></u>	<b>43.50</b>
	Ours	<b>36.41</b>	<b><math>9.726 \cdot 10^{-1}</math></b>	<b><math>8.245 \cdot 10^{-2}</math></b>	46.82
PaviaU	BM4D	33.62	$9.499 \cdot 10^{-1}$	$1.142 \cdot 10^{-1}$	48.10
	FastHyDe	35.34	$9.680 \cdot 10^{-1}$	$9.259 \cdot 10^{-2}$	<u>39.39</u>
	Deep hs prior	33.80	$9.522 \cdot 10^{-1}$	$1.107 \cdot 10^{-1}$	46.95
	第 5 章提案手法	<u>35.95</u>	<b><math>9.718 \cdot 10^{-1}</math></b>	<u><math>8.412 \cdot 10^{-2}</math></u>	<b>36.77</b>
	Ours	<b>36.20</b>	<u><math>9.717 \cdot 10^{-1}</math></u>	<b><math>6.544 \cdot 10^{-2}</math></b>	41.01

手法よりも良好な復元精度を示した。提案手法は、Frisco と PaviaC では 3 つの指標で、PaviaU では 2 つの指標で、Stanford では 1 つの指標で従来手法を上回った。Frisco と PaviaU に対する提案手法は、SSIM の観点で 2 番目に高い復元精度を示した。

つぎに、標準偏差 0.1 と 0.15 の白色ラプラシアンノイズによって劣化した観測画像におけるノイズ除去精度を比較し、その実験結果を表 6.7, および表 6.8 に示す。表 6.7 および表 6.8 は、提案手法が従来手法を上回っていることを示している。具体的には、標準偏差 0.1 を用いた場合、提案手法は全てのデータにおいて 3 つの評価指標で従来手法を

表 6.7: 標準偏差  $\sigma = 0.1$  におけるランダム化 3D パターン化マスクを用いたラプラシアンノイズ除去実験結果

Images	Methods	PSNR $\uparrow$	SSIM $\uparrow$	SAM $\downarrow$	ERGAS $\downarrow$
Frisco	BM4D	30.46	$8.561 \cdot 10^{-1}$	$1.734 \cdot 10^{-1}$	95.72
	FastHyDe	39.90	$9.845 \cdot 10^{-1}$	$5.755 \cdot 10^{-2}$	34.10
	Deep hs prior	37.57	$9.737 \cdot 10^{-1}$	$7.366 \cdot 10^{-2}$	43.64
	第 5 章提案手法	<u>40.61</u>	<b><math>9.881 \cdot 10^{-1}</math></b>	<u><math>5.011 \cdot 10^{-2}</math></u>	<u>31.22</u>
	Ours	<b>40.65</b>	<u><math>9.864 \cdot 10^{-1}</math></u>	<b><math>3.921 \cdot 10^{-2}</math></b>	<b>24.60</b>
Stanford	BM4D	30.91	$8.628 \cdot 10^{-1}$	$3.207 \cdot 10^{-1}$	133.8
	FastHyDe	<u>40.58</u>	<u><math>9.844 \cdot 10^{-1}</math></u>	<u><math>1.002 \cdot 10^{-1}</math></u>	<b>43.78</b>
	Deep hs prior	37.72	$9.713 \cdot 10^{-1}$	$1.408 \cdot 10^{-1}$	60.57
	第 5 章提案手法	40.12	$9.841 \cdot 10^{-1}$	$1.033 \cdot 10^{-1}$	<u>45.89</u>
	Ours	<b>40.72</b>	<b><math>9.851 \cdot 10^{-1}</math></b>	<b><math>5.497 \cdot 10^{-2}</math></b>	50.00
PaviaC	BM4D	28.98	$8.355 \cdot 10^{-1}$	$2.196 \cdot 10^{-1}$	98.27
	FastHyDe	35.75	$9.683 \cdot 10^{-1}$	$1.011 \cdot 10^{-1}$	<u>45.00</u>
	Deep hs prior	34.24	$9.548 \cdot 10^{-1}$	$1.194 \cdot 10^{-1}$	53.44
	第 5 章提案手法	<u>36.55</u>	<u><math>9.726 \cdot 10^{-1}</math></u>	<u><math>9.202 \cdot 10^{-2}</math></u>	<b>41.08</b>
	Ours	<b>36.62</b>	<b><math>9.737 \cdot 10^{-1}</math></b>	<b><math>8.118 \cdot 10^{-2}</math></b>	46.03
PaviaU	BM4D	29.01	$8.419 \cdot 10^{-1}$	$1.943 \cdot 10^{-1}$	81.90
	FastHyDe	35.59	$9.697 \cdot 10^{-1}$	$8.994 \cdot 10^{-2}$	<u>38.27</u>
	Deep hs prior	34.03	$9.553 \cdot 10^{-1}$	$1.079 \cdot 10^{-1}$	45.67
	第 5 章提案手法	<u>36.29</u>	<u><math>9.730 \cdot 10^{-1}</math></u>	<u><math>8.237 \cdot 10^{-2}</math></u>	<b>35.38</b>
	Ours	<b>36.53</b>	<b><math>9.736 \cdot 10^{-1}</math></b>	<b><math>6.348 \cdot 10^{-2}</math></b>	39.70

上回り, 提案手法による Frisco の復元結果は SSIM の観点で 2 番目に優れた復元精度であった. 標準偏差を 0.15 とした場合, PaviaC と PaviaU は 3 つの評価指標で, Frisco と Stanford は 2 つの評価指標で優れた復元精度を示した. さらに, Frisco の結果は PSNR と SSIM の点で 2 番目に優れた復元精度を示し, Stanford の結果は SSIM の点で 2 番目に優れた結果を示した.

つぎにポアソンノイズによって劣化した観測画像におけるノイズ除去結果を示す. この

表 6.8: 標準偏差  $\sigma = 0.15$  におけるランダム化 3D パターン化マスクを用いたラプラシアンノイズ除去実験結果

Images	Methods	PSNR $\uparrow$	SSIM $\uparrow$	SAM $\downarrow$	ERGAS $\downarrow$
Frisco	BM4D	28.00	$7.690 \cdot 10^{-1}$	$2.278 \cdot 10^{-1}$	125.5
	FastHyDe	37.96	$9.765 \cdot 10^{-1}$	$7.068 \cdot 10^{-2}$	42.66
	Deep hs prior	35.54	$9.588 \cdot 10^{-1}$	$9.328 \cdot 10^{-2}$	54.88
	第 5 章提案手法	<u>38.61</u>	<b><math>9.806 \cdot 10^{-1}</math></b>	<u><math>6.320 \cdot 10^{-2}</math></u>	<u>39.39</u>
	Ours	<b>38.63</b>	<u><math>9.784 \cdot 10^{-1}</math></u>	<b><math>4.611 \cdot 10^{-2}</math></b>	<b>30.95</b>
Stanford	BM4D	28.51	$7.794 \cdot 10^{-1}$	$4.136 \cdot 10^{-1}$	176.3
	FastHyDe	<u>38.49</u>	<u><math>9.758 \cdot 10^{-1}</math></u>	<u><math>1.237 \cdot 10^{-1}</math></u>	<b>55.56</b>
	Deep hs prior	35.66	$9.554 \cdot 10^{-1}$	$1.770 \cdot 10^{-1}$	76.77
	第 5 章提案手法	38.40	$9.790 \cdot 10^{-1}$	$1.237 \cdot 10^{-1}$	<u>55.69</u>
	Ours	<b>38.60</b>	<b><math>9.783 \cdot 10^{-1}</math></b>	<b><math>5.578 \cdot 10^{-2}</math></b>	58.54
PaviaC	BM4D	26.70	$7.494 \cdot 10^{-1}$	$2.828 \cdot 10^{-1}$	127.7
	FastHyDe	33.87	$9.532 \cdot 10^{-1}$	$1.252 \cdot 10^{-1}$	<u>55.92</u>
	Deep hs prior	32.25	$9.296 \cdot 10^{-1}$	$1.498 \cdot 10^{-1}$	67.23
	第 5 章提案手法	<u>33.31</u>	<u><math>9.363 \cdot 10^{-1}</math></u>	<u><math>1.345 \cdot 10^{-1}</math></u>	<b>59.73</b>
	Ours	<b>34.47</b>	<b><math>9.590 \cdot 10^{-1}</math></b>	<b><math>9.623 \cdot 10^{-2}</math></b>	57.24
PaviaU	BM4D	26.67	$7.550 \cdot 10^{-1}$	$2.526 \cdot 10^{-1}$	107.2
	FastHyDe	33.89	$9.566 \cdot 10^{-1}$	$1.091 \cdot 10^{-1}$	<u>46.52</u>
	Deep hs prior	32.12	$9.302 \cdot 10^{-1}$	$1.343 \cdot 10^{-1}$	57.02
	第 5 章提案手法	<u>33.82</u>	<u><math>9.497 \cdot 10^{-1}</math></u>	<u><math>1.091 \cdot 10^{-1}</math></u>	<b>47.08</b>
	Ours	<b>34.37</b>	<b><math>9.600 \cdot 10^{-1}</math></b>	<b><math>7.467 \cdot 10^{-2}</math></b>	50.00

実験では、観察された劣化画像を得るために以下の式を利用した。

$$\mathbf{y} = P(s \cdot \mathbf{x})/s \quad (6.2)$$

ここで、 $P(s \cdot \mathbf{x})$  は平均  $s \cdot \mathbf{x}$  のポアソン分布に基づいたノイズを適用する関数であり、 $s$  はノイズの強度を決定する定数である。式 (6.2) で示されているように、ノイズは画像によって異なる。この実験で使用されたノイズ強度の定数は  $s = 15$  および  $6$  であり、それぞれの結果は表 6.9、および表 6.10 に示されている。他の実験で従来の方法として使用さ

表 6.9: ノイズ強度定数  $s = 15$  におけるランダム化 3D パターン化マスクを用いたポアソンノイズ除去実験結果

Noise strength constant $s$		15			
Images	Methods	PSNR $\uparrow$	SSIM $\uparrow$	SAM $\downarrow$	ERGAS $\downarrow$
Frisco	FastHyDe	40.14	$9.895 \cdot 10^{-1}$	$4.866 \cdot 10^{-2}$	30.96
	Deep hs prior	38.88	$9.808 \cdot 10^{-1}$	$6.228 \cdot 10^{-2}$	36.50
	第 5 章提案手法	<u>41.76</u>	<b><math>9.915 \cdot 10^{-1}</math></b>	<u><math>4.405 \cdot 10^{-2}</math></u>	<u>26.13</u>
	Ours	<b>41.96</b>	<u><math>9.906 \cdot 10^{-1}</math></u>	<b><math>3.042 \cdot 10^{-2}</math></b>	<b>20.87</b>
Stanford	FastHyDe	41.63	<u><math>9.909 \cdot 10^{-1}</math></u>	$7.250 \cdot 10^{-2}$	38.24
	Deep hs prior	40.22	$9.848 \cdot 10^{-1}$	$9.696 \cdot 10^{-2}$	45.20
	第 5 章提案手法	<u>42.08</u>	$9.908 \cdot 10^{-1}$	<u><math>7.213 \cdot 10^{-2}</math></u>	<u>36.02</u>
	Ours	<b>42.84</b>	<b><math>9.925 \cdot 10^{-1}</math></b>	<b><math>3.537 \cdot 10^{-2}</math></b>	<b>31.82</b>
PaviaC	FastHyDe	36.46	$9.773 \cdot 10^{-1}$	$8.554 \cdot 10^{-2}$	41.30
	Deep hs prior	35.38	$9.666 \cdot 10^{-1}$	$1.034 \cdot 10^{-1}$	46.80
	第 5 章提案手法	<u>37.11</u>	<u><math>9.784 \cdot 10^{-1}</math></u>	<u><math>8.016 \cdot 10^{-2}</math></u>	<b>38.40</b>
	Ours	<b>37.66</b>	<b><math>9.801 \cdot 10^{-1}</math></b>	<b><math>7.005 \cdot 10^{-2}</math></b>	<u>40.37</u>
PaviaU	FastHyDe	36.10	$9.765 \cdot 10^{-1}$	$7.722 \cdot 10^{-2}$	35.99
	Deep hs prior	34.97	$9.648 \cdot 10^{-1}$	$9.434 \cdot 10^{-2}$	40.98
	第 5 章提案手法	<u>36.59</u>	<u><math>9.762 \cdot 10^{-1}</math></u>	<u><math>7.218 \cdot 10^{-2}</math></u>	<b>34.08</b>
	Ours	<b>37.23</b>	<b><math>9.783 \cdot 10^{-1}</math></b>	<b><math>5.641 \cdot 10^{-2}</math></b>	<u>35.93</u>

れた BM4D は、他の方法と比べて顕著に低い復元精度であったため、表から除外されている。

表 6.9, および表 6.10 から、本提案手法はすべてのテストデータと実験条件に対して優れた復元精度を持っていることがわかる。  $s = 15$  の場合、我々の方法は SSIM の Frisco, ERGAS の PaviaC と PaviaU を除いて、すべての評価基準で従来の方法を上回った。  $s = 6$  の場合、我々の方法はすべての評価基準で従来の方法よりも優れている。特に、我々の方法は PSNR と SAM の点で従来の方法を大幅に上回り、スペクトル情報の復元精度が従来の方法よりも優れていることを示している。これらの結果は、加法性ノイズまたはピクセル値に依存するノイズを持つ画像の復元に対して、本提案手法の有効性を示している。

表 6.10: ノイズ強度定数  $s = 6$  におけるランダム化 3D パターン化マスクを用いたポアソンノイズ除去実験結果

Noise strength constant $s$		6			
Images	Methods	PSNR $\uparrow$	SSIM $\uparrow$	SAM $\downarrow$	ERGAS $\downarrow$
Frisco	FastHyDe	35.66	$9.797 \cdot 10^{-1}$	$6.512 \cdot 10^{-2}$	49.15
	Deep hs prior	36.53	$9.686 \cdot 10^{-1}$	$8.152 \cdot 10^{-2}$	47.72
	第 5 章提案手法	<u>39.10</u>	<u><math>9.840 \cdot 10^{-1}</math></u>	<u><math>5.670 \cdot 10^{-2}</math></u>	<u>35.16</u>
	Ours	<b>39.69</b>	<b><math>9.842 \cdot 10^{-1}</math></b>	<b><math>3.673 \cdot 10^{-2}</math></b>	<b>26.69</b>
Stanford	FastHyDe	37.70	$9.824 \cdot 10^{-1}$	$9.266 \cdot 10^{-2}$	59.87
	Deep hs prior	37.45	$9.727 \cdot 10^{-1}$	$1.293 \cdot 10^{-1}$	61.79
	第 5 章提案手法	<u>39.30</u>	<u><math>9.844 \cdot 10^{-1}</math></u>	<u><math>8.075 \cdot 10^{-2}</math></u>	<u>50.52</u>
	Ours	<b>40.05</b>	<b><math>9.867 \cdot 10^{-1}</math></b>	<b><math>4.200 \cdot 10^{-2}</math></b>	<b>43.30</b>
PaviaC	FastHyDe	33.34	<u><math>9.618 \cdot 10^{-1}</math></u>	$1.091 \cdot 10^{-1}$	59.07
	Deep hs prior	33.24	$9.463 \cdot 10^{-1}$	$1.304 \cdot 10^{-1}$	59.86
	第 5 章提案手法	<u>34.27</u>	<u><math>9.608 \cdot 10^{-1}</math></u>	<u><math>1.035 \cdot 10^{-1}</math></u>	<u>53.30</u>
	Ours	<b>35.15</b>	<b><math>9.672 \cdot 10^{-1}</math></b>	<b><math>8.666 \cdot 10^{-2}</math></b>	<b>51.86</b>
PaviaU	FastHyDe	32.79	<u><math>9.609 \cdot 10^{-1}</math></u>	$9.947 \cdot 10^{-2}$	52.68
	Deep hs prior	32.76	$9.433 \cdot 10^{-1}$	$1.207 \cdot 10^{-1}$	52.85
	第 5 章提案手法	<u>33.56</u>	<u><math>9.536 \cdot 10^{-1}</math></u>	<u><math>9.389 \cdot 10^{-2}</math></u>	<u>48.37</u>
	Ours	<b>34.78</b>	<b><math>9.653 \cdot 10^{-1}</math></b>	<b><math>6.937 \cdot 10^{-2}</math></b>	<b>46.56</b>

つぎに実世界の劣化 HSI に対するノイズ除去を行う。この実験では、実世界のノイズに対して提案手法と従来手法を適用することにより、パフォーマンスを比較した。この実験のために、実世界のノイズを持つ 4 つの HSI データ (Urban, Botswana, Loukia, Nefeli) を使用した。これらのデータにおいて無劣化 HSI は利用できないため、定量的な評価を使用することができない。そのため、定性的評価を行う。Deep hs prior で早期停止の最適なタイミングを事前に決定するのが難しいため、実験にはデフォルト設定の 1800 エポックを使用した。前章提案手法の手法で用いられる劣化過程に関する情報が利用できないため、劣化過程をガウシアン分布と仮定し、標準偏差を推定するために、Bioucas-Dias らの研究 [100] で述べられている手法を使用した。BM4D と FastHyDe は、他の方法と比べてより良い復元結果を得られなかったため、実験結果から除外されている。



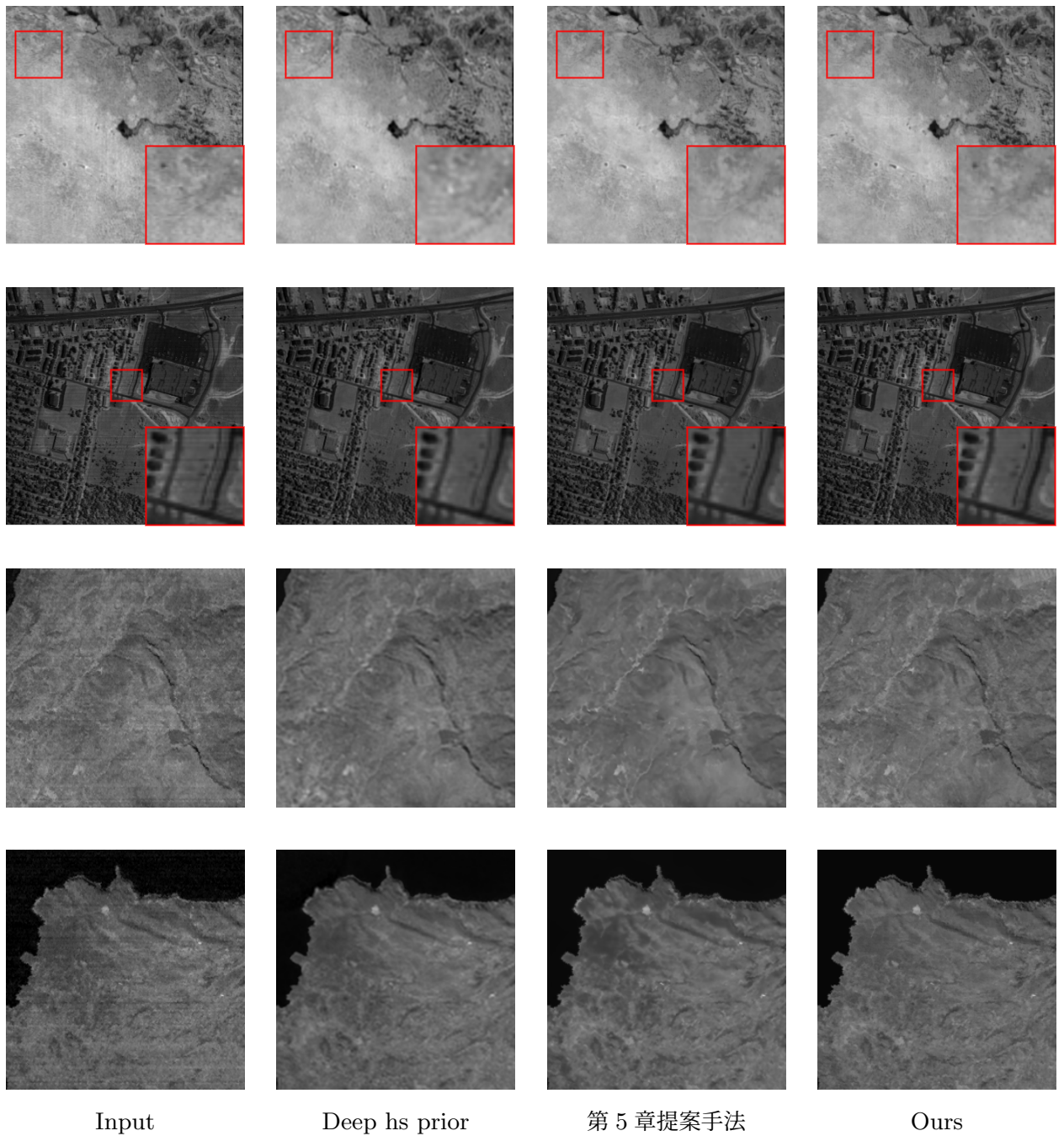


図 6.7: 実世界の劣化 HSI に対するノイズ除去結果

実世界のノイズ除去の結果を，図 6.7 に示す．図 6.7 の画像は上の行から順に，Botswana の 62 バンド目，Urban の 103 バンド目，Loukia の 52 バンド目，Nefeli の 51 バンド目である．表示されている画像は，ノイズ除去された HSI から単一のチャンネルを

表 6.11:  $K = 2$  における畳み込みの比較

$\sigma$	畳み込み手法	PSNR	SSIM	SAM	ERGAS
0.1	Standard	<b>40.04</b>	<b><math>9.853 \cdot 10^{-1}</math></b>	<b><math>5.080 \cdot 10^{-2}</math></b>	<b>33.84</b>
	Separable	39.59	$9.826 \cdot 10^{-1}$	$5.631 \cdot 10^{-2}$	36.06
0.15	Standard	36.24	$9.624 \cdot 10^{-1}$	$9.300 \cdot 10^{-2}$	45.06
	Separable	<b>38.06</b>	<b><math>9.775 \cdot 10^{-1}</math></b>	<b><math>6.160 \cdot 10^{-2}</math></b>	<b>42.19</b>

表 6.12:  $K = 3$  における畳み込みの比較

$\sigma$	畳み込み手法	PSNR	SSIM	SAM	ERGAS
0.1	Standard	<b>40.28</b>	<b><math>9.857 \cdot 10^{-1}</math></b>	<b><math>5.084 \cdot 10^{-2}</math></b>	<b>33.22</b>
	Separable	40.07	$9.843 \cdot 10^{-1}$	$5.596 \cdot 10^{-2}$	34.70
0.15	Standard	35.66	$9.539 \cdot 10^{-1}$	$2.995 \cdot 10^{-1}$	66.56
	Separable	<b>38.27</b>	<b><math>9.781 \cdot 10^{-1}</math></b>	<b><math>6.208 \cdot 10^{-2}</math></b>	<b>41.77</b>

抽出し、それを切り取り、正規化し、表示したものである。図 6.7 の 1 行目および 2 行目の右下の赤い枠で囲まれた拡大画像は、画像中の赤い枠で囲われた部分の拡大図を示している。ノイズ除去された画像は、Deep hs prior がぼやけた結果を生成する傾向があるのに対し、第 5 章提案手法と本提案手法はより鮮明な画像を生成することを示している。画像を個別に調べると、1 行目の画像では、Deep hs prior と第 5 章提案手法は、拡大画像の左上にある黒い穴のようなオブジェクトを復元することに失敗し、周囲のパターンに置き換えられている。2 行目の画像では、画像中の道路が Deep hs prior の復元結果でぼやけて表示されているが、水平方向のストライプ状のノイズはすべての方法で除去されている。3 行目、および 4 行目の画像では、第 5 章提案手法と本提案手法を比較すると、第 5 章提案手法は本提案手法に比べてテクスチャが崩れていることがわかる。これらの結果は、本提案手法が HSI のテクスチャを保持しながら効果的にノイズを除去し、従来の方法を上回ることを示している。

### 6.5.3 ランダム 3D パターン化マスクの有効性に関する検証

本小節では、本提案手法のモデルアーキテクチャと Blind-spot 戦略の有効性を検証するために実施されたアブレーション研究について説明する。モデルアーキテクチャのアブ

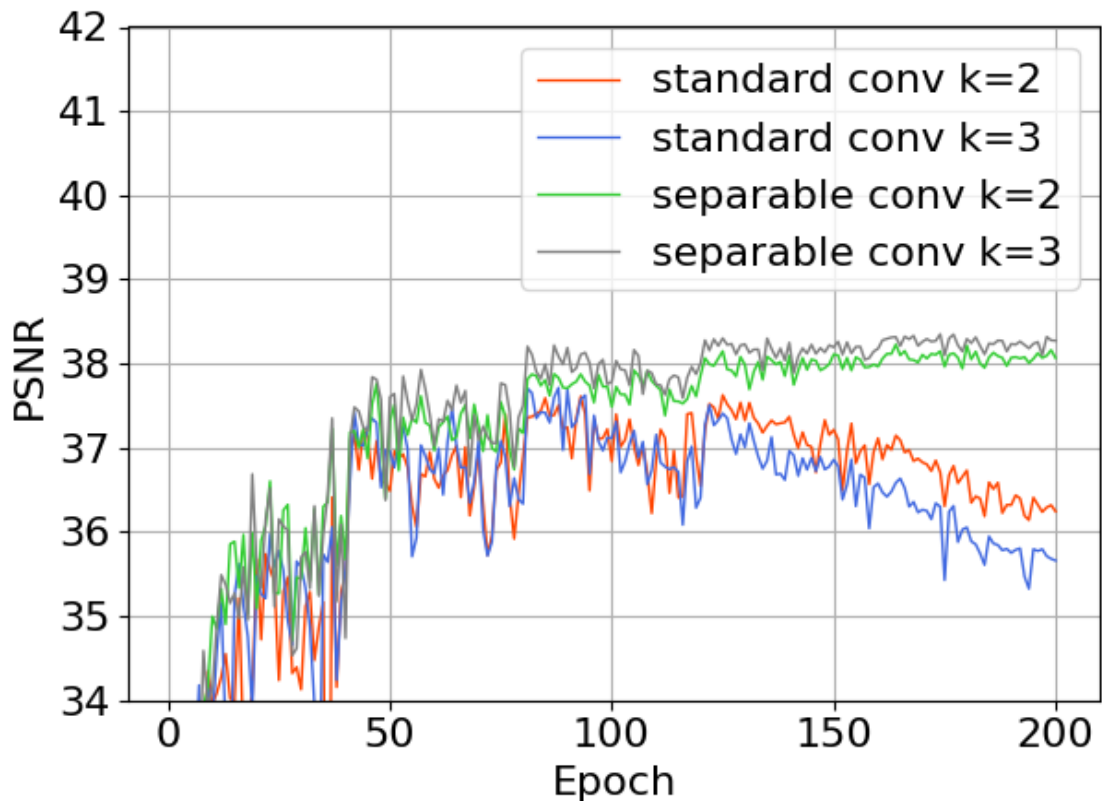


図 6.8: 標準偏差 0.15 のガウシアンノイズ除去におけるそれぞれの畳み込み層を用いたモデルの PSNR プロット

レーション研究は、通常の畳み込み (Standard) と Separable convolution (Separable) の違いを説明する。Blind-spot 戦略の研究は、Blind2Unblind の Blind-spot 戦略の 3D 拡張、第 6.2 節で提案されている 3D パターン化マスク、および本提案手法の 3 つの Blind-spot 戦略を比較した。アブレーション研究は、第 6.5.1 小節で説明されているガウシアンノイズ除去のための実験設定を使用し、4 つのデータ (Frisco, PaviaC, PaviaU, Stanford) を使用する。各実験結果の表では、各テストデータに対する評価指標の平均値を表している。

ランダム 3D パターン化マスクにおける Separable convolution の効果を検証を行う。  $K = 2$  および  $K = 3$  におけるガウシアンノイズ除去実験の結果は、表 6.11 および 表 6.12 に示されており、標準偏差 0.15 における PSNR プロットは 図 6.8 に示されている。

表 6.11 および 表 6.12 の結果から、標準偏差  $\sigma = 0.1$  の場合、通常の畳み込みを使用

したモデルは Separable convolution と同等またはわずかに良い精度を達成したことがわかる。ただし、標準偏差  $\sigma = 0.15$  の場合、Separable convolution を使用したモデルは、通常の畳み込みを持つモデルを上回っている。特に、標準偏差  $\sigma = 0.15$  の場合、通常の畳み込みを使用したモデルは精度が大幅に低下した。

図 6.8 の結果は、標準偏差  $\sigma = 0.15$  の場合、通常の畳み込みを使用するモデルは過学習する傾向があったことを示している。これの一因として、通常の畳み込みを使用したモデルは、パラメータの数かなり多く、 $C_m = 200$  で約 160M あり、Separable convolution を用いたモデルは約 18M のパラメータしか持っていないことが考えられる。通常の畳み込みを使用したモデルは低い標準偏差で高い復元精度を示すが、Separable convolution を用いたモデルは激しく劣化した HSI に対するロバスト性が高い。通常の畳み込みを用いたモデルが低い標準偏差でわずかに改善するだけであるのに対して、Separable convolution を用いたモデルが高い標準偏差においてかなり良い精度を達成することから、Separable convolution の有効性が確認できる。

つぎに、ランダム 3D パターン化マスクへの拡張の有効性を評価するために、Blind2Unblind の Blind-spot 戦略、3D パターン化マスクの Blind-spot 戦略と比較する。

実験結果を、表 6.13 および 図 6.9 に示す。表 6.13 において、B2U は Blind2Unblind、Pixel-wise は 3D パターン化マスク、Ours ( $K = 2$ ) と Ours ( $K = 3$ ) は、それぞれ  $K = 2$  と  $K = 3$  の場合の実験結果を示している。表 6.13 の結果から、ランダム 3D パターン化マスクが最も良い復元精度を達成したことが示されている。3D パターン化マス

表 6.13: ランダム 3D パターン化マスクとその他の Blind-spot 戦略の比較

$\sigma$	Blind-spot 戦略	PSNR	SSIM	SAM	ERGAS
0.1	B2U [86]	37.81	$9.711 \cdot 10^{-1}$	$6.798 \cdot 10^{-2}$	45.89
	Pixel-wise	39.58	$9.816 \cdot 10^{-1}$	$6.175 \cdot 10^{-2}$	37.10
	Ours ( $K = 2$ )	39.59	$9.826 \cdot 10^{-1}$	$5.631 \cdot 10^{-2}$	36.06
	Ours ( $K = 3$ )	<b>40.07</b>	<b><math>9.843 \cdot 10^{-1}</math></b>	<b><math>5.596 \cdot 10^{-2}</math></b>	<b>34.70</b>
0.15	B2U [86]	36.69	$9.715 \cdot 10^{-1}$	$6.285 \cdot 10^{-2}$	47.82
	Pixel-wise	38.14	$9.776 \cdot 10^{-1}$	$6.332 \cdot 10^{-2}$	42.48
	Ours ( $K = 2$ )	38.06	$9.775 \cdot 10^{-1}$	<b><math>6.160 \cdot 10^{-2}</math></b>	42.19
	Ours ( $K = 3$ )	<b>38.27</b>	<b><math>9.781 \cdot 10^{-1}</math></b>	$6.208 \cdot 10^{-2}$	<b>41.77</b>

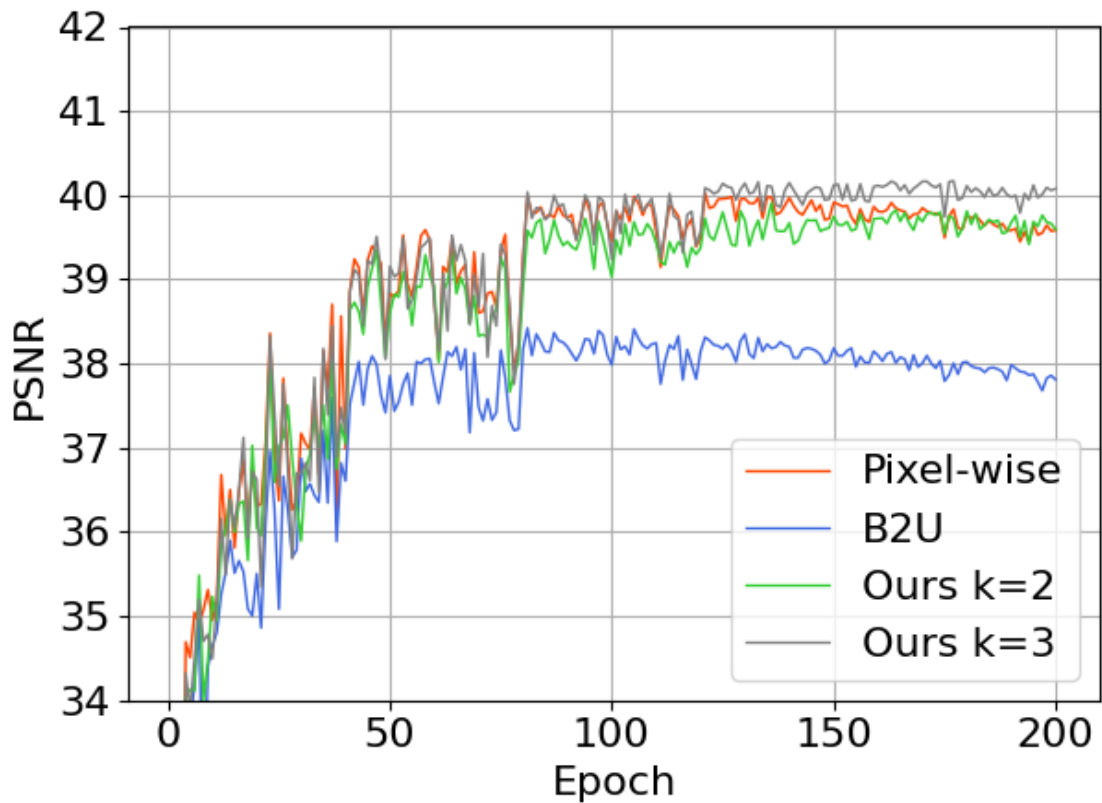


図 6.9: 標準偏差 0.1 のガウシアンノイズ除去における Blind-spot 戦略ごとの PSNR のプロット

ク (Pixel-wise) およびランダム 3D パターン化マスクの復元精度は B2U よりも高く、マスクの 3D 化によるスペクトル情報の組み込みが復元精度の向上に寄与したことを示している。さらに本提案手法は、Pixel-wise 方法が  $K^3$  のマスク ( $K = 2$  の場合は 8 つのマスク) を使用するのに対し、ランダム 3D パターン化マスクでは  $K^2$  のマスク ( $K = 2$  の場合は 4 つのマスク,  $K = 3$  の場合は 9 つのマスク) を利用するため、計算コストの面で優れている。

また、図 6.9 に示されているように、Blind2Unblind の Blind-spot 戦略と 3D パターン化マスクの Blind-spot 戦略における PSNR は、約 120 エポック後に徐々に減少しており、これは過学習していることを示している。しかし、ランダム化 3D パターン化マスクの PSNR は収束しており、ランダム性を取り入れることで学習のロバスト性が向上したことを示唆している。これらの結果は、ランダム 3D パターン化マスクの有効性を示して

いる.

従来の方法と比較して, 本提案手法は計算時間がおおよそ 2 から 3 倍必要である. しかし我々の方法は, 訓練中の最適なモデル選択とハイパーパラメータの選択においてまだ改善の余地が残されている. これらのハイパーパラメータ最適化タスクは, 膨大な計算コストが必要なため, 今回は実行していない.

## 6.6 結論

### 6.6.1 本章のまとめ

本論文では, 深層学習ベースのゼロショット HSI ノイズ除去手法を提案した. モデルのアーキテクチャは U-Net に基づいており, Separable convolution が組み込まれている. 空間的な方向情報に加えて, スペクトル方向情報の情報を利用して復元するために, 3D パターン化マスクを提案した. このマスクでは  $2 \times 2 \times 2$  の領域ごとに欠損位置を変えた 8 つのマスク済み HSI を学習に用いる. また, 学習のロバスト性を向上させるため, Blind-spot の選択をランダム化したランダム 3D パターン化マスクも提案した.  $K^2$  種類の 3D パターンマスクを使用して, 各  $K \times K \times K$  領域で  $K$  ピクセルをランダムに重複なしで欠落させることで, よりロバストな学習を実現する. 提案したゼロショット HSI ノイズ除去手法の優位性を示すために, 他の最適化ベースおよび深層学習ベースのゼロショット HSI 復元方法と比較するいくつかの実験を実施した. また同時に, 提案手法のモデルアーキテクチャおよびブラインドスポット戦略の効果検証を行った. 我々の研究は, 3D マスクがゼロショット HSI 復元に効果的であり, ランダムな Blind-spot の選択がロバストな学習の実現に寄与することを示している.

### 6.6.2 提案手法の課題

Blind-spot ベースのゼロショット学習手法にしたことにより, これまで必要だったエポック数の事前決定, および劣化過程の情報なしに HSI ノイズ除去を行うことができるようになった. しかし, ランダム化 3D パターン化マスクでは学習ステップごとにマスクを変更するため, ゼロショットであるものの計算コストが高く, 学習に時間がかかってしまう. そのため, 復元精度を維持しつつ, より高速で使用メモリ量の少ない手法を開発していく必要がある. また, 本提案手法はノイズ除去以外のタスクに対して適用することができないため, その他のタスクに対しては別の手法を開発する必要がある,

## 第7章

# 結論

本論文では、整備された大規模データセットを利用することが困難なデータであるハイパースペクトル画像 (HSI: Hyperspectral image) に対し、深層学習を用いた高精度なノイズ除去手法を提案することを目的として研究を行った。深層学習モデルは高い関数近似能力を有しており、適切に学習することができれば深層学習ベースの HSI ノイズ除去において高い復元精度が期待できる。しかし、HSI は無劣化画像の利用が困難であること、データごとにバンド数が異なること、データサイズおよびバンド数が膨大であることなどから、モデルの学習が困難であった。

上記の問題を解決するため、我々は事前学習、先験情報の利用、ゼロショット学習の3つの方向性に沿ってノイズ除去モデルを提案した。事前学習を用いる手法では、HSI が2次元的な空間情報と1次元的なスペクトル情報によって構成されていることに着目し、Separable convolution と呼ばれる畳み込みの空間方向のカーネルとチャンネル方向のカーネルを個別に事前学習する手法を提案した。空間方向のカーネルはグレースケール画像を用いて学習し、チャンネル方向のカーネルはスペクトル情報を用いて学習する。この手法では特に劣化が激しい場合において、事前学習が有効に機能することを実験的に示した。しかし、復元対象の HSI の空間的情報やスペクトル情報が事前学習のデータと大きく異なる場合、深層学習モデルの復元精度が低下する恐れがあり、事前学習データに対する画像復元モデル精度の依存性が問題となる可能性がある。

先験情報を利用する手法では、凸最適化における画像復元アーキテクチャを利用している。凸最適化では、復元対象の画像の特性をモデル化し正則化項として利用することで、学習データなしでの画像復元を可能にする。我々は深層学習ベースの画像復元モデルの出力部に先験情報を組み込んだモデルを用いることで、画像復元モデルの出力を制約する。提案手法は画像前処理のための深層学習モデルと、凸最適化を解く反復アルゴリズムを深

層展開によって拡張した深層展開モデルによって構成される。グレースケール画像復元実験において、学習データが 50 枚以下の条件下にて、提案手法が深層学習モデルのみの場合と比較してより高い汎化性能を持つことを示した。しかし、反復アルゴリズムは漸化式の形であるため高速な学習が難しく、また凸最適化において用いられる正則化項の性能がボトルネックになってしまう点が問題点として残っている。

ゼロショット学習をベースとした手法では、観測劣化 HSI の劣化仮定を利用してデータペアを作成し学習する手法と、特定のパターンを用いた Blind-spot 戦略によってデータペアを作成し学習する手法の 2 つを提案した。劣化過程を利用する手法では、観測劣化 HSI を劣化過程に沿ってさらに劣化させた二重劣化 HSI を用いて学習を行う。ゼロショット学習を用いた深層学習ベースの手法、および最適化ベースの手法と比較して、優れた復元精度を示した。Blind-spot 戦略を用いた手法では、3 次元的な領域において異なる欠損位置を持つマスク済み HSI から欠損した画素を復元するように学習する。劣化過程を利用する手法を含むゼロショット学習ベースの手法、および最適化ベースの手法と比較して優れた復元精度を示した。また、Blind-spot をランダムに選択するように拡張した手法も提案した。そちらの手法ではより過学習を抑制する効果があることを実験的に確認した。

これらの研究結果から、我々はゼロショット学習をベースとした手法が最も有望ではないかと考えている。Zhang らの研究 [101] のように、露光時間を変更することで無劣化 HSI に近い画像を取得し、ノイズをモデリングする手法も存在するが、雲などの影響によってそもそも無劣化画像を取得することが困難な状況も考えられる。我々の手法は現実世界のノイズ除去においても有望な結果を示しており、学習ペアの作成方法やモデル駆動型とデータ駆動型を組み合わせるアプローチ [102–105] 等とゼロショット学習を組み合わせることで、よりロバストで精度の高い HSI 復元手法を提案できるようにしていきたい。今後は、精度を維持しつつ学習ペアの作成方法やモデル構造を変更することによる計算量の削減や高速化を行うことが課題であると考えられる。



# 謝辞

本論文は、私が北九州市立大学国際環境工学部情報メディア工学科、北九州市立大学大学院国際環境工学研究科、同志社大学大学院理工学研究科に在籍していた期間に行った研究をまとめたものである。本研究を遂行するにあたり、多くの方々に多大なるご指導、ご支援を賜りましたことを心より感謝申し上げます。

指導教員である奥田正浩教授には、研究の進め方、研究アイデアに関するディスカッションなどの研究に関する指導をはじめ、研究成果を投稿するにあたっての論文執筆指導などについて丁寧にご指導いただきましたことを心より感謝申し上げます。

北九州市立大学の松岡諒准教授は、修士課程2年からの1年間、研究や論文執筆に関する指導をしていただきました。学術論文誌への投稿にあたり、提出書類のご指導をいただきましたことを感謝いたします。

工学院大学の京地清介准教授には、修士論文に関するアドバイスや、国際会議でのサポートなどをしていただきました。一人での国際会議参加の際にお声がけいただいたことが非常にありがたく感謝申し上げます。

また、日頃から研究に関するディスカッションや発表練習、国内学会や国際学会などの発表で協力いただきました、北九州市立大学の奥田研究室、松岡研究室、同志社大学知的機構研究室的の諸氏にお礼を申し上げます。その中でも、プログラム作成や研究を進めていくにあたって熱心にアドバイスをしてくださった今村竜司氏に深く感謝申し上げます。

最後に博士後期課程への進学に際し、励ましの言葉や援助をしていただいた両親に感謝いたします。

## 参考文献

- [1] Joachim Hill, Henning Buddenbaum, and Philip A Townsend. Imaging spectroscopy of forest ecosystems: perspectives for the use of space-borne hyperspectral earth observation systems. *Surveys in Geophysics*, Vol. 40, No. 3, pp. 553–588, 2019.
- [2] Liangfu Chen, Husi Letu, Meng Fan, Huazhe Shang, Jinhua Tao, Laixiong Wu, Ying Zhang, Chao Yu, Jianbin Gu, Ning Zhang, et al. An introduction to the chinese high-resolution earth observation system: Gaofen-1~7 civilian satellites. *Journal of Remote Sensing*, 2022.
- [3] Sima Peyghambari and Yun Zhang. Hyperspectral remote sensing in lithological mapping, mineral exploration, and environmental geology: an updated review. *Journal of Applied Remote Sensing*, Vol. 15, No. 3, pp. 031501–031501, 2021.
- [4] R Booyesen, S Lorenz, M Kirsch, R Jackisch, R Zimmermann, and R Gloaguen. Hyperspectral imaging with uavs for mineral exploration. In *Second EAGE Workshop on Unmanned Aerial Vehicles*, Vol. 2021, pp. 1–3. European Association of Geoscientists & Engineers, 2021.
- [5] Asgeir Bjorgan and Lise Lyngsnes Randeberg. Towards real-time medical diagnostics using hyperspectral imaging technology. In *European Conference on Biomedical Optics*, p. 953712. Optica Publishing Group, 2015.
- [6] Guolan Lu and Baowei Fei. Medical hyperspectral imaging: a review. *Journal of biomedical optics*, Vol. 19, No. 1, pp. 010901–010901, 2014.
- [7] Ndubisi A Aviara, Jacob Tizhe Liberty, Ojo S Olatunbosun, Habib A Shoyombo, and Samuel K Oyeniya. Potential application of hyperspectral imaging in food grain quality inspection, evaluation and control during bulk storage. *Journal of Agriculture and Food Research*, Vol. 8, p. 100288, 2022.

- [8] S Karthika Shree, Vaishali Vijayarajan, B Sathya Bama, and S Mohammed Mansoor Roomi. Milk quality inspection using hyperspectral imaging. In *2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT)*, pp. 1–6. IEEE, 2023.
- [9] Hongyan Zhang, Jiayi Li, Yuancheng Huang, and Liangpei Zhang. A nonlocal weighted joint sparse representation classification method for hyperspectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 7, No. 6, pp. 2056–2065, 2014.
- [10] Xiangyong Cao, Feng Zhou, Lin Xu, Deyu Meng, Zongben Xu, and John Paisley. Hyperspectral image classification with markov random fields and a convolutional neural network. *IEEE Transactions on image processing*, Vol. 27, No. 5, pp. 2354–2367, 2018.
- [11] Akrem Sellami and Salvatore Tabbone. Deep neural networks-based relevant latent representation learning for hyperspectral image classification. *Pattern Recognition*, Vol. 121, p. 108224, 2022.
- [12] Swalpa Kumar Roy, Ankur Deria, Danfeng Hong, Behnood Rasti, Antonio Plaza, and Jocelyn Chanussot. Multimodal fusion transformer for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 2023.
- [13] Jos V’e MBIoucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gader, Jocelyn Chanussot. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE journal of selected topics in applied earth observations and remote sensing*, Vol. 5, No. 2, pp. 354–379, 2012.
- [14] David WJ Stein, Scott G Beaven, Lawrence E Hoff, Edwin M Winter, Alan P Schaum, and Alan D Stocker. Anomaly detection from hyperspectral imagery. *IEEE signal processing magazine*, Vol. 19, No. 1, pp. 58–69, 2002.
- [15] Mia Rizkinia and Masahiro Okuda. Joint local abundance sparse unmixing for hyperspectral images. *Remote Sensing*, Vol. 9, No. 12, p. 1224, 2017.
- [16] Mia Rizkinia, Tatsuya Baba, Keiichiro Shirai, and Masahiro Okuda. Local spectral component decomposition for multi-channel image denoising. *IEEE Transactions on Image Processing*, Vol. 25, No. 7, pp. 3208–3218, 2016.
- [17] H Zhang. Hyperspectral image denoising with cubic total variation model.

- ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 7, pp. 95–98, 2012.
- [18] Qiangqiang Yuan, Liangpei Zhang, and Huanfeng Shen. Hyperspectral image denoising employing a spectral–spatial adaptive total variation model. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 50, No. 10, pp. 3660–3677, 2012.
- [19] Stamatiou Lefkimmatis, Anastasios Roussos, Petros Maragos, and Michael Unser. Structure tensor total variation. *SIAM Journal on Imaging Sciences*, Vol. 8, No. 2, pp. 1090–1122, 2015.
- [20] VB Surya Prasath, Dmitry Vorotnikov, Rengarajan Pelapur, Shani Jose, Guna Seetharaman, and Kannappan Palaniappan. Multiscale tikhonov-total variation image restoration using spatially varying edge coherence exponent. *IEEE Transactions on Image Processing*, Vol. 24, No. 12, pp. 5220–5235, 2015.
- [21] Shunsuke Ono, Keiichiro Shirai, and Masahiro Okuda. Vectorial total variation based on arranged structure tensor for multichannel image restoration. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pp. 4528–4532. IEEE, 2016.
- [22] Hongyan Zhang, Wei He, Liangpei Zhang, Huanfeng Shen, and Qiangqiang Yuan. Hyperspectral image restoration using low-rank matrix recovery. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 52, No. 8, pp. 4729–4743, 2014.
- [23] Hemant Kumar Aggarwal and Angshul Majumdar. Hyperspectral image denoising using spatio-spectral total variation. *IEEE Geoscience and Remote Sensing Letters*, Vol. 13, No. 3, pp. 442–446, 2016.
- [24] Wei He, Hongyan Zhang, Liangpei Zhang, and Huanfeng Shen. Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 54, No. 1, pp. 178–188, 2016.
- [25] Saori Takeyama, Shunsuke Ono, and Itsuo Kumazawa. Hyperspectral image restoration by hybrid spatio-spectral total variation. In *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, pp. 4586–4590. IEEE, 2017.
- [26] Ryuji Kurihara, Shunsuke Ono, Keiichiro Shirai, and Masahiro Okuda. Hyper-

- spectral image restoration based on spatio-spectral structure tensor regularization. In *Signal Processing Conference (EUSIPCO), 2017 25th European*, pp. 488–492. IEEE, 2017.
- [27] Ryota Yuzuriha, Ryuji Kurihara, Ryo Matsuoka, and Masahiro Okuda. Tnng: Total nuclear norms of gradients for hyperspectral image prior. *Remote Sensing*, Vol. 13, No. 4, p. 819, 2021.
- [28] Viren Jain and Sebastian Seung. Natural image denoising with convolutional networks. *Advances in neural information processing systems*, Vol. 21, , 2008.
- [29] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. *IEEE Transactions on Image Processing*, Vol. 27, No. 9, pp. 4608–4622, 2018.
- [30] Xiaojiao Mao, Chunhua Shen, and Yu-Bin Yang. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. *Advances in neural information processing systems*, Vol. 29, , 2016.
- [31] Ying Tai, Jian Yang, Xiaoming Liu, and Chunyan Xu. Memnet: A persistent memory network for image restoration. In *Proceedings of the IEEE international conference on computer vision*, pp. 4539–4547, 2017.
- [32] Saeed Anwar and Nick Barnes. Real image denoising with feature attention. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 3155–3164, 2019.
- [33] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5728–5739, 2022.
- [34] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, Vol. 26, No. 7, pp. 3142–3155, 2017.
- [35] Yunjin Chen and Thomas Pock. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 39, No. 6, pp. 1256–1272, 2016.
- [36] Yang Liu, Zhenyue Qin, Saeed Anwar, Pan Ji, Dongwoo Kim, Sabrina Caldwell, and Tom Gedeon. Invertible denoising network: A light solution for real noise removal. In *Proceedings of the IEEE/CVF conference on computer vision and*

- pattern recognition*, pp. 13365–13374, 2021.
- [37] Yao Wang, Jiangjun Peng, Qian Zhao, Yee Leung, Xi-Le Zhao, and Deyu Meng. Hyperspectral image restoration via total variation regularized low-rank tensor decomposition. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 11, No. 4, pp. 1227–1243, 2018.
- [38] Lina Zhuang and José M Bioucas-Dias. Fast hyperspectral image denoising and inpainting based on low-rank and sparse representations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 11, No. 3, pp. 730–742, 2018.
- [39] Daniel Gabay and Bertrand Mercier. A dual algorithm for the solution of non-linear variational problems via finite element approximation. *Computers & mathematics with applications*, Vol. 2, No. 1, pp. 17–40, 1976.
- [40] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, Vol. 40, pp. 120–145, 2011.
- [41] Tatsuki Itasaka, Ryo Matsuoka, and Masahiro Okuda. Constrained design of fir filters with sparse coefficients. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. 104, No. 11, pp. 1499–1508, 2021.
- [42] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pp. 448–456. pmlr, 2015.
- [43] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016.
- [44] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807–814, 2010.
- [45] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, Vol. 15, No. 1, pp. 1929–1958, 2014.
- [46] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learn-

- ing for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.
- [47] Shen Cheng, Yuzhi Wang, Haibin Huang, Donghao Liu, Haoqiang Fan, and Shuaicheng Liu. Nbnnet: Noise basis learning for image denoising with subspace projection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4896–4906, 2021.
- [48] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
- [49] Weihao Yu, Mi Luo, Pan Zhou, Chenyang Si, Yichen Zhou, Xinchao Wang, Jiashi Feng, and Shuicheng Yan. Metaformer is actually what you need for vision. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10819–10829, 2022.
- [50] David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, Vol. 2, pp. 416–423. IEEE, 2001.
- [51] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 126–135, 2017.
- [52] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3291–3300, 2018.
- [53] Abdelrahman Abdelhamed, Stephen Lin, and Michael S Brown. A high-quality denoising dataset for smartphone cameras. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1692–1700, 2018.
- [54] Jun Xu, Hui Li, Zhetong Liang, David Zhang, and Lei Zhang. Real-world noisy image denoising: A new benchmark. *arXiv preprint arXiv:1804.02603*, 2018.
- [55] Benoit Brummer and Christophe De Vleeschouwer. Natural image noise dataset. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 0–0, 2019.

- [56] Yoonsik Kim, Jae Woong Soh, Gu Yong Park, and Nam Ik Cho. Transfer learning from synthetic to real-noise denoising with adaptive instance normalization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 3482–3492, 2020.
- [57] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1712–1722, 2019.
- [58] Jingwen Chen, Jiawei Chen, Hongyang Chao, and Ming Yang. Image blind denoising with generative adversarial network based noise modeling. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3155–3164, 2018.
- [59] Kai Zhang, Yawei Li, Wangmeng Zuo, Lei Zhang, Luc Van Gool, and Radu Timofte. Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, No. 10, pp. 6360–6376, 2021.
- [60] Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. Learning deep cnn denoiser prior for image restoration. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3929–3938, 2017.
- [61] Kai Zhang, Yawei Li, Wangmeng Zuo, Lei Zhang, Luc Van Gool, and Radu Timofte. Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, No. 10, pp. 6360–6376, 2021.
- [62] Stanley H Chan, Xiran Wang, and Omar A Elgendy. Plug-and-play admm for image restoration: Fixed-point convergence and applications. *IEEE Transactions on Computational Imaging*, Vol. 3, No. 1, pp. 84–98, 2016.
- [63] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of naacL-HLT*, Vol. 1, p. 2, 2019.
- [64] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, Vol. 33, pp. 1877–1901, 2020.
- [65] Hanting Chen, Yunhe Wang, Tianyu Guo, Chang Xu, Yiping Deng, Zhenhua



- Liu, Siwei Ma, Chunjing Xu, Chao Xu, and Wen Gao. Pre-trained image processing transformer. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12299–12310, 2021.
- [66] Kataoka Hirokatsu, Matsumoto Asato, Yamagata Eisuke, Yamada Ryosuke, Inoue Nakamasa, Akio Nakamura, and Satoh Yutaka. Pre-training without natural images. *International Journal of Computer Vision*, Vol. 130, No. 4, pp. 990–1007, 2022.
- [67] Emmanuel Asiedu Brempong, Simon Kornblith, Ting Chen, Niki Parmar, Matthias Minderer, and Mohammad Norouzi. Denoising pretraining for semantic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4175–4186, 2022.
- [68] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255. Ieee, 2009.
- [69] Aidan Nicholas Gomez, Lukasz Mieczyslaw Kaiser, and Francois Chollet. Depth-wise separable convolutions for neural machine translation, December 1 2020. US Patent 10,853,590.
- [70] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251–1258, 2017.
- [71] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [72] Matteo Maggioni, Vladimir Katkovnik, Karen Egiazarian, and Alessandro Foi. Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE transactions on image processing*, Vol. 22, No. 1, pp. 119–133, 2012.
- [73] Jonas Adler and Ozan Öktem. Learned primal-dual reconstruction. *IEEE transactions on medical imaging*, Vol. 37, No. 6, pp. 1322–1332, 2018.
- [74] Filippos Kokkinos and Stamatios Lefkimmiatis. Deep image demosaicking using a cascade of convolutional residual denoising networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 303–319, 2018.
- [75] Uwe Schmidt and Stefan Roth. Shrinkage fields for effective image restoration. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2774–2781, 2014.

- [76] Oren Solomon, Regev Cohen, Yi Zhang, Yi Yang, Qiong He, Jianwen Luo, Ruud JG van Sloun, and Yonina C Eldar. Deep unfolded robust pca with application to clutter suppression in ultrasound. *IEEE transactions on medical imaging*, Vol. 39, No. 4, pp. 1051–1063, 2019.
- [77] Jian Sun, Huibin Li, Zongben Xu, et al. Deep admm-net for compressive sensing mri. *Advances in neural information processing systems*, Vol. 29, , 2016.
- [78] Kai Zhang, Luc Van Gool, and Radu Timofte. Deep unfolding network for image super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 3217–3226, 2020.
- [79] Adam Coates, Andrew Ng, and Honglak Lee. An analysis of single-layer networks in unsupervised feature learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pp. 215–223. JMLR Workshop and Conference Proceedings, 2011.
- [80] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, Vol. 13, No. 4, pp. 600–612, 2004.
- [81] J Lehtinen, J Munkberg, J Hasselgren, S Laine, Tero Karras, M Aittala, and Timo Aila. Noise2Noise: Learning image restoration without clean data. *ICML*, 2018.
- [82] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug. Noise2void-learning denoising from single noisy images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2129–2137, 2019.
- [83] Jason Lequyer, Reuben Philip, Amit Sharma, Wen-Hsin Hsu, and Laurence Pelletier. A fast blind zero-shot denoiser. *Nature Machine Intelligence*, Vol. 4, No. 11, pp. 953–963, 2022.
- [84] Nick Moran, Dan Schmidt, Yu Zhong, and Patrick Coady. Noisier2noise: Learning to denoise from unpaired noisy data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12064–12072, 2020.
- [85] Tao Huang, Songjiang Li, Xu Jia, Huchuan Lu, and Jianzhuang Liu. Neighbor2neighbor: a self-supervised framework for deep image denoising, 2022.
- [86] Zejin Wang, Jiazheng Liu, Guoqing Li, and Hua Han. Blind2unblind: Self-supervised image denoising with visible blind spots. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2027–

2036, 2022.

- [87] Zhilu Zhang, Rongjian Xu, Ming Liu, Zifei Yan, and Wangmeng Zuo. Self-supervised image restoration with blurry and noisy pairs. *Advances in Neural Information Processing Systems*, Vol. 35, pp. 29179–29191, 2022.
- [88] Adria Font Calvarons. Improved noise2noise denoising with limited data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 796–805, 2021.
- [89] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 9446–9454, 2018.
- [90] Yuhui Quan, Mingqin Chen, Tongyao Pang, and Hui Ji. Self2self with dropout: Learning self-supervised denoising from single image. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1890–1898, 2020.
- [91] Youssef Mansour and Reinhard Heckel. Zero-shot noise2noise: Efficient image denoising without any data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14018–14027, 2023.
- [92] Zhi-Qin John Xu, Yaoyu Zhang, Tao Luo, Yanyang Xiao, and Zheng Ma. Frequency principle: Fourier analysis sheds light on deep neural networks. *arXiv preprint arXiv:1901.06523*, 2019.
- [93] Zhi-Qin John Xu, Yaoyu Zhang, and Yanyang Xiao. Training behavior of deep neural network in frequency domain. In *Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12–15, 2019, Proceedings, Part I 26*, pp. 264–274. Springer, 2019.
- [94] Chaewon Kim, Jaeho Lee, and Jinwoo Shin. Zero-shot blind image denoising via implicit neural representations. *arXiv preprint arXiv:2204.02405*, 2022.
- [95] Pravendra Singh, Vinay Kumar Verma, Piyush Rai, and Vinay P Namboodiri. Hetconv: Heterogeneous kernel-based convolutions for deep cnns. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4835–4844, 2019.
- [96] Oleksii Sidorov and Jon Yngve Hardeberg. Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pp. 0–0, 2019.

- [97] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pp. 234–241. Springer, 2015.
- [98] Qian Du, Nicholas H Younan, Roger King, and Vijay P Shah. On the performance evaluation of pan-sharpening techniques. *IEEE Geoscience and Remote Sensing Letters*, Vol. 4, No. 4, pp. 518–522, 2007.
- [99] Roberta H Yuhas, Alexander FH Goetz, and Joe W Boardman. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm. In *JPL, Summaries of the Third Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop*, 1992.
- [100] José M Bioucas-Dias and José MP Nascimento. Hyperspectral subspace identification. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No. 8, pp. 2435–2445, 2008.
- [101] Tao Zhang, Ying Fu, and Cheng Li. Hyperspectral image denoising with realistic data. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2248–2257, 2021.
- [102] Qiang Zhang, Qiangqiang Yuan, Jie Li, Fujun Sun, and Liangpei Zhang. Deep spatio-spectral bayesian posterior for hyperspectral image non-iid noise removal. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 164, pp. 125–137, 2020.
- [103] Qiang Zhang, Qiangqiang Yuan, Meiping Song, Haoyang Yu, and Liangpei Zhang. Cooperated spectral low-rankness prior and deep spatial prior for hsi unsupervised denoising. *IEEE Transactions on Image Processing*, Vol. 31, pp. 6356–6368, 2022.
- [104] Qiang Zhang, Yaming Zheng, Qiangqiang Yuan, Meiping Song, Haoyang Yu, and Yi Xiao. Hyperspectral image denoising: From model-driven, data-driven, to model-data-driven. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.
- [105] Qiang Zhang, Yushuai Dong, Qiangqiang Yuan, Meiping Song, and Haoyang Yu. Combined deep priors with low-rank tensor factorization for hyperspectral image restoration. *IEEE Geoscience and Remote Sensing Letters*, Vol. 20, pp.

1–5, 2023.