

修士論文

機械学習による CNT フォレスト断面画像の解析と
光反射率に影響を及ぼす構造の推定

Machine Learning-Based Cross-Sectional Structural Analysis of
CNT Forests and Prediction of Structural Factors Affecting
Optical Reflectance

報告者

学籍番号: 1275051

氏名: 神生龍一

指導教員

古田 寛 教授

令和7年2月18日

高知工科大学大学院工学研究科

基盤工学専攻電子・光工学コース

内容

第1章	序論	1
1.1	研究背景	1
1.1.1	カーボンナノチューブの構造と物性の関係	1
1.1.2	機械学習の材料開発の分野への応用	2
1.2	目的	2
1.3	本論文の構成	3
第2章	使用した機械学習アルゴリズム	4
2.1	主成分分析	4
2.2	混合ガウス法	5
2.3	ニューラルネットワーク	6
2.3.1	オートエンコーダ	8
2.3.2	畳み込みニューラルネットワーク	9
2.3.3	ResNet	10
第3章	CNT フォレスト断面構造の特徴量抽出と判別	12
3.1	実験方法	12
3.1.1	データセット	12
3.1.2	モデルと訓練条件	13
3.2	結果	15
3.2.1	抽出した特徴ベクトルの主成分分析	15
3.2.2	クラスタ分布に対応する構造画像	15
3.3	考察	17
3.3.1	主成分分析	17
3.3.2	クラスタに対応する構造画像	17
第4章	光学反射率の分類による光学特性の影響を与える構造の予測	20
4.1	実験方法	20
4.1.1	データセット	20
4.1.2	モデルと訓練条件の設定	21
4.2	結果	23
4.2.1	ラベル付きデータに対する分類精度	23
4.2.2	各ラベルに分類された構造画像の枚数	23
4.2.3	構造画像の特徴分類器から得られた特徴ベクトルの分布	26
4.3	考察	27
4.3.1	ラベルをつけたサンプルの分類結果	27
4.3.2	ラベルをつけていないサンプルの分類結果	28

第5章 結論.....	29
5.1 CNT フォレスト断面構造の特徴量抽出と判別	29
5.1.1 目的との対応	29
5.1.2 今後の展望	29
5.2 光学反射率の分類による光学特性の影響を与える構造の予測	30
5.2.1 目的との対応	30
5.2.2 今後の展望	30
謝辞	31
参考文献	32
業績	33
主著	33
共著	33
付録	35
A モデル条件の選定について	35

第1章 序論

1.1 研究背景

1.1.1 カーボンナノチューブの構造と物性の関係

カーボンナノチューブ(CNT: Carbon Nanotubes)は、炭素の sp^2 結合で構成された原子一層分の厚さのグラフェンシートを円筒状に丸めた単純な構造の物質であり、高い電子移動度や電流耐性、高引張強度など多くの有用な物性を持っていることから、幅広い分野で応用が期待されている。

これらの物性には CNT がもつ構造的な特徴が寄与しており、一本の CNT の場合では、同軸円筒グラフェンシートが多層となった構造の層数は機械強度などに寄与している。単層 CNT では、構成するグラフェンシートの巻き方により、軸と垂直円周方向のイオンポテンシャル周期を反映したバンド構造をとるために金属・半導体性質を分ける。複数の CNT が集積した場合には一本一本の向きが揃っているかを示す配向性や、単位体積あるいは単位面積当たりの本数である本数密度などがあり、CNT フォレストと呼ばれる CNT が基板上に垂直高密度に合成された CNT 構造体では、可視光から赤外領域での反射率が密度に依存して吸収率が変化することや、CNT の向きが電磁波の電界の振動方向と一致したときに共振して長波長域で反射率が高くなることが報告されている [1]。

CNT の応用デバイス開発において優れた物性を引き出すためには、特性制御の手法に加えて構造と物性の相関を解析することも重要である。

CNT の構造解析には、ラマン分光法による光学的な応答を解析する手法や電子顕微鏡を使って構造を直接観察する手法がある。ラマン分光法では CNT への入射光に対する散乱光の波長シフト(ストークスシフト)をもとに結晶性を評価できる。電子顕微鏡を使った手法は電子線を用いた撮像の仕組みによって分けられる。透過型電子顕微鏡(TEM: Transmission Electron Microscopy)を用いた方法では数 nm 以下の領域を観察できるため、CNT の一本一本を観察し層数や欠陥などを評価できる。走査型電子顕微鏡(SEM: Scanning Electron Microscopy)を用いた方法では数 nm から数百 μm の領域を観察できるため、CNT の集合体を観察し密度や配向性などを評価できる。

CNT フォレストの構造解析において、配向性や密度、成長高さの解析には SEM が用いられており、SEM での CNT フォレストの断面画像を撮影したのち、各種画像解析ソフトを用いて評価される。この画像解析ソフトではフィルタリングやフーリエ変換などの画像処理を用いて CNT のエッジを検出し、上記の配向性・密度・成長高さなどの構造を定量的に評価している。

しかし、この SEM 画像の解析手法では、解析時に設定するパラメータによって解析結果が大きく変化するため、パラメータチューニングに慣れている人と初心者で解析の結果が変わったり、解析に時間がかかったりなど解析者への負担と依存が大きいこと、そして解析結果から物性との相関を直接解析することができないという課題があった。

1.1.2 機械学習の材料開発の分野への応用

機械学習は既知のデータを用いて自身で性能を改善していく学習アルゴリズムおよびそのアルゴリズムが実装されたモデルであり、応用指向の強いデータ分析手法として利用されている。身近なものでは、工学の分野に限らず広く利用されている、実験条件と得られた結果の関係を導く最小二乗法がある。

近年ではコンピュータの処理能力の大幅な向上と、インターネットの発展によりあらゆる情報へのアクセスが可能になり大量のデータを扱えるようになったことで、計算量の多いアルゴリズムの導入や汎化性能の高いモデルを作れるようになったため社会への活用が始まっている。日常生活では、人とモノの識別を高精度に行えるようになったことで自動運転に、大規模言語モデルの登場により自然言語処理能力が大幅に向上したことで文書要約や翻訳などに用いられている。学術分野では、論文の内容要約や実験条件抽出で広く使われており、特に材料開発の分野では開発の効率化と高速化を目指した、マテリアルズ・インフォマティクスにおいて活用されている。例えば、階層的クラスタリングを用いた組成が異なる結晶の XRD データの結晶性判別 [2]や、重回帰分析を用いた材料の物性からの電池特性の予測 [3]があり、CNT に関連するものではニューラルネットワークを用いた CNT 塗布膜の SEM 画像からの電気特性の予測 [4]が行われている。

1.2 目的

CNT フォレストの物性と構造には相関がありそれらを解析することは応用デバイス開発において重要であるものの、従来の構造解析手法では解析者への負担が大きいことや、物性と構造の直接的な相関解析ができないという課題があった。

これらの課題に対して、材料の分野でも活用されている機械学習を用いた解決方法の提案として次の 2 つを本研究の目的とする。

1 つ目は解析者への負担軽減のための手法の提案として、描画範囲が異なる画像サイズの構造画像に対するオートエンコーダを用いた異常検知による構造判

別の性能を明らかにすること。2つ目は物性と構造の相関解析として、従来報告がなかった CNT フォレストの光学反射率の予測と、その予測に影響を与える構造を明らかにすることである。

1.3 本論文の構成

本論文では、第 1 章では序論として CNT の構造と物性の相関と構造の解析方法の課題、および近年の材料開発分野における機械学習の活用について概説し、機械学習を用いた課題解決策について研究目的を述べた。第 2 章では本研究で扱った機械学習のアルゴリズムの基礎部分について述べた。第 3 章では CNT フォレスト断面 SEM 画像オートエンコーダを用いた異常検知による構造判別の手法と性能評価の結果について述べた。第 4 章では CNT フォレスト断面 SEM 画像を光学反射率の高低によって分類予測を行う手法とその結果について述べた。第 5 章では全体の結果を結論としてまとめた。

第2章 使用した機械学習アルゴリズム

この章では本研究で扱った機械学習アルゴリズムについて紹介する。

扱ったものを大きく分けると3種類あり、1つ目は、変数の多いデータを変数間の関係をもとに、より少ない変数で表現することで解析の効率化や可視化に用いられる次元削減、2つ目は、データ間の類似度をもとにデータをグループ分けするクラスタリング、3つ目は、データの分類や予測など汎用的に用いられるニューラルネットワークに関する手法である。

2.1 主成分分析

主成分分析(PCA)は相間のある変数同士をまとめることで元の変数の数よりも少ない変数でデータを表現する次元削減手法の一つである。

これは、図 2.1(a)のような青い点で表される2つの変数を持つデータに対して、分散が最も大きくなる赤矢印の方向を軸に持つ、図 2.1 (b)のような新たな座標系に元のデータを線形変換している。図 2.1 では2変数の場合を例に挙げたが、 D 個の変数をもつデータでも変換で、変換後の座標の各軸は、元のデータでの分散が大きい軸から順に第一主成分(PC1)、第二主成分(PC2)... 第 D 主成分(PCD)と決められていく。この時、第 j 主成分に対応する分散 σ_j^2 に対して、式(2.1)適用した値は寄与率と呼ばれ、第 j 主成分が元のデータの情報をいくら持っているかを示す指標になる。特に第一主成分から第 j 主成分までの寄与率の総和は累積寄与率と呼び、第 j 主成分までで元のデータをどれだけ再現できるかを表すため、次元削減を行う際の指標に用いられる。

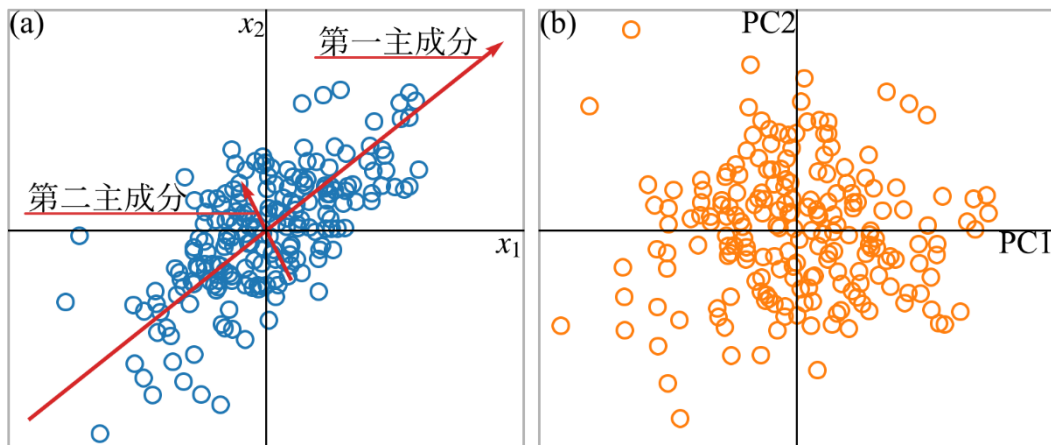


図 2.1 2変数データにおける主成分分析の例. (a) 青い点がデータの分布であり、赤い矢印がこのデータ分布における第一、第二主成分. (b) 主成分分析後の変換されたデータの分布.

$$\frac{\sigma_j^2}{\sum_{d=1}^D \sigma_d^2} \quad (2.1)$$

2.2 混合ガウス法

混合ガウス法 [5]は、データ全体の分布が複数の正規分布 p_j の重ね合わせで式(2.2)の混合ガウス分布で表せるとして、データ間の類似度をもとに図 2.2 のようなデータのグループ(クラスタ)ごとの重み π_j , 平均 μ_j , 分散共分散行列 Σ_j の3つのパラメータを求めるクラスタリング手法である.

$$p(\mathbf{x}) = \sum_{k=1}^K \pi_k p_k(\mathbf{x}) = \sum_{k=1}^K \pi_k \frac{1}{\sqrt{(2\pi)^m |\Sigma_k|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu_k)^T \Sigma_k^{-1}(\mathbf{x} - \mu_k)\right) \quad (2.2)$$

クラスタリングにはクラスタの境界が明確なハードクラスタリングと曖昧なソフトクラスタリングに分けられる. 混合ガウス法はデータが各クラスタに所属する確率でクラスタを決めるため, 後者のソフトクラスタリングに対応する.

データの分布の混合ガウス分布へのフィッティングには大きく分けて EM アルゴリズムを用いたものと変分ベイズを用いたものの2種類ある. EM アルゴリズムは正規分布自体の平均値や分散などのパラメータを調整してフィッティングさせる手法で, 変分ベイズを用いた方法ではそれらのパラメータ自体が確率分布で与えられると仮定して, それらの代表値を用いてフィッティングさせる手法である. クラスタリングではあらかじめクラスタ数を指定する必要があるが, EM アルゴリズムでは決められたクラスタ数の分だけデータを分けようとするため, 過剰なクラスタ数を指定すると一つのクラスタを複数のクラスタに振り分けようとして不自然な結果になる. 変分ベイズではクラスタ数を過剰に設定した場合でも疎なデータ分布に対して適応的にフィッティングするため, 余分なクラスタに対応する正規分布が出てこなくなる. 本研究では後者の変分ベイズによるフィッティングを行った混合ガウス法を用いた.

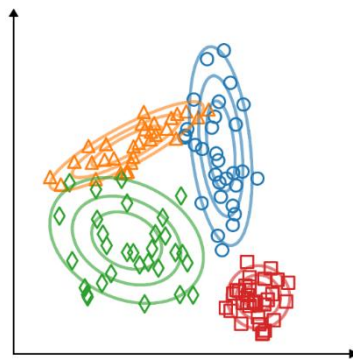


図 2.2 混合ガウス法による複数の正規分布に対応する各クラスタの分布の例

2.3 ニューラルネットワーク

ニューラルネットワークとは人間の脳神経同士の結合を模して考えられた機械学習アルゴリズムで、入力されたデータの種別を判別する分類問題や値を予測する回帰問題のほか、入力をもとに新たに文章や画像を作り出す生成モデルに用いられている。最も単純なニューラルネットワークは図 2.3 のような多層パーセプトロン(MLP)と呼ばれるモデルである。円がニューロンを模したパーセプトロンであり、このパーセプトロンが縦に並んだものを層とよぶ。特にデータが入力される左端のものを入力層、データをもとに予測値を出力する右端の層を出力層と呼ぶ。パーセプトロンは前後の層のすべてのパーセプトロンと接続されており、信号がこの接続を通る際にそれぞれに設定された荷重が掛けられてパーセプトロンに入力される。そして、荷重が掛けられた信号は総和がとられ、活性化関数による非線形変換が行われて後ろの層へ伝わっていく。出力層まで到達した信号は予測値と正解との誤差を損失関数によって計算され、より誤差が小さくなるように荷重の値が更新される。この更新には勾配降下法またはその発展手法および、勾配計算を支える誤差逆伝播法とよばれる手法が用いられる。勾配降下法は t 回目の更新の第 l 層の荷重を W_l^t 、損失の値を L 、更新の幅を調整する学習率を η とし、式(2.3)で表される式が用いられる。これは、横軸を重み、縦軸を損失としたときに、損失の極小値で勾配が 0 になることを用いた手法である。

$$W_l^{t+1} = W_l^t - \eta \frac{\partial L}{\partial W_l^t} \quad (2.3)$$

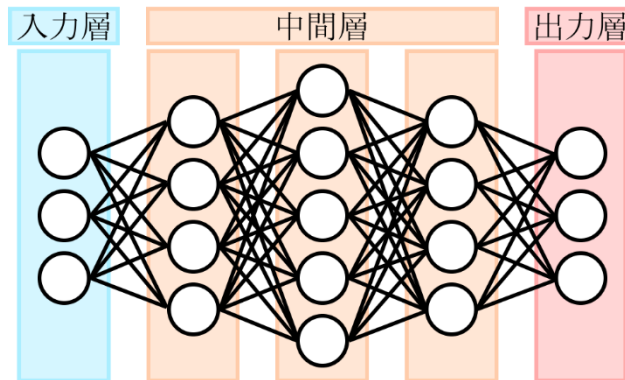


図 2.3 ニューラルネットワークの基本的なモデルの一つである多層パーセプトロンの構造

ここで層が深くなると、浅い層の荷重での微分は変数同士が複雑に絡まっているため複雑になる。そこで誤差逆伝播法では、出力層の出力を Y 、第 l 層の出力を U_l とし、式(2.4)表される微分の連鎖率を用いて効率的に微分の計算を行っている。これは入力層から出力層への順伝播の際に各層での出力値を計

算すると同時に出力値の入力値での微分を計算しておき、荷重を更新するとき、この微分を出力層から入力層に逆伝播させている。

$$\frac{\partial L}{\partial W_l^t} = \frac{\partial L}{\partial U_l^t} \frac{\partial U_l^t}{\partial W_l^t} = \frac{\partial L}{\partial U_{l+1}^t} \frac{\partial U_{l+1}^t}{\partial U_l^t} \frac{\partial U_l^t}{\partial W_l^t} = \frac{\partial L}{\partial Y} \cdots \frac{\partial U_{l+1}^t}{\partial U_l^t} \frac{\partial U_l^t}{\partial W_l^t} \quad (2.4)$$

勾配降下法はシンプルな手法であり学習速度や精度向上を目指し、データをランダムに小分けにして更新を行う確率的勾配降下法や、損失の微分を力学の速度と対応づけて慣性や抵抗力に相当する項を付け加えたモーメンタムなどの発展させた手法がある。勾配と誤差逆伝播法を用いた荷重の更新方法は MLP に限った話ではなく層ごとの微分さえ計算できれば適応できるため、多くのモデルで適用可能である。近年では収束速度と安定性から、モーメンタムと学習率を適応的に変化させる AdaDelta を組み合わせた Adam [6]とよばれる式(2.5)表される最適化手法が用いられる。ここで、 m_t と v_t の初期値はそれぞれ $m_0 = 0$, $v_0 = 0$ で、 β_1 と β_2 は訓練を行う前に設定するハイパーパラメータ、 ϵ はゼロ除算を防ぐための小さな値の定数、演算子 \odot は行列の要素ごとの積である。

$$\begin{aligned} g_t &= \frac{\partial L}{\partial W_l^t} \\ m_t &= \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t \\ v_t &= \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t \odot g_t \\ \hat{m}_t &= \frac{m_t}{1 - (\beta_1)^t} \\ \hat{v}_t &= \frac{v_t}{1 - (\beta_2)^t} \\ W_l^{t+1} &= W_l^t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \end{aligned} \quad (2.5)$$

活性化関数について、様々なものが使われるが今回使用した次の 2 つを説明する。1 つ目は中間層で用いた SELU 関数で、式(2.6)で表される図 2.4 (a) のようなグラフになる。ここで、 λ と α は定数で、それぞれ約1.05と約1.67 である。入力が正の場合は線形な変換、負の場合は指数関数的な変換が行われる特徴を持つ。

$$\text{SELU}(x) = \begin{cases} \lambda\alpha(e^x - 1), & x \leq 0 \\ \lambda x, & x > 0 \end{cases} \quad (2.6)$$

2 つ目は出力層で用いた Sigmoid 関数で、これは式(2.7)で表される図 2.4(b)のようなグラフになる。出力値が 0 から 1 の間の値をとるため、2つのクラスどちらに所属しているかを推測する二値分類において確率として使われる。

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (2.7)$$

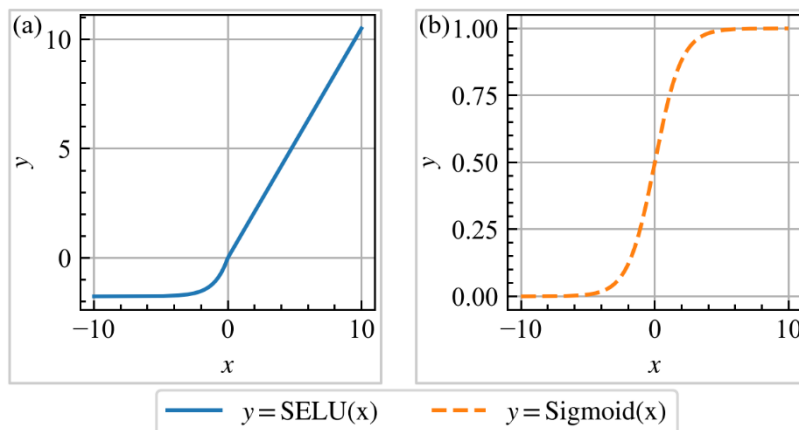


図 2.4 活性化関数の比較. (a)SELU, (b)Sigmoid

以下ではこのニューラルネットワークを応用したモデルである、オートエンコーダ、CNN、ResNet について説明する。

2.3.1 オートエンコーダ

オートエンコーダとは、データをより低次元なものに圧縮し抽象化することでデータ間の関係をまとめる手法である。データを低次元のものにするというのは PCA と同様であるが、あちらは行列を用いた線形変換であるのに対し、オートエンコーダは内部にニューラルネットワークを用いているため非線形変換を行っており、効率的なデータの圧縮を行うことができる。

オートエンコーダの構造は図 2.5 のようになっており、圧縮したい元データ X_{original} をエンコーダに入力することで圧縮された低次元のデータである特徴ベクトル z を得る。その後 z をデコーダに入力して X_{original} を再構成した X_{reconst} を得る。

このオートエンコーダを用いた有名なものには word2vec がある。これは言葉の単語をオートエンコーダを用いてベクトル化するモデルで、複数の単語をベクトル化し、それらを足し引きしたものをデコーダを通して単語に変換すると、その複数の単語の意味が足し引きされた意味を持つ単語を出力することができる。大規模言語モデルも基本的には同様で、単語ではなく文章をベクトルに変換し、このベクトルに対し演算を行うことで文章を生成している。

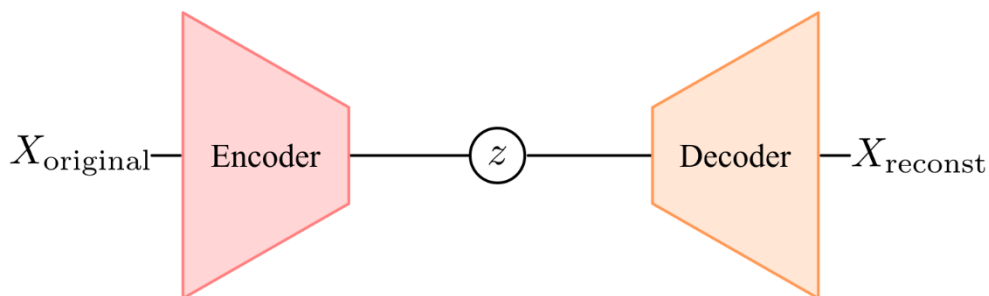


図 2.5 エンコーダで入力したデータを圧縮し，デコーダで再構成を行うオートエンコーダの構造.

2.3.2 畳み込みニューラルネットワーク

畳み込みニューラルネットワーク(CNN)はデータ処理においてエッジ強調や平滑化などで行われる畳み込み処理を利用したニューラルネットワークである。周期的なパターンがあるようなデータ，特に画像データを扱うモデルでベースとして扱われており，入力された画像に何が主に映っているかを分類する物体認識や，画像内のどこに何があるかを識別する物体検知など，幅広いタスクで用いられている。

畳み込みの仕組みは図 2.6 のようになっている。ニューラルネットワークにあった荷重は荷重フィルタ(一般にはカーネル)が対応しており，特徴マップは入力データが畳み込み層によって輪郭情報や色見の情報などに抽象化されたものである。畳み込み層では，特徴マップの左上から右下にかけて重なった領域で内積をとり，内積をとった領域と対応する場所に値を格納していき次の層に渡される特徴マップを得る。内積の部分では掛け算と足し算が行われているため，この畳み込み層でも入出力の値で微分を計算することができる。

学習に使われる画像データは，多くの場合は横方向の幅と縦方向の高さに加えて，RGBの色味(チャンネル方向)を持つ3次元配列として扱われる。特徴マップも同様であり，荷重フィルタでもチャンネル方向のフィルタの数を変えることで，特徴マップのRGBの3種類の値を分解・統合してチャンネル方向の成分を増減できる。

また，CNNでは畳み込み層以外にも，特徴マップ中の代表値を取り出すプーリング層や特徴マップの値を調整し性能向上を図る正規化層といった層での処理が行われる。

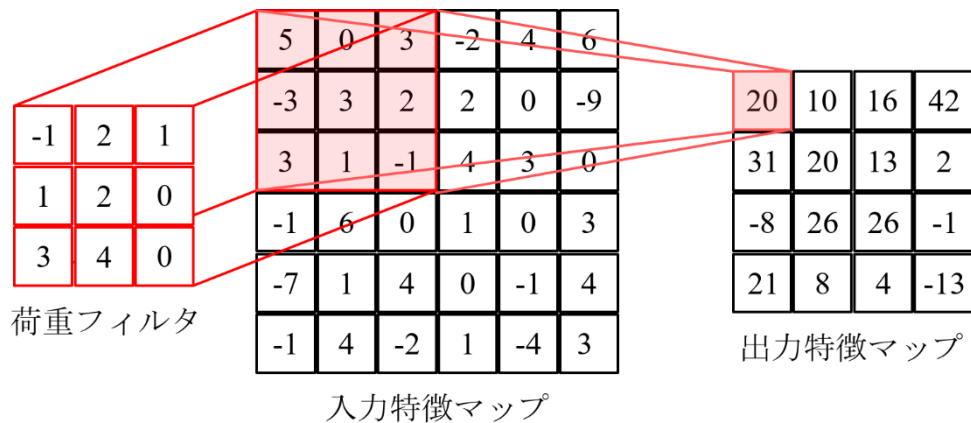


図 2.6 CNN の仕組み. 荷重フィルタで特徴マップを走査して新たな特徴マップを作る.

2.3.3 ResNet

ResNet [7] は CNN を発展させたモデルの一つである. 従来の CNN を基にしたモデルは, 畳み込みやプーリングによる入力画像の抽象化などの処理を順番に実行していくのに対し, ResNet は図 2.7 のような, 層への入力とその出力に加算される残差ブロックという構造が使われていることが特徴である. 畳み込み層や正規化層による変換を $f(x)$, 残差ブロック全体による入力の変換を $h(x)$ として数式で表すと式 (2.8) となる.

$$h(x) = x + f(x) \quad (2.8)$$

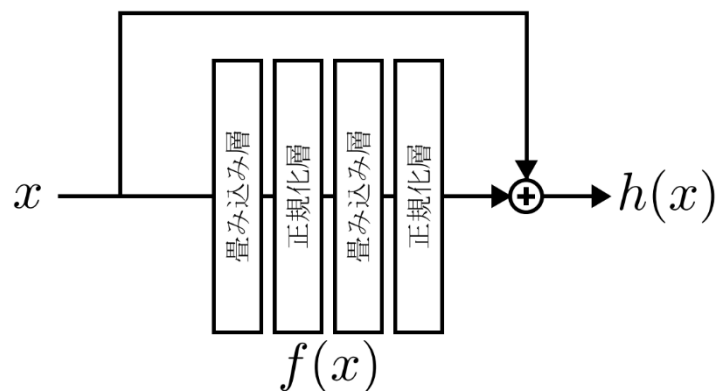


図 2.7 ResNet における残差ブロックの構造

ニューラルネットワークでは層を深くすればするほど性能が向上すると考えられているが, CNN や MLP などの単純なモデルの場合, 各層の微分が小さくなった場合に誤差逆伝播での勾配が小さくなるため, 荷重の更新幅が小さくな

りモデルの学習が止まってしまう勾配消失問題と、層を深くしても性能の頭打ちが起こり、さらに深くした場合に性能が低下してしまうモデル劣化という問題が生じる。

これらの問題に対して残差ブロックは次のように機能する。勾配消失問題では、畳み込み層や正規化層による変換 $f(x)$ の勾配が十分小さい場合、この層の入力 x における損失 L の勾配は式(2.9) のようになる。つまり、残差ブロック全体での勾配は 1 で近似でき、深い層からの勾配がそのまま浅い層に伝わるため勾配消失が起りにくくなる。

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial h(x)} \frac{\partial h(x)}{\partial x} = \frac{\partial L}{\partial h(x)} \frac{\partial (f(x) + x)}{\partial x} \approx \frac{\partial L}{\partial h(x)} \quad (2.9)$$

モデル劣化では、変換 $f(x) \approx 0$ となるように荷重を更新することで、残差ブロックは $h(x) = x + f(x) \approx x$ となり恒等関数の役割を果たす。つまり、入力をそのまま出力することで不要な層を無視することができ、性能低下を防ぐことができる。ResNet は層をさらに増やすことを可能にし、様々なケースでモデルの性能向上に貢献している。

残差ブロックにはその内部構造によって種類があり、元論文 [7] では BasickBlock と Bottleneck の 2 つが提案されている。BasickBlock では畳み込み層と正規化層が交互に 2 回繰り返されたシンプルな構造なのに対し、Bottleneck では入力データのチャンネル数を一度圧縮しその後入力時のチャンネル数に戻す構造をしている。これは計算時のコストを同程度にしつつ、層数を増やすことで性能向上を図ったものになる。

第3章 CNT フォレスト断面構造の特徴量抽出と判別

この章では解析者の負担軽減を目指して、描画範囲の異なる 3 種類の画像サイズの構造画像に対する、オートエンコーダとクラスタリングを用いた構造特徴の抽出と異常検知による構造判別を行い、その性能評価として判別された画像の構造について評価した。

3.1 実験方法

3.1.1 データセット

CNT フォレストサンプルの断面 SEM 画像 1 枚に対して、図 3.1 のように画像の縦方向と横方向それぞれに 16 px ずつずらしながら 64 × 64 px (約 600 × 600 nm) および 32 × 32 px (約 300 × 300 nm), 16 × 16 px (約 150 × 150 nm) の画像サイズを切り出すことで、64 × 64 px の構造画像を 4398 枚、32 × 32 px の構造画像を 4661 枚、16 × 16 px の構造画像を 18921 枚取得した。これを異なる条件で合成した 9 つのサンプルに対して行いデータセットを作成した。作成したデータセットの内 8 割を訓練データに、2 割を荷重更新に使用しない検証データとして使用した。

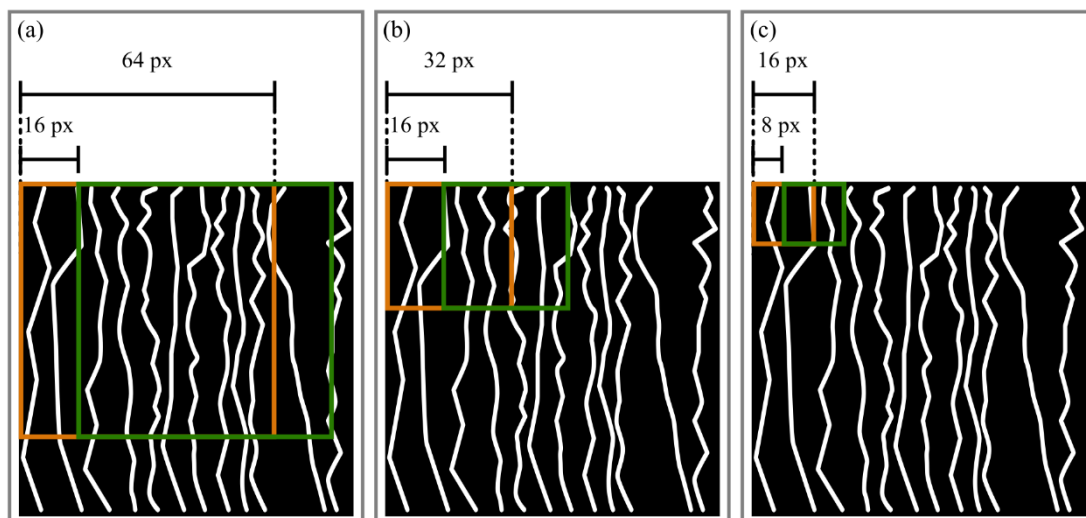


図 3.1 CNT フォレスト断面 SEM 画像からのデータセットの作成方法. 16 px および 8px ずつずらして、(a) 64 × 64px , (b) 32 × 32px, (c) 16 × 16px の画像を切り出した。

3.1.2 モデルと訓練条件

それぞれの画像サイズごとに図 3.2 のようにモデルを作成した [8]. すべてのモデルに共通して、エンコーダでは畳み込み層, バッチ正規化層, SELU 関数を繰り返しており, デコーダでは畳み込み層, バッチ正規化層, SELU 関数, デコーダに入力して転置畳み込み層, バッチ正規化層, SELU を繰り返して出力層では SELU 関数代わりに Sigmoid 関数を使用した. 層の数が異なるのは, 畳み込み層を通るたびに画像が半分になるため, 層が多いモデルで小さなサイズの画像を入力する際に生じるエラーを避けるためである. 出力層の活性化関数に Sigmoid 関数を使用しているのは, 入力画像の値を 0 から 1 の間に正規化して

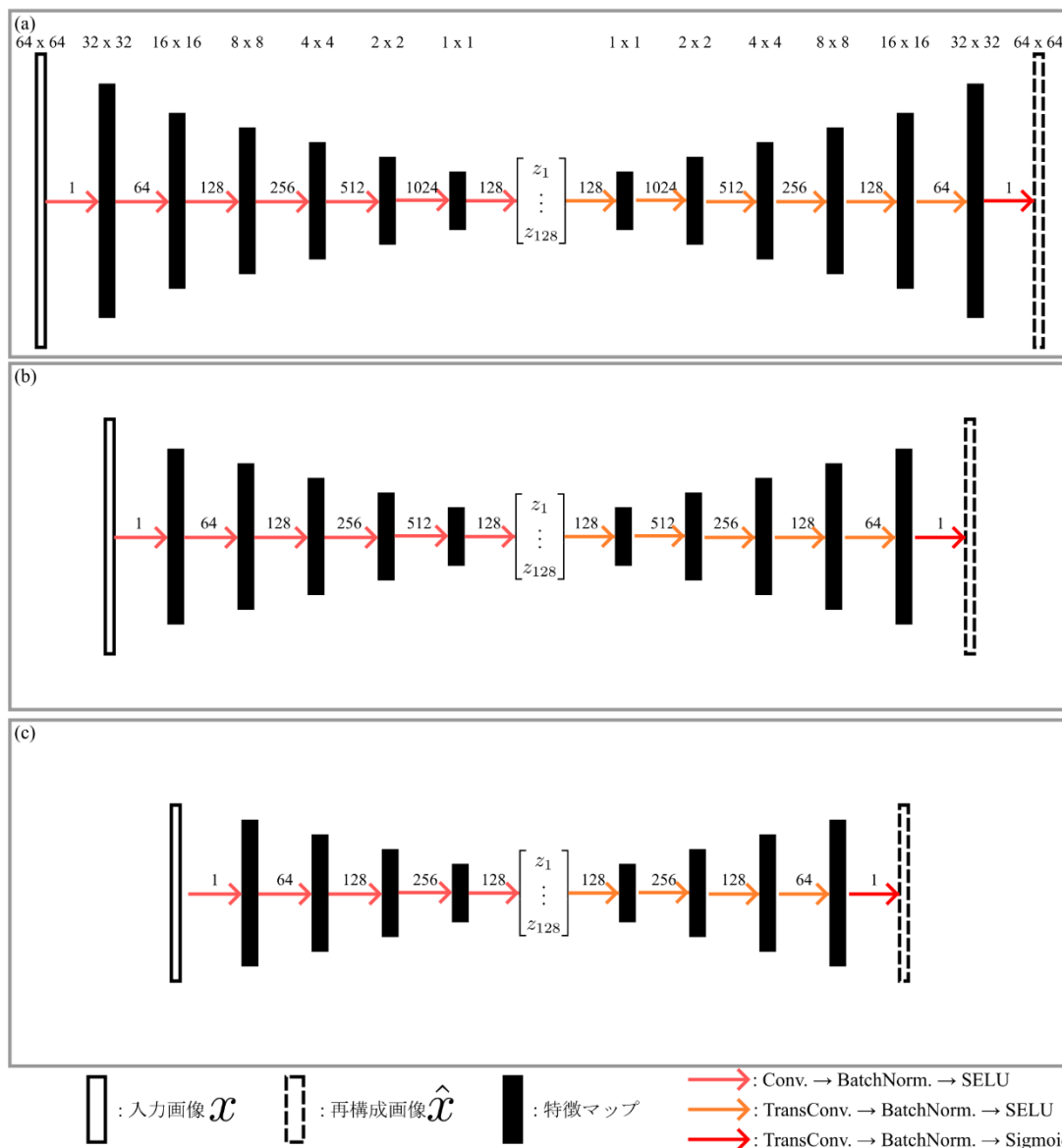


図 3.2 各画像サイズにおける特徴量抽出に用いたオートエンコーダのモデル. (a) 64 × 64px 用のモデル, (b) 32 × 32px 用のモデル, (c) 16 × 16px 用のモデル

おり、同じ範囲の値を出力する必要があるためである。転置畳み込み層は通常の畳み込みの処理と逆向きの処理を行うことで特徴マップのサイズを大きくする層である。矢印の上を書いてある数値は、矢印の根本の画像および特徴マップにおけるチャンネル数を意味しており、ベースを 64 として層を通るたびに倍々に増加させた。そのほかのモデルのハイパーパラメータを表 3.1 に示す。

損失関数には入力画像 x と再構成画像 \hat{x} の M 個の同じ位置のピクセルに対するバイナリークロスエントロピー(BCE)の総和である式(3.1)を用いた。

$$\text{BCE}(x, \hat{x}) = - \sum_{m=1}^M \{ \hat{x}_m \log x_m + (1 - \hat{x}_m) \log(1 - x_m) \} \quad (3.1)$$

訓練時のハイパーパラメータを表 3.2 に示す。早期終了はモデルの過学習による性能低下を防ぐもので、早期終了には荷重の更新に使われない検証データでの損失が増加し汎化性能が低下する際に何回目の荷重更新で訓練を打ち切る役割がある。エポック数は一つのデータセットを何週繰り返すかを表す単位である。

表 3.1 オートエンコーダのハイパーパラメータ

設定項目名	設定内容
畳み込み層フィルタサイズ	4×4
特徴ベクトルの次元	128
ベースのチャンネル数	64
層間の活性化関数	SELU
出力層の活性化関数	Sigmoid
正規化層	バッチ正規化層

表 3.2 訓練時のハイパーパラメータ

設定項目名	設定内容
学習率	0.001
バッチサイズ	128
最大エポック	1000 エポック
早期終了	50 エポック
荷重の最適化手法	Adam ($\beta_1 = 0.9, \beta_2 = 0.999$)

以上のようにしてオートエンコーダを訓練したのち、エンコーダ部分を用いて構造画像の特徴ベクトルを抽出した。抽出した特徴ベクトルはそのままでは次元が大きいため、主成分分析を用いて累積寄与率が 0.5 になる次元までを残し

て次元削減を行った。次元削減後の特徴ベクトルに対して、混合ガウス法を用いて最大クラスタ数を 10 個に設定しクラスタリングを行った。

3.2 結果

3.2.1 抽出した特徴ベクトルの主成分分析

各画像サイズにおける構造画像から抽出された特徴ベクトルに対して主成分分析を行った際の累積寄与率を図 3.3 に示す。横軸の値の番号までの主成分を第一主成分から順番に選んだ時の、その番号までの累積寄与率を縦軸に表している。画像サイズが大きいと累積寄与率の推移が緩やかであるのに対し、画像サイズが小さくなるにつれて立ち上がりが急になっており、それぞれの画像サイズで累積寄与率が 0.5 に到達するのに必要な主成分の個数は 64×64 px で 42 個、 32×32 px で 30 個、 16×16 px で 25 個となった。

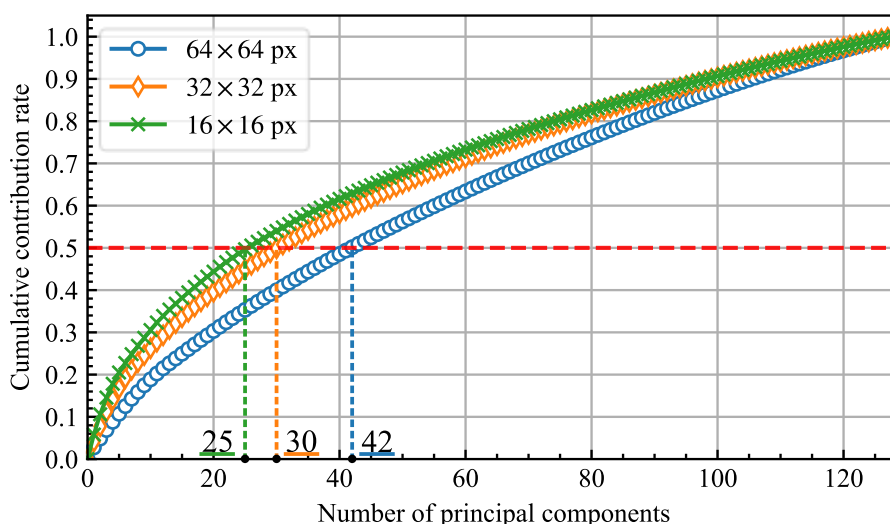


図 3.3 各画像サイズで取り出された特徴ベクトルの主成分分析を行った際の累積寄与率の推移と、累積寄与率が 0.5 に到達する主成分の数。

3.2.2 クラスタ分布に対応する構造画像

主成分分析を行った後の、構造画像から抽出された特徴ベクトルの分布を図 3.4 に示す。クラスタリングに用いた主成分の数は 64×64 px で 42 個、 32×32 px で 30 個、 16×16 px で 25 個である。

クラスタ番号の順番は、式(2.2)の混合ガウス分布の各クラスタに対応する正規分布同士の重み π_j に基づき昇順に並べている。この順番はおおよそクラスタに属するデータの数とそろっており、クラスタ番号とそのクラスタに属するデ

ータの割合を図 3.5 に示す。

それぞれのクラスタの分布について、すべての画像サイズに共通して属するデータ数が多いクラスタほど原点付近に、データ数が少ないクラスタほど原点から離れた特徴的な分布がみられた。図 3.4 (a) の 64×64 px では 5 番以前のクラスタがデータ数が少なくまばらな分布になっている中、4 番のクラスタは第一主成分軸状のほかクラスタの外周部に分布していた。図 3.4 (b) の 32×32 px では、0 番と 1 番のクラスタが特徴的な分布になっており、0 番は第一主成分軸状の第二主成分負の部分に、1 番は第一主成分負の部分に広く分布していた。図 3.4 (c) の 16×16 px では、3 番クラスタが第一主成分負の部分に偏って分布しているものの、ほとんどの分布が原点付近に集中していた。

クラスタに属するデータの割合について、画像サイズが小さい場合の方がクラスタ間での割合の差が小さくなっていった。また、割合がほぼ 0 のクラスタ番

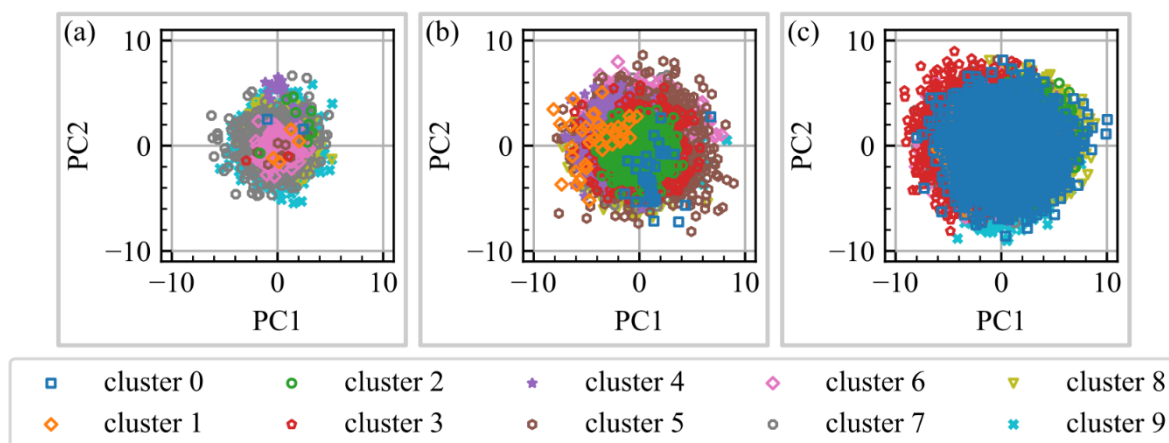


図 3.4 各画像サイズにおける、構造画像の特徴ベクトルの第一および第二主成分分布。
(a) 64×64 px. (b) 32×32 px. (c) 16×16 px.

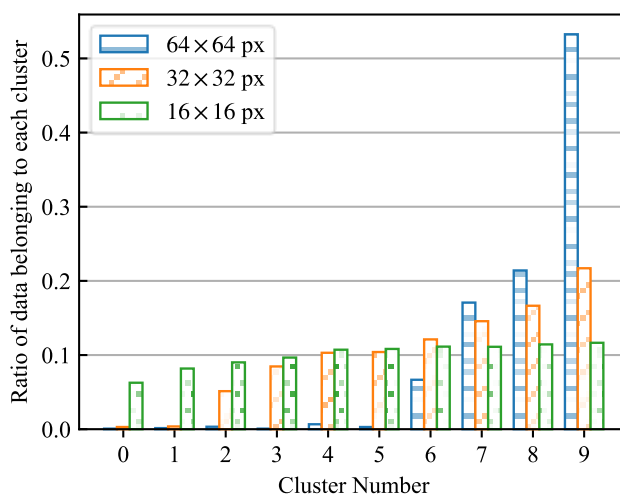


図 3.5 各画像サイズにおける、全データに対するクラスタに属するデータの割合。

号は $64 \times 64 \text{ px}$ では 5 番まで、 $32 \times 32 \text{ px}$ では 1 番までなのに対して、 $16 \times 16 \text{ px}$ はすべてのクラスタにおいて割合がほぼ一定になっていた。

各画像サイズにおける、クラスタリングされた特徴ベクトルに対応する構造画像 16 枚を図 3.6 に示す。画像が 16 枚に満たないクラスタについては足りないところを黒で塗りつぶしている。

すべての画像サイズにおいて各クラスタ内に対応する構造画像は類似した画像が含まれていた。 $64 \times 64 \text{ px}$ の画像サイズにおいて、4 番のクラスタでは左側に束状 CNT が、6 番のものでは高密度の構造に対応していた。 $32 \times 32 \text{ px}$ のサイズでは、0 番のクラスタで高密度な構造に対応していた。 $16 \times 16 \text{ px}$ のサイズでは、他のサイズで見られる高密度な構造のクラスタは見られなかった。

3.3 考察

3.3.1 主成分分析

図 3.3 の主成分の数ごとの累積寄与率において画像サイズが小さいほど早く推移が急であったことについて、画像サイズが大きいと構造の情報が多く含まれており、逆に画像サイズが小さいと構造の情報が少ないことによって、エンコーダによる画像情報の特徴ベクトルへの圧縮率に差が生まれたためだと考える。

3.3.2 クラスタに対応する構造画像

各クラスタに属するデータの割合について考察する。

$64 \times 64 \text{ px}$ や $32 \times 32 \text{ px}$ の場合にクラスタ間で偏りが生じたのは、画像サイズが大きいと、画像一枚の中に複数の構造が含まれており、その含まれている構造の組み合わせごとにクラスタに分けられたためだと考える。それに対して、 $16 \times 16 \text{ px}$ のように画像サイズが小さいと構造が個々に分割され、構造ごとにクラスタに振り分けられたため、クラスタ内のデータ数の割合が一様になったと考える。

クラスタの分布と構造画像について考察する。すべての画像サイズの各クラスタにおいて類似した構造の画像に対応していたため、画像サイズによらず CNT フォレスト断面 SEM 画像の構造特徴を特徴ベクトルとして圧縮することができると思う。

それぞれの画像サイズについて考察する。 $64 \times 64 \text{ px}$ では、対応する構造画像が少なかった 0 番から 5 番のクラスタについて、4 番のクラスタに対応する構造画像は画像の左右での本数密度の差や中央部分に束状 CNT がある特徴的

構造を持っており、この構造のため他のクラスタの外周部分布し、異常として検知されたのではないかと考える。1番と2番は他の構造と比べて配向性の良い構造をしており、5番は配向性が悪く霧がかかった構造になっているため分けられたと考える。0番と3番は、7番や8番のクラスタに対応する構造と似ているもののCNTの太さが異なるため、それぞれで別のクラスタとして分けられたと考える。32×32 px では、第一主成分軸状に分布していた0番のクラスタは他のものと比べて高密度になっている構造のため、第一主成分の負の部分に分布していた1番のクラスタは画像の内の左側の束状構造と右側の低密度構造のため、特徴的な分布の別のクラスタに分けられたと考える。16×16 px では、大きな画像で見られた密度異常の構造に対応するクラスタはなく、どの画像も画像内にCNTが数本あるような構造になっており、これは密度異常の構造を持っていた画像が分割されたためだと考える。また、1, 2, 6番のクラスタは低密度であることは共通しているものの、CNTの場所によって別のクラスタに分けられており、同様に他のクラスタでもCNTを含んでいるものの、その場所や向きによってクラスタが分けられていた。これは画像サイズを過度に小さくしたことで、CNTの画像内にある場所やCNTが含まれているかどうかなど、構造のバリエーションが増えたためであると考えられる。

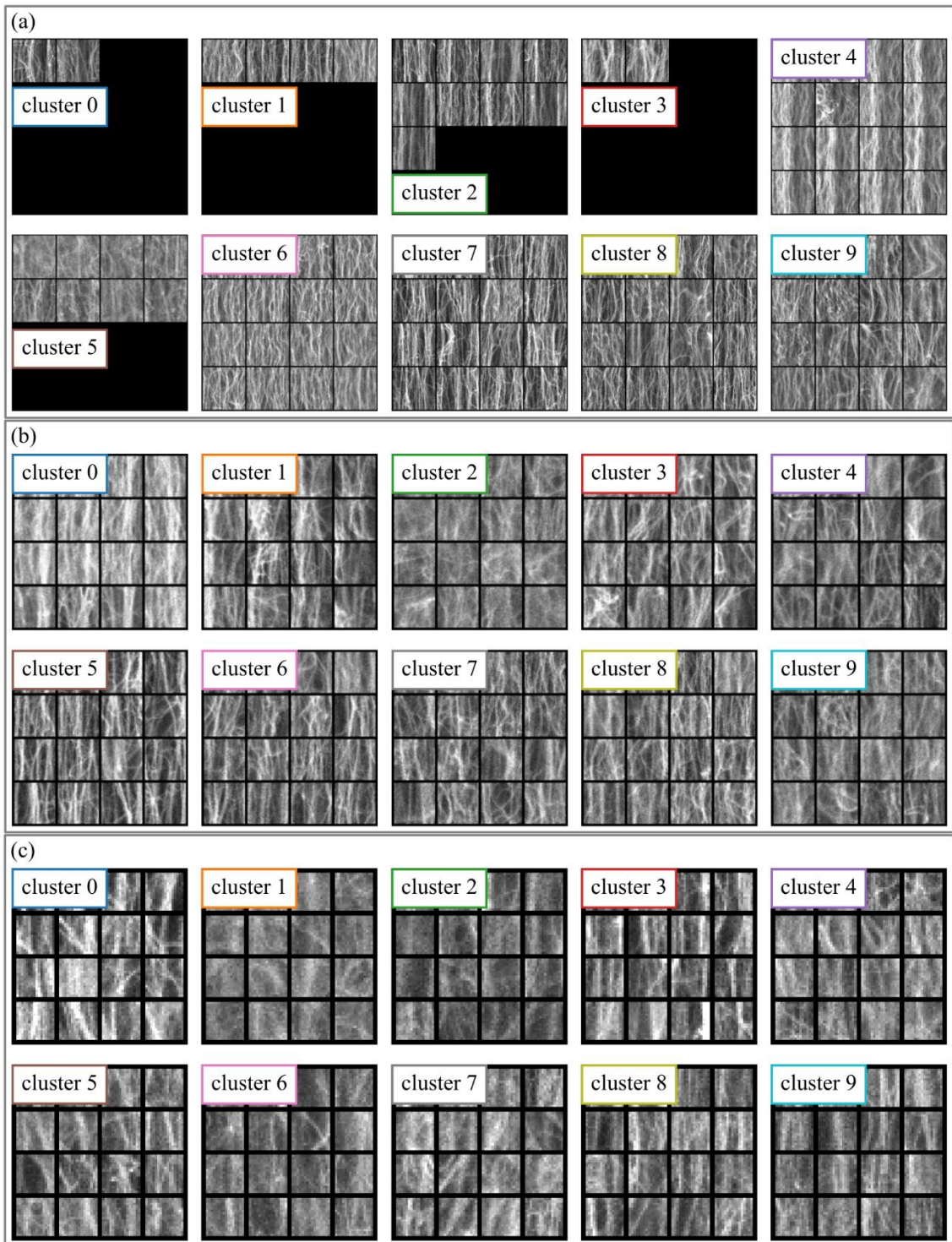


図 3.6 各画像サイズにおいて、クラスタリングされた特徴ベクトルに対応する構造画像(a) 64×64 px. (b) 32×32 px. (c) 16×16 px [8].

第4章 光学反射率の分類による光学特性の影響を与える構造の予測

この章では従来報告の無いCNTフォレストの光学反射率の予測を行うことを目的として、機械学習モデルを使用したCNTフォレストの断面構造からの反射率の高低分類を行い、その分類結果に影響を与える構造について評価した。

4.1 実験方法

4.1.1 データセット

データセットは、入力データとして第3章で用いたCNTフォレスト断面SEM画像の構造画像を、教師データとして図4.1の各サンプルの全反射率スペクトルを基に「High」と「Low」のラベルをつけたものを使用した。

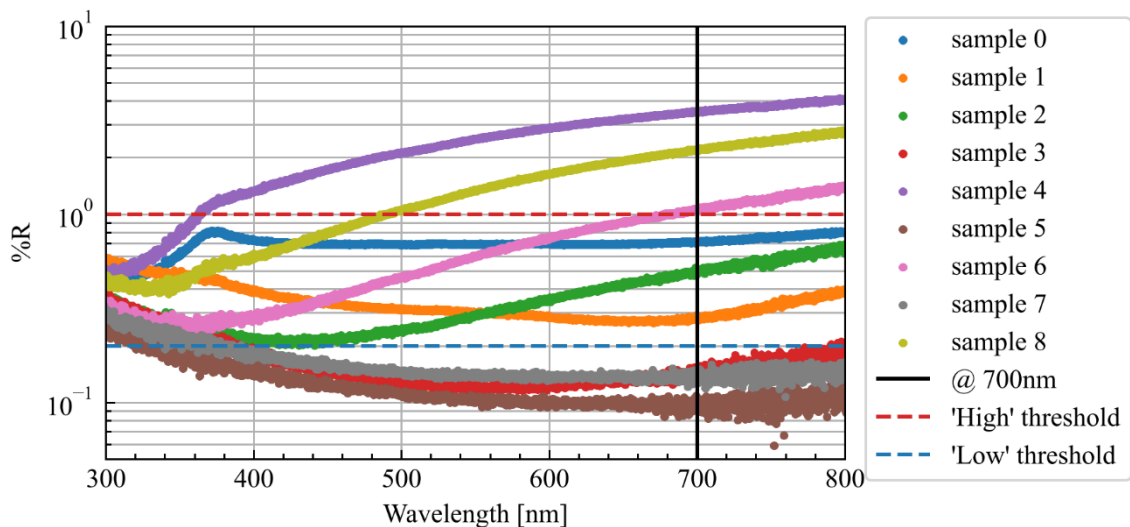


図 4.10 番から 8 番のサンプルの紫外可視分光光度計による全反射率スペクトル [1] [9].

入力する画像データは、切り取り位置がそれぞれの画像サイズの倍数になるような画像をテストデータとして除外し、それ以外の画像のうち、8割を訓練データ、残りの2割を検証データとして利用した。教師データのラベル付けの条件は、波長 700 nm における反射率が 1% 以上のサンプルを「High」、反射率が 0.2% 以下のサンプルを「Low」、0.2% より大きく 1% 未満のサンプルにはラベルを付けなかった。その結果、各サンプル番号と対応するラベルは以下の表 4.1 となった。

表 4.1 教師データとして扱うラベルとそれに対応する CNT フォレストのサンプル番号 [9]

ラベル	ラベル付けの条件	サンプル番号
High	$\%R \geq 1$	4, 6, 8
Low	$\%R \leq 0.2$	3, 5, 7
無し	$0.2 < \%R < 1$	0, 1, 2

4.1.2 モデルと訓練条件の設定

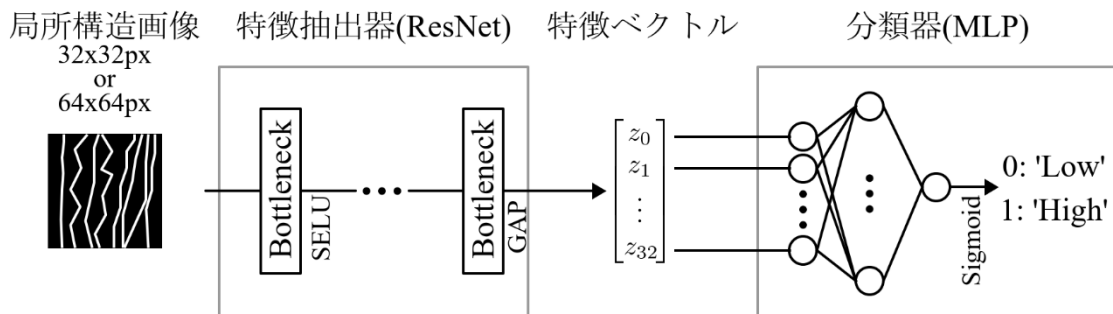


図 4.2 反射率分類に用いたモデルの構造. 前段が特徴抽出器. 後段が分類器 [9]

分類に使用したモデルは図 4.2 のような二段に分かれるものを使用した. 1 段目は残差ブロックに Bottleneck を使用した畳み込み層が 50 層ある ResNet50 をベースにした, 画像の特徴ベクトルを得るための特徴抽出器である. 内部の構成には, 活性化関数に SELU, 正規化層にバッチ正規化層を使用し, 最後 Bottleneck から出力された特徴マップに対して, 各チャンネルの特徴マップ全体の平均を求めるグローバルアベレージプーリング (GAP) を使用した. 2 段目は中間層が一層の MLP を用いた分類器である. 内部の構成には, 層間の活性化関数には SELU, 出力層には Sigmoid を使用した. 詳細なモデルの条件については, 特徴抽出器を表 4.2 に, 分類器を表 4.3 に示す.

損失関数には, 0 から 1 の間の値をとるモデルの予測値を y_{pred} , 0 または 1 のどちらかの値をとる正解のラベルを y_{true} として, 式 (4.1) で表されるバイナリクロスエントロピーを使用した.

$$\text{BCE}(y_{\text{pred}}, y_{\text{true}}) = -\{y_{\text{true}} \log y_{\text{pred}} + (1 - y_{\text{true}}) \log(1 - y_{\text{pred}})\} \quad (4.1)$$

訓練条件については両方の画像サイズの場合で, 付録 A での調整をもとに表 4.4 の条件で訓練を行った. 荷重減衰はモデルの過学習による性能低下を防ぐもので, モデルの荷重が過度に大きくなるのを防ぐ効果がある.

以上の条件でモデルの訓練を行った後、テストデータによる分類精度の評価と、特徴抽出器から得られる特徴ベクトルの主成分の分布を確認した。

表 4.2 1 段目の特徴抽出器のハイパーパラメータ

設定項目名	設定内容
残差ブロック	Bottleneck
出力される特徴マップのチャンネル数	32
残差ブロック内の活性化関数	SELU
出力層における活性化関数	恒等関数 ($f(x) = x$)
正規化層	バッチ正規化

表 4.3 2 段目の分類器のハイパーパラメータ

設定項目	設定内容
入力層のパーセプトロン数	32
中間層のパーセプトロン数	128
出力層のパーセプトロン数	1
ドロップアウト率	0.3
層間の活性化関数	SELU
出力層の活性化関数	Sigmoid

表 4.4 訓練時のハイパーパラメータ

設定項目	設定内容
学習率	0.0001
バッチサイズ	64
最大エポック	10000 エポック
早期終了	50 エポック
荷重減衰	0.001
荷重の最適化手法	Adam ($\beta_1 = 0.9, \beta_2 = 0.999$)

4.2 結果

4.2.1 ラベル付きデータに対する分類精度

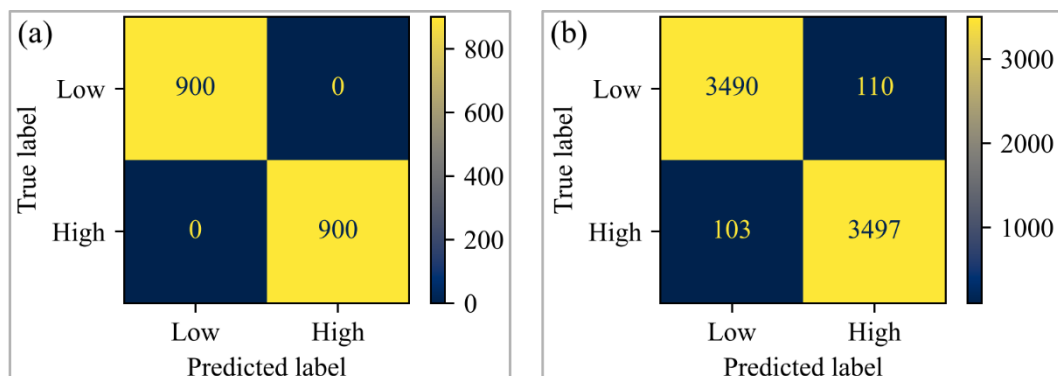


図 4.3 それぞれの画像サイズにおける、モデルの分類精度を意味する混同行列. (a) 64 × 64 px (b) 32 × 32 px [9].

訓練完了後のモデルに対して 64 × 64 px と 32 × 32 px それぞれの画像サイズにおけるテストデータを分類させたときの混同行列を図 4.3 に示す。混同行列はモデルの分類精度を可視化したもので、教師データの真のラベルに対してモデルによって予測されたラベルがどれだけ対応しているかを行列形式で表したものである。64 × 64px では「High」と「Low」の両方のラベルで正確に分類できていた。それに対して 32 × 32px では、両方のラベルで 100 枚強の画像で分類に誤りがあり、全体の分類精度は 97.0%、となった。

4.2.2 各ラベルに分類された構造画像の枚数

各サンプルの画像をモデルに分類させたときに各ラベルに分類された構造画像について、それぞれの枚数を表 4.5 に示す。

ラベルを付けた画像について、64 × 64 px のサイズの画像は混同行列に示されたものと同じで異なるラベルに分類されることはなかった。32 × 32 px のサイズの画像はそれぞれのサンプルで一部の画像が異なるラベルに分類されるた。「Low」ラベルである 5 番のサンプルで最も異なるラベルに分類された数が少なくなっていた。

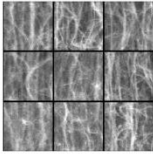
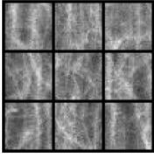
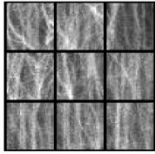
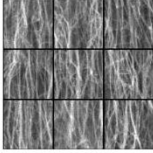
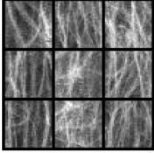
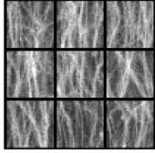
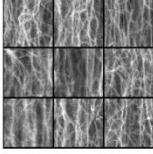
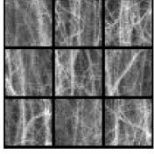
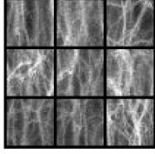
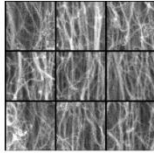
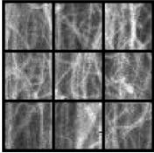
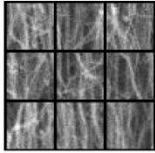
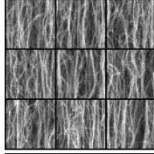
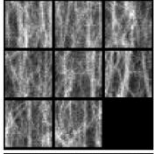
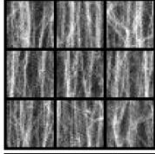
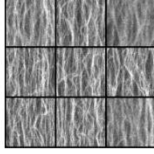
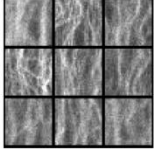
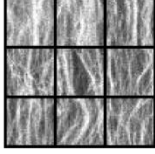
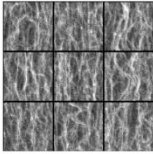
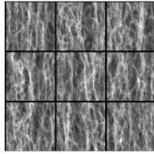
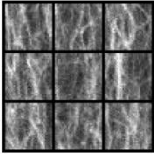
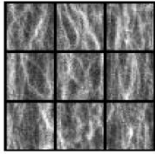
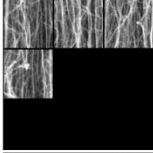
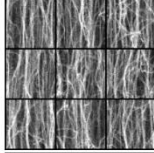
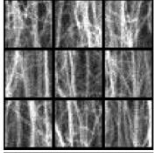
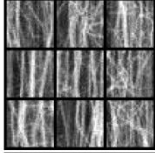
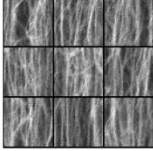
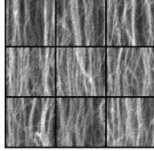
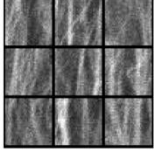
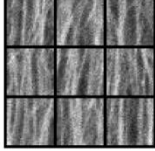
ラベルを付けなかったサンプルについて、両方の画像サイズにおいて、0 番のサンプルは「High」に、1 番のサンプルは「Low」に多く分類されていた。しかし、2 番のサンプルは 64 × 64 px の場合には「Low」に、32 × 32 px の場合

には「High」に多く分類されており，逆に分類される傾向が見られた．また，各サンプルの構造画像をモデルに入力した際に「High」と「Low」に分類された画像を表 4.6 に示す．64 × 64 px のサイズにおいては誤分類がなかったため，「High」を「Low」と，「Low」を「High」と分類した画像の部分は空白にしている．画像サイズに限らず，それぞれに分類された画像の特徴について，「High」と分類されたものは，白線で示された CNT が画像に対して斜めに伸びていることや左右に振れて波打っていることから配向性が悪いことや，靄がかかったような CNT の形状が曖昧になっているような特徴を持っていた．「Low」と分類されたものは，全体的に画像に対して上下に真っすぐ伸びた配向性の良い CNT が写っており，「High」のものとは比べて密度が高くなっているような特徴を持っていた．

表 4.5 各サンプル画像をモデルに分類させたときの各ラベルに分類された画像の枚数 [9]

ラベル	サンプル番号	64 × 64px		32 × 32px	
		High	Low	High	Low
High	4	300	0	1182	18
	6	300	0	1151	49
	8	300	0	1164	36
Low	3	0	300	39	1161
	5	0	300	8	1192
	7	0	300	63	1137
無し	0	224	76	1085	115
	1	4	296	136	1064
	2	81	219	732	468

表 4.6 各サンプルの画像をモデルに分類させたときの各ラベルに分類させた画像 [9]

ラベル	サンプル番号	64 x 64 px		32 x 32 px	
		High	Low	High	Low
High	4				
	6				
	7				
Low	3				
	5				
	7				
無し	0				
	1				
	2				

4.2.3 構造画像の特徴分類器から得られた特徴ベクトルの分布

訓練完了後の特徴抽出器で得られた特徴ベクトルの、横軸を第一主成分、縦軸を第二主成分としたときの分布は、 $64 \times 64\text{px}$ と $32 \times 32\text{px}$ の画像サイズについてそれぞれ図 4.4 の(a) と (b) のようになった。両方の画像サイズに共通して第一主成分方向が支配的になっていた。ラベルを付けたサンプルの画像の分布は、 $64 \times 64\text{px}$ では、高反射率のサンプルが第一主成分負の領域に、低反射率のサンプルが第一主成分正の領域に集中しているのに対し、 $32 \times 32\text{px}$ では、反射率ごとの分布が正負反対になっていた。ラベルを付けなかったサンプルの画像の分布は、0 番と 2 番サンプルは 0 付近に、1 番サンプルは第一主成分負の領域に分布していた。

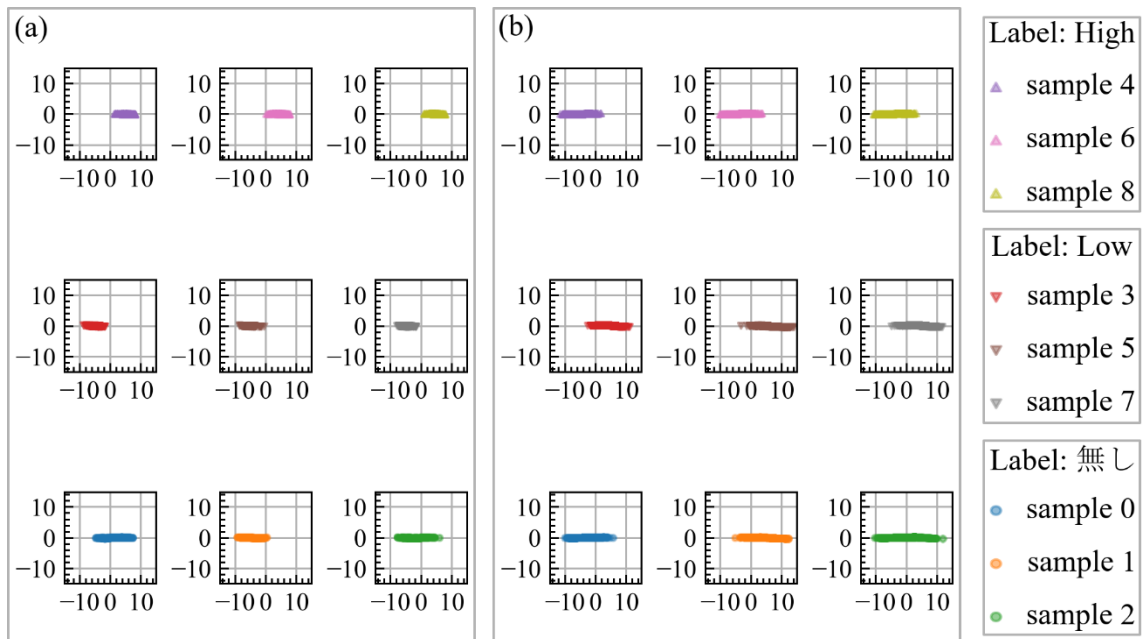


図 4.4 横軸を第一主成分、縦軸を第二主成分としたときの特徴ベクトルの分布. (a) $64 \times 64\text{px}$. (b) $32 \times 32\text{px}$ [9].

図 4.4 では第一主成分が支配的であったため、第二主成分の方向の分布を確認するために、各主成分の値を -1 から 1 の間でリスケールした分布を図 4.5 に示す。ラベルを付けたサンプルの画像の分布は、ラベルごとに第一主成分で左右に分かれているのは両方の画像サイズで同様だが、 $64 \times 64\text{px}$ では原点を通る負の傾きを持つ直線で分離できるような分布になっているのに対し、 $32 \times 32\text{px}$ ではいずれのサンプル画像の分布も曲線的になっており、第一主成分の 0

付近の第二主成分正の部分でラベルごとの分布が重なっていた。ラベルを着けなかったサンプルの画像の分布は、0 番のサンプルでは両方の画像サイズで第一主成分 0 付近に分布しつつも $64 \times 64 \text{ px}$ においては正の領域に、 $32 \times 32 \text{ px}$ においては負の領域に分布していた。1 番のサンプルでは他のサンプルと比較して正と負の領域の両方に偏って分布していた。2 番のサンプルでは $64 \times 64 \text{ px}$ においてはやや負の領域に偏って分布しているが、 $32 \times 32 \text{ px}$ においては「High」と「Low」の両方のラベルにまたがった分布をしていた。

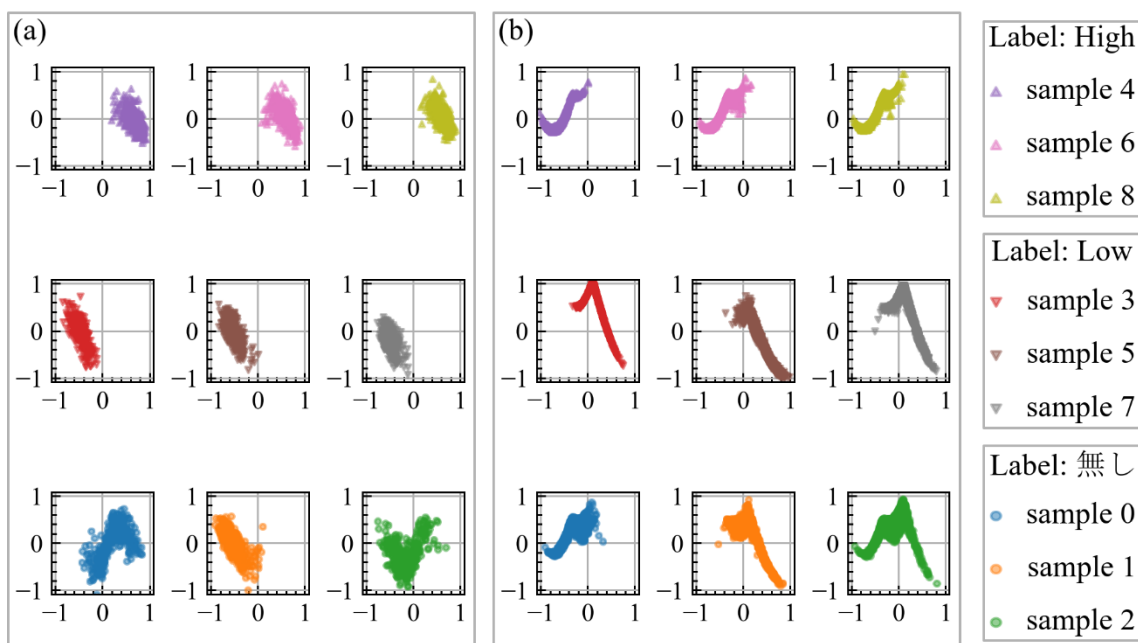


図 4.5 -1 から 1 の間でリスケールしたとしたときの特徴ベクトルの主成分の分布。(a) $64 \times 64 \text{ px}$. (b) $32 \times 32 \text{ px}$ [9].

4.3 考察

4.3.1 ラベルをつけたサンプルの分類結果

ラベルをつけたサンプルの分類結果について考察する。 $64 \times 64 \text{ px}$ の場合に誤分類が無かったことについて、CNT フォレストの光学特性は CNT の集合体としての構造の特徴が由来であることから、画像サイズが大きいデータの方が CNT の集合体としての情報を持っているため反射率の高低に関する分類で高い精度になったのではないかと考える。 $32 \times 32 \text{ px}$ の場合はいくつか誤りがあったことについて、画像サイズが小さくなったことで構造にグラデーションが

生じ、図 4.5 の特徴ベクトルの分布のように各ラベルの境界が曖昧になったため、境界付近に分布した画像が別のラベルに分類されたと考える。また、表 4.6 の分類された構造を見ると、配向性の悪い構造を「High」、良い構造を「Low」と分類していたことから、CNT フォレストの構造と光学特性の関係に近い情報で分類できたと考える。

図 4.4 の特徴ベクトルの分布において画像サイズによらず「High」と「Low」のラベルの分布が第一主成分に支配的に分かれたことについて、ラベルの数が2つと少なかったため一つの軸で十分分離することができたと考える。また、画像サイズによって分布の場所が逆になったことについて、モデルの初期値あるいは主成分分析で回転行列をかける際の回転方向が原因だと考える。

4.3.2 ラベルをつけていないサンプルの分類結果

ラベルをつけていないサンプルの分類結果について考察する。両方の画像サイズに共通して 0 番のサンプルは「High」に、1 番は「Low」に分類される傾向があり、これはそれぞれのサンプルの反射率において 0 番のものが 1 番のものより高反射率であったことと対応がとれていると考える。

これに対して 2 番のサンプルは 64×64 px では「Low」に多く分類されていたのに対し、 32×32 px では「High」に多く分類されており、さらに各ラベルに分類された画像枚数の比は $732 : 468$ と他のサンプルの比率よりも拮抗していた。表 4.6 のそれぞれの画像サイズで各ラベルに分類された画像を見ると、 64×64 px の画像サイズでは両方のラベルにおいて配向性が高く疎と密両方の構造を 1 枚の画像に含んでいるのに対し、 32×32 px の画像サイズでは、「High」と「Low」それぞれに疎な構造と密な構造が対応していた。

このことから 2 番のサンプルの構造は他のサンプルのものとは比べて配向性は高いものの密度が不均一なため、画像サイズが大きい場合は疎密の判定ができず、配向性をもとに「Low」に多く分類され、画像サイズが小さい場合は疎密で判定できたため分類結果が拮抗したと考える。

また、2 番のサンプルにおける反射スペクトルは、図 4.1 を見ると 420 nm から長波長域にかけて反射率が上昇している。これは CNT の π 結合を形成する電子が入射光によって共振することで生じる光の影響で、CNT と平行な電界の成分が大きくなりやすい、配向性の低い 4 番のサンプルのような構造でみられる現象であるが、2 番のサンプルは配向性が高いにも関わらずこの現象が見られた。これは共振によって生じた光が CNT フォレスト内部の疎な構造を通過し、外に漏れ出たことで長波長域での反射率が高くなったと考える。

第5章 結論

5.1 CNT フォレスト断面構造の特徴量抽出と判別

5.1.1 目的との対応

描画範囲が異なる $64 \times 64 \text{ px}$, $32 \times 32 \text{ px}$, $16 \times 16 \text{ px}$ の3種類の画像サイズの構造画像に対して、オートエンコーダと混合ガウス法を用いた CNT フォレスト断面の構造特徴の抽出と異常検知による構造判別の性能を評価した。

結果としては、すべての画像サイズで類似した構造ごとに画像を分けることができしており、 $64 \times 64 \text{ px}$ と $32 \times 32 \text{ px}$ では密度構造に関する異常な構造を判別することができしており、 $16 \times 16 \text{ px}$ では密度構造が分けられたことで CNT の有無や向きといった細かな構造について判別できていることが明らかになった。また、画像サイズを小さくするにつれて構造が細かく分かれるため、クラスタ数を多く指定する必要があることが示唆された。

このことから、密度の異常構造を判定するには大きな画像サイズが適しており、細かな構造を判別するには小さな画像サイズを用いることが適しているがクラスタリング時のクラスタ数を多く指定する必要があると結論づける。

5.1.2 今後の展望

今後の展望として次の3つを挙げる。

1つ目は、クラスタリングの最適化である。 $16 \times 16 \text{ px}$ の画像サイズにおけるクラスタリングで、クラスタ間に属するデータの割合が一様であったことから、設定したクラスタ数が不適で異なる構造が一つのクラスタに属している可能性がある。そのため、他のクラスタ数を設定してクラスタリングを行い最適化を行いそのうえで構造を改めて評価することは重要だと考える。

2つ目は、特徴抽出に用いるモデルの条件の公平にすることである。特徴抽出に用いたオートエンコーダのモデルについて、層数が画像サイズごとに異なっているため、モデルの性能に差が出た可能性がある。そのため、異なる画像サイズであっても特徴抽出を行えるようなモデルを構築し、条件を揃えた上で再度上記の手法による解析を行うことは重要だと考える。

3つ目は、物性との相関を解析することである。今回の実験ではクラスタごとの構造の評価を行った。その発展として、サンプルにどのクラスタが多く含まれているかを解析し、それをもとに物性との相関を解析することは重要だと考える。

5.2 光学反射率の分類による光学特性の影響を与える構造の予測

5.2.1 目的との対応

従来報告の無い、機械学習を用いた CNT フォレストの光学反射率の予測を行うために、反射率の高低に関して $64 \times 64 \text{ px}$ と $32 \times 32 \text{ px}$ の 2 種類の画像サイズの断面画像を分類し、分類された画像に含まれている構造を評価した。

結果としては、配向性が悪い構造画像は反射率が高いラベルに、配向性の良い構造画像は反射率が低いラベルに分類されており、従来の構造と物性の関係と同じ傾向になったことが明らかになった。また、 $32 \times 32 \text{ px}$ の画像サイズにおいて、分類されたラベルごとの画像枚数が拮抗していたサンプルの構造画像と反射率スペクトルとの相関から、配向性の良い構造でありながら CNT 内部で反射した光が低密度構造によって CNT フォレストから漏れたことで長波長域にかけて反射率が増加するスペクトルが得られたと考察した。

以上のことより機械学習を用いて、CNT フォレストの構造から物性予測を行うこと、得られた結果をもとに相関を解析することは可能であると結論づける。

5.2.2 今後の展望

今後の展望について次の 3 つを挙げる。

1 つ目はモデルの汎化性能の検証を行うことである。本研究ではデータセットを訓練用、検証用、テスト用に分けてモデルの訓練を行ったが、偶然高精度な結果が得られた可能性がある。そのため訓練用と検証用データの組み合わせを変えて交差検証による汎化性能の評価を行う必要がある。

2 つ目は分類ラベルの付け方を改善することである。本研究では特定波長での反射率の高低予測を行ったが、反射率スペクトルの波形は吸収の密度依存と長波長域反射の配向性依存 2 つの要素に分けられる。そのため、反射率の値だけでなく、反射スペクトルの波形によってもラベルを考慮することでどのような構造がどのラベルに分類されるかを評価することは重要だと考える。

3 つ目は分類に影響を与えた構造の場所を評価することである。今回の実験では、分類された画像の傾向を主観での配向性や密度の評価を行ったため、その部分が機械学習モデルの分類に影響を与えているか不明である。そのため、Grad-CAM [10]をはじめとした説明可能 AI の枠組みを用いてモデルの分類に影響を与えた領域を可視化し評価することは重要だと考える。

謝辞

本研究を進めるにあたり、学部 3 年次の配属以来、勉強会の開催や研究の進め方など、研究室で前例のない内容にもかかわらず多くの困難のなか相談に乗っていただいた古田寛教授に心より感謝申し上げます。

副査を務めていただき、ご指導いただいたプラズマ応用研究室の八田章光教授と Soft Intelligent SoC 研究室の星野孝総教授に感謝申し上げます。

CNT やメタマテリアルにおけるご指導をしていただいた元助教 Adam Pander 先生(現 NTT 先端集積デバイス研究所・現高知工科大学客員講師)に深く感謝いたします。

講義や研究活動などでお世話になった電子系の橘昌良教授、牧野久雄教授、密山幸男教授、山本真行教授、李朝陽教授、小林弘和教授、田上周路准教授、野田聡人准教授、岩下克名誉教授、綿森道夫先生、就職活動において面接練習や履歴書の添削などサポートしていただいた元教育講師の福島史郎先生心より感謝申し上げます。

電子事務室の中山愛様、谷口真奈美様ほか事務局の方々に感謝申し上げます。

本研究を進めるにあたり古田寛研究室の OB の沢田侑斗様、西森秀人様、亀岡伸義様、博士課程の先輩の Md. Saiful Islam 様、Muhammad Shahbaz 様、研究室の同期である山本十夢様、研究室後輩の大江亮介様、坂田健伸様、迫田北斗様、杉山洸生様、佐藤佑紀様、中村迅様、フロリダ国際大学から短期留学に来られた Aaron Font 様、Luciana Di Campli 様に深く感謝いたします。

最後に、大学生活を共に過ごしてきた学友、そして、生前に成果を伝えることがかなわなかった祖父、祖母、父に改めてここで報告するとともに、その中で自身が多忙だったにもかかわらず応援し、支えていただいた母と姉に深く感謝を申し上げます。

本研究は JSPS 基盤研究(C) JP23K04383 の支援を受けました。

参考文献

- [1] A. Pander, K. Ishimoto, A. Hatta and H. Furuta, "Significant decrease in the reflectance of thin CNT forest films tuned by the Taguchi method," *Vacuum*, vol. 154, pp. 285-295, 2018.
- [2] Y. Iwasaki, A. G. Kusne and I. Takeuchi, Comparison of dissimilarity measures for cluster analysis of X-ray diffraction data from combinatorial libraries, vol. 3, 2017, p. 4.
- [3] K. Sodeyama, Y. Igarashi, T. Nakayama, Y. Tateyama and M. Okada, Liquid electrolyte informatics using an exhaustive search with linear regression, vol. 20, Royal Society of Chemistry, 2018, pp. 22585-22591.
- [4] T. Honda, S. Muroga, H. Nakajima, T. Shimizu, K. Kobashi, H. Morita, T. Okazaki and K. Hata, Virtual experimentations by deep learning on tangible materials, vol. 2, Springer Nature, 2021, p. 88.
- [5] D. Reynolds, "Gaussian Mixture Models," 2009. [Online]. Available: http://leap.ee.iisc.ac.in/sriram/teaching/MLSP_16/refs/GMM_Tutorial_Reynolds.pdf. [Accessed 30 1 2025].
- [6] D. P. Kingma and J. L. Ba, Adam: A Method for Stochastic Optimization, International Conference on Learning Representations, ICLR, 2014.
- [7] K. He, X. Zhang, S. Ren and J. Sun, Deep Residual Learning for Image Recognition, Vols. 2016-December, IEEE Computer Society, 2015, pp. 770-778.
- [8] R. Shinsei, H. Furuta, "Anomaly detection of CNT forest cross-sectional SEM image using machine learning," FNTG67, Eikokuji Campus, Kochi Univ. of Tech., 2024.
- [9] R. Shinsei, H. Furuta, "Machine Learning-Based Cross-Sectional Structural Analysis of CNT Forests and Prediction of Structural Factors Affecting Optical Reflectance," FNTG68, Tenpaku Campus, Meijo Univ., 2025.
- [10] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, vol. 128, Springer, 2016, pp. 336-359.

業績

主著

1. R. Shinsei and H. Furuta, “Machine Learning-Based Cross-Sectional Structural Analysis of CNT Forests and Prediction of Structural Factors Affecting Optical Reflectance”, FNTG68 (Mar. 5, 2025, Poster, Tenpaku Campus, Meijo Univ.)
2. R. Shinsei and H. Furuta, “Anomaly detection of CNT forest cross-sectional SEM image using machine learning”, FNTG67 (Sep. 3, 2024, Poster, Eikokuji Campus, Kochi Univ. of Tech.)
3. R. Shinsei and H. Furuta, “Autoencoder machine learning for feature extraction and clustering of cross-sectional SEM images of CNT forests”, FNTG65 (Sep. 6, 2023, Poster, Nishijin Plaza, Kyushu Univ.)
4. R. Shinsei and H. Furuta, “Feature extraction and clustering using autoencoder from CNT forest cross-sectional SEM images”, FNTG64 (Mar. 2, 2023, Poster, Nagoya Univ.)
5. 神生 龍一, 古田 寛, “CNT フォレスト断面の局所構造特徴の自動判別に適した画像サイズ最適化”, 2024 年度 応用物理・物理系学会 中国四国支部 合同学術講演会(2024 年 7 月 28 日(日) 徳島大学(常三島キャンパス))
6. 神生龍一, 古田 寛, “機械学習を用いた CNT フォレスト断面 SEM 画像からの特徴量抽出とクラスタリング”, 2023 年度 応用物理・物理系学会 中国四国支部 合同学術講演会 (2023 年 7 月 29 日(土) 島根大学(松江キャンパス))
7. 神生龍一, 沢田侑斗, 西森秀人, Rebekah Arias, 古田寛, ” 機械学習を用いた CNT フォレスト断面 SEM 画像からの特徴抽出”, 2022 年度 応用物理・物理系学会 中国四国支部 合同学術講演会,(2022.07.30, 香川大学教育学部 (幸町キャンパス))

共著

1. Yuki Sato, Ryuichi Shinsei, Hiroshi Furuta, ”Memristive resistance in Transverse Electrical Conduction in CNT Forest”, FNTG67 (Sep. 2, 2024, Poster, Eikokuji Campus, Kochi Univ. of Tech.)
2. Hiroshi Furuta^{1,2}, Taichi Marui¹, Rebekah Arias³, Yuto Sawada¹, Ryuichi Shinsei¹, Nobuyuki Kameoka¹, Tomu Yamamoto¹, Kenshin Sakata¹, Ryosuke Ooe¹, Hokuto Sakoda¹, Nathan Chung³, Nezhil Pala³, Chunlei Wang³ (1. School

of Systems Eng., Kochi Univ. Tech. (Japan), 2. Res. Inst., Kochi Univ. Tech. (Japan), 3. Florida Int'l. Univ. (United States of America)), “[G1-O402-04] Plasma assisted formation of nano catalysts for the high-density growth of carbon nanotube forests and its device applications”, MRM2023/IUMRS-ICA2023, (Dec. 14, 2023, Oral, (Kyoto International Conference Center, Kyoto)

2. 山本 十夢, 神生 龍一, 迫田 北斗, 大江 亮介, 佐藤 佑紀, 古田 寛, “DC パルススパッタ堆積による高面密度触媒微粒子形成”, 2024 年度 応用物理・物理系学会 中国四国支部 合同学術講演会 (2024 年 7 月 28 日(日) 徳島大学(常三島キャンパス))
3. 大江亮介, 神生 龍一, 山本十夢, 古田 寛, “ 小直径高密度 CNT 成長に及ぼす基板下地前処理の効果 ”, 2024 年度 応用物理・物理系学会 中国四国支部 合同学術講演会 (2024 年 7 月 28 日(日) 徳島大学(常三島キャンパス))
4. 佐藤 佑紀, 神生 龍一, 古田 寛, “CNT フォレスト構造体の 横方向電気伝導におけるメモリスタ現象”, 2024 年度 応用物理・物理系学会 中国四国支部 合同学術講演会 (2024 年 7 月 28 日(日) 徳島大学(常三島キャンパス))
5. 大江亮介, 亀岡信義, 神生龍一, 山本十夢, 坂田健伸, 古田寛, “スパッタ環境内の変化による、触媒微粒子への影響”, 2023 年度 応用物理・物理系学会 中国四国支部 合同学術講演会 (2023 年 7 月 29 日(土) 島根大学(松江キャンパス))

付録

A モデル条件の選定について

第 4 章のモデルの条件の決め方について、 32×32 px の画像サイズにける活性化関数・学習率・ドロップアウト率・荷重減衰のハイパーパラメータの組み合わせを、訓練時の損失の値をもとに調整した。また、訓練時の他のハイパーパラメータについて、最大のエポック数を 10000、早期終了するエポック数を 50 とした。

これらのハイパーパラメータのうち、最初に活性化関数を SELU 関数・ReLU 関数・SiLU 関数の 3 種類から選定を行った。この時、学習率は 0.0001、ドロップアウト率は 0、荷重減衰は 0 に設定して学習を行った。SELU 関数は第 2 章で述べた式(2.6) で表される関数で、ReLU 関数は正の入力はそのまま、負の入力は 0 を返す式(A.1)で表される関数である。SiLU 関数は入力が十分大きい場合と十分小さい場合では ReLU 関数と似た入出力の関係を持つが 0 で滑らかになっていることが特徴の式(A.2)で表される関数である。

$$\text{ReLU}(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (\text{A.1})$$

$$\text{SiLU}(x) = \frac{x}{1 + e^{-x}} \quad (\text{A.2})$$

それぞれの活性化関数を用いた場合の訓練時の損失の推移を図 A.1 に示す。色の濃い曲線が検証データにおける損失、薄い曲線が訓練データにおける損失であり、点線は生の損失の推移、実線は 5 点での移動平均をとった推移である。SELU 関数の場合は両方のデータで徐々に損失が減少しているのに対して、ReLU 関数と SiLU 関数では数エポック経ったのちに検証データの損失が増加し過学習の傾向がみられ、そのまま 50 エポックを過ぎたため早期終了が行われ

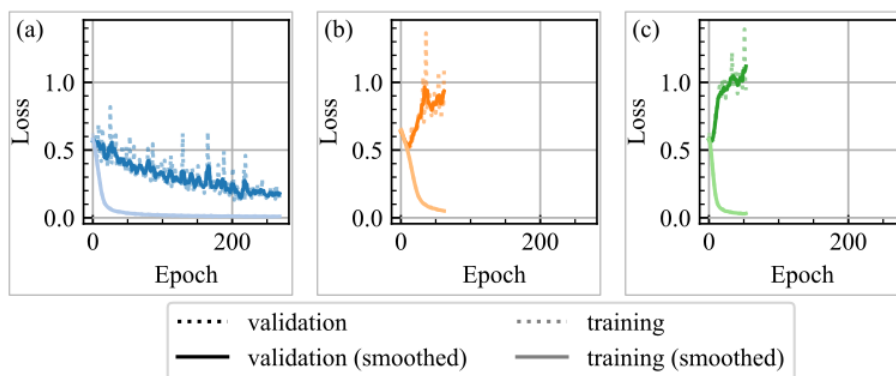


図 A.1 3 種類の活性化関数それぞれを用いた場合の損失関数の推移. (a) SELU 関数. (b) ReLU 関数. (c) SiLU 関数.

た。この結果から、過学習が起きず損失が減少したSELU 関数を採用した。

活性化関数の次に、学習率・ドロップアウト率・荷重減衰を調整した。その際に試した組み合わせを表 A.1 その訓練時の検証データにおける損失を図 A.2 に示す。条件 1 と 4 では学習初期で検証データでの損失が急上昇しており過学習が起きており、学習率が小さかったことで局所解に陥り、荷重減衰が小さいことで荷重の値柄の影響が小さく汎化性能が下がったためだと考える。条件 3 と 5 と 6 では急上昇することはなかったが停滞しており、これは学習率を大きかったため荷重の更新幅が大きくなり、更新後の荷重の値が極小値をまたいだもの担っていたと考える。条件 2 は他の条件では早期終了していたエポックでも損失が減少しており、結果的に最も損失が小さくなった。

以上の結果より、損失関数には SELU 関数、学習率には 0.0001、ドロップアウト率には 0.3、荷重減衰には 0.001 を採用した。

表 A.1 学習率・ドロップアウト率・荷重減衰の組み合わせの条件

条件番号	学習率	ドロップアウト率	荷重減衰
1	0.0001	0.3	0.0001
2	0.0001	0.3	0.001
3	0.001	0.3	0.0001
4	0.0001	0.5	0.0001
5	0.003	0.3	0.0001
6	0.001	0.3	0.001

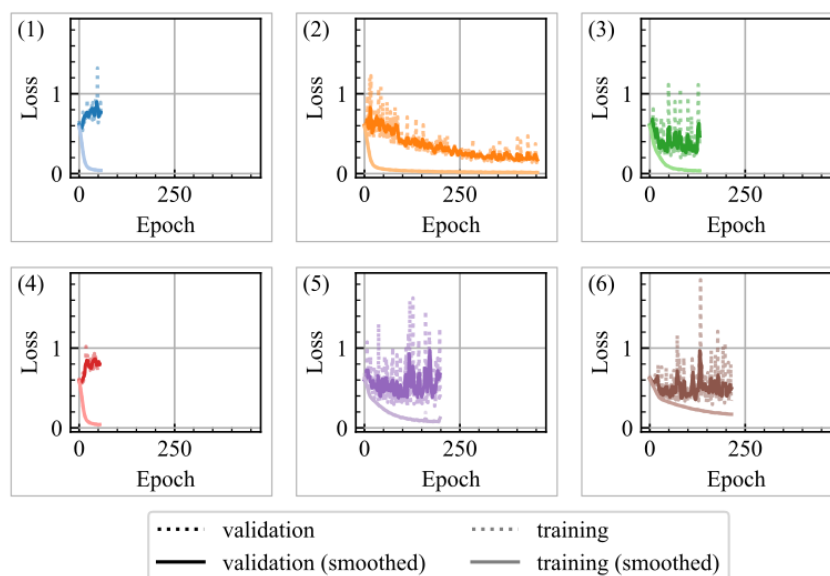


図 A.2 学習率・ドロップアウト率・荷重減衰の組み合わせにおける損失の推移。左上の数字は表における条件番号。