

深層学習による

分光反射スペクトルからの変質鉱物同定

田中 壮一郎

深層学習による
分光反射スペクトルからの変質鉱物同定
(Identification of Alteration Minerals from Reflectance
Spectra Using a Deep Learning Method)

田中 壮一郎
(Soichiro TANAKA)

名古屋大学大学院環境学研究科
地球環境科学専攻学位論文 博士(理学)

2019 年

主論文の要旨

本研究は、深層学習を用いて分光反射スペクトルデータから変質鉱物同定を自動的に行う方法の開発と、その有効性の検討を目的としている。黄銅鉱などの鉱石鉱物そのものは変質鉱物より広がり狭く、地表に露頭が出ていない限り鉱床探査を行う上で直接的な探査指標にはならないため、一般に広がり示す変質鉱物を探査指標として追跡する。このため、変質鉱物同定は金属鉱床探査において非常に有用である。鉱床形成時、鉱床周辺の変質帯には特定の変質鉱物が広く分布し、時には同心円状を示し、中心から酸性変質帯、中性変質帯、アルカリ性変質帯などの変質ハローを示す。これらの変質鉱物は、短波長赤外域に特徴的な吸収ピークを持ち、金属鉱床探査において重要な指標となる。

変質鉱物の吸収ピークは鉱物特有の波長位置に認められ、時には吸収ピークの位置が数 nm の違いで他の変質鉱物と区別される。しかし、近年普及が進んでいる安価な分光器では、波長校正後も波長方向の誤差が数 nm 程度ある場合が認められ、誤同定の要因になっている。また、目視で変質鉱物を同定する際には、どの波長域の吸収スペクトルのどのような性状に注目すべきかを、鉱物の反射スペクトルに関する専門知識を持った人間が判断しなくてはならない。

深層学習は人間の脳神経回路をモデルにした多層構造アルゴリズムで、自ら学習しデータに潜む特徴量を抽出する。この深層学習を用いることで、波長方向の誤差の問題及び吸収スペクトルのどこに注目して特徴を取り出すかは、深層学習に任せることができる。

本研究では、熱水性鉱床探査の指標となる変質鉱物など 24 種類を同定対象とし、各変質鉱物を 30 回測定した後、これらのデータに対して 50 倍のデータ拡張（オーギュメンテーション）を実施した。また分光反射スペクトルデータをそのまま用いる場合と、分光反射スペクトルデータに包絡線（以下、HQ: Hull quotient）処理を適用したものを、用いる場合の二つのケースを検討した。データ数は 1 セット 36,000 個で二つのケースで総計 72,000 個である。深層学習の多層構造アルゴリズムとしては、多層パーセプトロン（以下、MLP: Multi-Layer Perceptron）と畳み込みニューラルネットワーク（以

下、CNN：Convolutional Neural Network）の二種類のネットワーク手法を鉱物同定に適用した。ネットワークの検証には交差検証法を用いた。データを5分割して4分割分を学習データとして用い、残りの1分割分を検証用のテストデータとした。検証は5回行い、適合率（precision）と再現率（recall）の調和平均であるf値及び正答率の平均値で評価した。

MLPは順伝播型の代表的なニューラルネットワークで、誤差逆伝播法を適用した非線形教師あり学習手法で、1次元のデータも取り扱える。全結合層とドロップアウト層の単純な層構造の繰り返しで構成されているため、学習時間は比較的短い、パラメータ数が多く調整が難しい。MLPは、一般には過学習を起こしやすくノイズに弱く、収束が遅いとされている。MLPの検討において、HQ処理を施していないデータを学習データとし、Small層、Medium層、Large層の層構造の異なる3タイプについて変質鉱物の同定処理を行った。交差検証を行ったf値は、Small層、Medium層、Large層に対して96.1%、96.4%、96.2%の値をそれぞれ示した。さらにHQ処理を施したSmall層、Medium層、Large層に対して交差検証を行った結果、f値は97.5%、96.8%、97.1%の値をそれぞれ示した。このようにHQ処理を実施することで、0.4~1.4%のf値の上昇が確認された。またHQ処理を施したネットワークには、過学習を起こさずモデル化が可能であるものがあり、HQ処理が過学習抑制に貢献していることが明らかになった。各f値に大きな違いはないが、正答率・損失関数とエポック数のグラフからは、HQ処理を行った4層のSmall層（全結合層4層、ドロップアウト層3層）が過学習を起こさずに安定していた。さらに層構造を薄い構造にして、どの程度の構造が最適か確認したところ、Small層よりも薄い構造のHQ処理を行った3層のMLPが変質鉱物自動同定のためのモデル化に最適であることが判明した。

CNNは近年、画像識別競技会などで良好な成績を示しており、畳み込み層やプーリング層などの新しいテクニックを使えることが有利な点である。1次元CNNは時系列の連続データにも用いられており、分光反射率は連続する1次元データ群であるため、適用を試みた。2次元CNNは形態等を学習することができ、人間が反射率パターンから鉱物を同定するときの思考に類似した解析法と考えられている。CNNは、一般に過学習が起きにくいと考えられている。HQ処理を施していない画像解像度を落としたテ

ストデータに対して、2次元 CNN は 96.7%という f 値を示した。一方、1次元 CNN では f 値が 89.8%であった。これは 1次元 CNN の特徴から、分光反射スペクトルの形態や特徴量を精度よく抽出できていない可能性があることを示している。MLP と同様、HQ 処理を実施することで、f 値の上昇が確認された。1次元 CNN は 92.3%の f 値を示し、2次元 CNN で 97.4%の f 値を示した。しかし、いずれの CNN も、過学習が起き、またエポック後半で損失関数の動きに乱れがあり変質鉱物自動同定のためのモデル化は困難であった。

深層学習による分光反射スペクトルからの変質鉱物の同定を行った結果、いくつかの MLP で高い f 値(例えば MLP_S_HQ の 97.5%)を示した。これは学習を進めていくなかで、深層学習が自ら分光反射スペクトルに潜む識別概念を作成し、特徴を抽出できたことに他ならない。このことは深層学習が変質鉱物同定に有効であることを示している。

本研究では、MLP の層数と同定能力の関係については、MLP の層数が多くなるに従い過学習の傾向を示した。HQ 処理により、正答率・f 値が上昇し、MLP では MLP_S_HQ よりもドロップアウト層が一層少ない 3 層の MLP、MLP_S_HQ、および MLP_M_HQ で学習モデルが生成できた。また、MLP の同定鉱物数と測定回数の関係は、同定鉱物数が少ないほど正答率は上がり、また測定回数が多いほど正答率は上がった。

変質鉱物同定に深層学習手法を適用して、その評価を行う際に問題となるのが、各種のパラメータであり、データ拡張処理を行うときの乱数の与え方も重要である。どの程度の乱数を発生させるかによって、深層学習の結果に大きな影響を及ぼす。今後は変質鉱物の混合比などの連続値を予想する回帰問題でも、深層学習の有効性を確認したい。

作成された学習済みモデルは、市販の携帯型分光放射計による変質鉱物同定システムに深層学習プログラムとして組み込むことが可能であり、コンピュータに実装することにより、現実的な変質鉱物同定システムが製作できると考えられる。

Abstract

The purpose of this research is to develop a method to automatically identify alteration minerals from spectral reflectance data by using deep learning and to examine the effectiveness of this method.

Ore minerals, such as chalcopyrite, cannot themselves be a direct index for mineral exploration because such minerals are spatially confined and have no outcrops on the surface. Thus, in general, alteration minerals are traced as exploration indices. When an ore deposit is being formed, specific alteration minerals are widely distributed in the alteration zone around the deposit, sometimes forming concentric rings. This results in an alteration halo, comprising an acidic zone in the center, followed by a neutral zone, and then an alkaline zone. Such alteration minerals exhibit distinctive absorption peaks in the short-wavelength infrared region, which are important indicators, for an exploration of metalliferous deposits. The wavelength of the absorption peaks is unique to the mineral. Even a difference of a few nanometers can change the mineral identification. The recent use of inexpensive spectrometers has often been found to lead to misidentification of minerals because of their large wavelength errors of a few nanometers even after calibration.

Moreover, when alteration minerals are identified by visual inspection, a person with expertise in mineral reflection spectra is needed in order to determine which absorption spectral characteristics and which wavelength region should be the focus. Deep learning is a multilayered algorithm based on human brain neural networks that conducts self-learning and extracts features from data. This allows deep learning to be used to adjust errors in wavelength and determine which part of the absorption spectrum should be the focus for extraction of features.

In this study, 24 kinds of alteration minerals, which are indicators for hydrothermal deposits, were identified. There were two cases that used spectral reflectance data as-is and spectral reflectance data to which Hull quotient (HQ) envelope processing was

applied. Also, two types of network methods, a multilayer perceptron (MLP) and a convolutional neural network (CNN), were employed for mineral identification.

The f-values given below represent the harmonic mean precision and the recall in cross-validation obtained when each alteration mineral was measured 30 times. The accuracy rate was also used in this study. These data were augmented 50 times, and the total number of data was 36,000 per set.

A MLP is a typical feedforward neural network that uses a nonlinear supervised backpropagation method for training. MLP can also handle one-dimensional data. It comprises a simple, repeating structure consisting of a fully connected layer (hereafter “dense layer”) and a dropout layer.

Data without HQ processing is used as learning data. Small, medium and large layer structures are used to identify alteration minerals. After cross-validation, the respective f-values were 96.1%, 96.4%, and 96.2%. Cross-validation was also conducted on small, medium, and large layers that were subjected to HQ processing, and the f-values were 97.5%, 96.8%, and 97.1%, respectively. This indicates that HQ processing increases the f-values, although not very significantly. However, the graphs of accuracy rate, loss function, and epoch number indicated that four small layers (four dense layers, three dropout layers) that had been subjected to HQ processing were stable and did not cause overfitting. Moreover, the MLP with HQ processing was capable of modeling without overfitting. After investigating which structure would be optimal for the identification of 24 types of alteration minerals, it was found that three layers were found to be most suitable, where HQ processing had been used.

In recent years, CNNs have yielded good results in, for example, major image identification competitions. CNNs have the advantage of allowing the use of new techniques, such as a convolutional layer and a max pooling layer. A one-dimensional CNN is also used for time-series data, and its spectral reflectance is a continuous data group, a two-dimensional CNN is capable of recognizing shapes. It is thought that this is a situation in which deep learning can be applied. A two-dimensional CNN yielded

an f-value of 96.7% against test data with a reduced image resolution without HQ processing. On the other hand, a one-dimensional CNN yielded an f-value of only 89.8%, indicating a failure to accurately extract the features and characteristics of the reflectance spectrum. As in the case of the MLP, an increase in the f-value was found when a CNN was performed with HQ processing. A one-dimensional CNN yielded an f-value of 92.3%, and a two-dimensional CNN yielded an f-value of 97.4%. In each CNN, the movement of the loss function was disturbed up and down in the second half of the epoch. CNN also showed overfitting behavior. Consequently the author concluded that the CNN method could not be used for modeling to identify minerals.

We identified alteration minerals using deep learning. Some network models showed a high accuracy and f-value. This means that deep learning itself creates a discrimination that is hidden in the spectral reflectance data, and can extract feature quantities. In this study, deep learning was found to be effective in identifying alteration minerals.

The relationship between the number of MLP layers and the identification ability showed a tendency for overfitting as the number of MLP layers increased.

Identification results of MLP with HQ processing could be modeled without overfitting.

In addition, the smaller the number of identified minerals was, the higher the degree of accuracy became. It was found that the degree of accuracy also increased with an increasing number of measurements. The deep learning method was applied to the identification of alteration minerals. In the assessment of this method, the selection of parameters, as well as the random numbers used for data expansion processing, is important. The degree to which random numbers are used has a large effect on the results of deep learning.

In the future, I would like to use deep learning to investigate not only classification problems but also regression problems that predict continuous values, such as the mixing ratio for minerals. The present learning model can be incorporated into a commercially portable available altered mineral identification system. It is thought that

a realistic alteration mineral identification system can be developed by interfacing it with a computer.

目次

1 序論	1
2 既存研究	3
3 深層学習の技術要素	7
3. 1 前処理 (preprocessing)	7
3. 2 畳み込み (convolution)	7
3. 3 全結合 (full connected, dense)	8
3. 4 ゼロパディング (zero padding)	8
3. 5 最大プーリング層 (max pooling layer)	9
3. 6 活性化関数 (activation function)	9
3. 7 誤差逆伝搬法 (Back Propagation)	11
3. 8 過学習制御 (Over-Fitting Control)	13
3. 9 勾配降下法 (Gradient Descent method)	14
3. 10 勾配降下法の最適化アルゴリズム (Optimization Algorithm of Gradient Descent method)	15
4 データ	17
4. 1 対象とする鉱物のデータセット	17
4. 2 ハルコーシェント処理 (Hull quotient 処理(HQ 処理))	22

4. 3 データ拡張処理 (Augmentation (AUG 処理)).....	23
5 手法	25
5. 1 深層学習 (Deep Learning).....	25
5. 2 多層パーセプトロン (Multi-Layer Perceptron (MLP))	26
5. 3 畳み込みニューラル ネットワーク (Convolution Neural Network (CNN))	27
5. 4 研究の流れ.....	28
5. 5 モデルの評価法	31
6 結果	34
6. 1 HQ 処理なしの深層学習の結果.....	34
6. 2 HQ 処理済みの深層学習の結果.....	42
7 議論	49
8 結論	54
謝辞.....	57
引用文献	58

1 序論

熱水金属鉱床が生成されるときに熱水により母岩は変質作用を受け、熱水の物理化学条件により酸性変質帯、中性変質帯、アルカリ性変質帯が比較的広く形成される(e.g. 石川ほか, 1976)。各々の帯に出現する変質鉱物はその帯特有のものが多い。それらの変質鉱物は可視～短波長赤外域に特徴的な分光学的吸収ピークを示すことが良く知られている。野外で鉱物をサンプリングし、研究室でX線回折格子実験を行うことで変質鉱物の同定は可能であるが、時間と経費はそれなりに必要である。そのため従来から変質の有無、変質鉱物の種類、変質ハローの中心を知るために変質鉱物の同定を現場で即時にできるような機器の開発が望まれてきた。この結果、1990年初頭より通商産業省(通商産業省資源エネルギー庁, 1992)を中心にソフトウェア、ハードウェアの開発が進み、携帯型分光放射計による変質鉱物同定装置(以下、POSAM: POrtable SpectrorAdiometer for Mineral identification)が開発された。

近年普及が進んでいる安価な分光器を用いることで変質鉱物同定装置を比較的安価に製作することが可能であるが、これらの分光器に対して波長校正を実施しても波長帯によっては数 nm の波長方向の誤差が認められた。現在の変質鉱物同定ソフトは吸収ピークの位置を微分により抽出して解析を行うものが多い。波長誤差を有する分光器を数値処理で正しい同定ができるようにするには、マスターデータベースのすべての吸収許容値を書き換える作業等が必要であり、またプログラミング面でも IF 文を多用しなくてはならないため現実的ではない。したがって、分光器ごとに吸収ピークの位置などの特徴量を収集しなおす必要がある。

深層学習は多段のニューラルネットワークからなり学習データから自動的に特徴量を抽出する。深層学習がビッグデータを学習し、抽出したデータ間の相関関係の特徴量(画像認識の場合、例えば各々の顔の作りが特徴量になり、分光スペクトルの場合、吸収ピークの位置・形・深さ等が特徴量になる)を抽出する。従来は、データの特徴量を読み解き、手法の選択、分析を行うのは研究者の力量に負うところが多かったが、深層学習は研究者の力量を不要にして、より高度な解析を可能にしたことが近年の大きなブ

レークスルーである。しかし、深層学習による解析の一部分は証明がなされておらずブラックボックス化しており、この分野の同定に適用できるかは未知数である。

本研究は、深層学習を用いて分光反射スペクトルデータから変質鉱物同定を自動的に行う方法の開発と、その有効性の検討を目的とした。用いた手法は多層パーセプトロン（以下、MLP：Multi-Layer Perceptron）と畳み込みニューラルネットワーク（以下、CNN：Convolutional Neural Network）の二種類のネットワーク手法で、これらを鉱物同定に適用した。詳細は 5. 手法で説明する。

2 既存研究

可搬型変質鉱物同定装置 POSAM(通商産業省資源エネルギー庁, 1992)は分光器と制御・同定を行うパームトップコンピュータからなっており数値計算によって検出した吸収ピークにマスターデータベースから得点を与え、得点の高い順に表示するシステムである。マスターデータベースには鉱物名、基本総スコア、吸収ピークの波長位置とその許容範囲、及び吸収の鋭さなどが保存されている。ハードウェアは約 6kg で回折格子型の分光器を内蔵してバッテリー駆動で約 4 時間、野外にて連続使用が可能なものであった。ソフトウェアはパームトップコンピュータで制御されていた。特徴量を研究者が作成するという点でやや古典的な手法と言える。

田辺ほか(1994)は専用ニューロボードを用いて 3 層の順伝播型のニューラルネットワークを変質鉱物同定に用いた。6 鉱物種が同定可能であり、240 点の反射スペクトルデータを用いている。手法はパーセプトロンのみで CNN など是用いられていない。専用ハードウェアボードを用いており、入力ノードも最大 255 に制限されており限定的な研究となっている。

Montero et al. (2001)では FSTSpecID(Fast Spectral Identification Algorithm) processing、continuum removal method、smoothing and derivationなどの処理を行い変質鉱物の吸収中心(NIR)、吸収の深さ、スペクトルのコントラスト、吸収の対称性などを多次元の座標系として設定している。ベクトル化及び基準化された参照スペクトルライブラリを作成し変質鉱物の同定を行っている。特徴量を研究者が作成するという点で古典的な手法である。

Solar et al. (2008)はニューラルネットワークを用いて、岩石薄片試料をデジタル化した画像から 6 種類の鉱物（黄銅鉱、輝銅鉱、銅藍、斑銅鉱、黄鉄鉱、および硫砒銅鉱）の認識とこれらの鉱物の量比を決定している。鉱物試料のデジタル化された写真から選択された 160 枚の画像で学習を行い、システムの正答率はネットワークの学習で使われなかった 240 枚の写真で算出された。結果は、ネットワークを学習するために使用された画像の 97%が正しく認識されたことを示した。新しい画像の場合、ネットワークはサンプルの 91%を正しく認識することができたとしている。鉱物の量比を決める

のに専門家が行うと 12.5 日が必要であるが、このシステムを使うと 20 分で 400 枚の写真の自動判別が可能である。これらの鉱物は肉眼的に色が異なり判別も容易であることから、鉱物の同定と量比をニューラルネットワークによって検出することができたと考えられる。

Ishikawa and Gulick (2013)はラマンスペクトルデータベース(RRUFF™)を用いて、自律的な鉱物分類法を提案している。この研究では、ニューラルネットワークで構築された鉱物分類器が、ラマン分光法によって取得されたスペクトルデータを使用して、火成岩の組成を特徴付ける主要な鉱物を正確に区別するように学習できることを示した。これらの鉱物は、かんらん石、石英、斜長石、カリ長石、雲母、及び輝石である。鉱物分類器は平均で 83%の正答率を示しており、石英、かんらん石、及び輝石は 100%の正答率を示した。RRUFF スペクトルおよび RELAB 反射スペクトルの両方から得られた結果は、斜長石とカリ長石を区別することが困難であることを示唆している。例えば、RRUFF 試験セットの斜長石スペクトルの 77%が誤ってカリ長石に分類されている。これらの傾向は、両方の鉱物間の化学的類似性に起因する可能性が高いとしている。

Liu et al. (2017)はラマン分光の RRUFF™データベースを用いて深層学習の適用を試みている。この中で CNN と Support Vector Machine (以下、SVM)などいくつかの手法の比較を行っている。ラマン分光はベースライン補正がしばしば処理として用いられているが CNN との組み合わせでは良い結果を示していない。SVM と比べて CNN はよりスペクトル同定に有効であると指摘している。

Cheng and Guo (2017)は薄片画像から鉱物の特徴を選択抽出し、細分性を認識する分類器を構築することで、CNN を用いた粒状性解析を実施している。オルドス盆地で採取された 4,800 試料は、偏光顕微鏡の下でそれぞれ HSV、YCbCr、RGB の色空間で実験に用いられており、RGB の色空間において正等率は 98.5%で、HSV および YCbCr 色空間においても信憑性があると指摘している。CNN が高い精度で岩石画像を分類できることを示している。

衛星画像の分類を目的にした深層学習の応用は多くの研究があり、吉原ほか(2017)では Geoeye-1 画像を用いたもので事前実験として、5 つの分類クラスすべてを含むサンプル領域 A (256×256[pixels])を抽出し、分類精度の検証を実施している。サンプル

領域 A に対し、手作業でラベル付けを行う。分類するクラスは、Buildings, Water, Vegetation, Asphalt, Ground の 5 種類である。その後、5 つの各クラスからランダムに 1,000 ピクセルずつ選択し、これらの 8 割を学習用データ、2 割をテスト用データに分割する。CNN モデルへの入力、各学習用データ 1 ピクセルを中心とする周囲 1m×1m、すなわち 4,000 枚の 1m×1m サイズの画像とした。これらから土地被覆分類を行っている。分類は softmax 関数と SVM を用いている。SVM を用いた結果が CNN よりわずかに高い結果を示している。学習データに対する正答率は 95%を超えているもののオープンテストでは分類の正答率はそれほど高くない結果となっている。著者らはデータ数が増えていくと結果が逆転する可能性を示唆している。

三輪ほか (2017) では火山灰顕微鏡画像を深層学習が実装された NEC Advanced Analytics - RAPID 機械学習で解析した。代表的試料として阿蘇火山において 2014 年 12 月 15 日に採取された火山灰粒子を用い、学習器を作成し、分類実験を行った。各カテゴリに含まれる火山灰粒子について、計 50 枚の顕微鏡写真を取得し、色味、光沢の有無、気泡の有無など、それぞれの特徴を学習させた。そして作成した色味・光沢・気泡に関する三つ学習器をカテゴリ既知の粒子に適用し分類を行った。作成した学習器を 6 試料から得られた火山灰粒子顕微鏡画像へ適用することで、画像内に含まれる粒子の種類とその量比を評価した。火山灰粒子はそのガラス光沢の有無、色、気泡の有無に基づいて、5 つのカテゴリに分類した (BGD (黒色ガラスデンス)、BGV (黒色ガラス発泡)、BN (黒色光沢無し)、Al (変質)、PL (斜長石))。分類実験により三つの学習器が有効であることが示された。色味に基づいた学習器によって PL、AL、B 系統 (BGV, BGD, BN) が正答率 100% で分類された。また光沢の有無に基づく学習器によって BN と BG 系統 (BGV、BGD) が正答率 95% で分類された。さらに、気泡の有無に基づいた学習器によって、BGD と BGV を正答率 78% で分類することが出来たとしている。

現在のところ既存の研究で短波長赤外域での変質鉱物同定に深層学習を用いた研究は少ない (例えば、田辺ほか、1994)。また、公開されているライブラリにおいて、深層学習に適した短波長赤外域のデータは不足しており、ビッグデータを自ら新規で構築する必要がある。そのため、X 線回折分析を伴う変質鉱物を保有しており、スペクトルの分光放射計での測定およびビッグデータの作成ができ、そして深層学習を実施できる環

境が必要である。また、鉍床地質学の基礎知識も必要になるため、先行研究において同類の研究が少ないものと考えられる。

3 深層学習の技術要素

本章の以下の説明は主に岡谷貴之（2015）、神嶋ほか（2015）、斎藤康毅（2016）に基づく。

3. 1 前処理 (preprocessing)

前処理は深層学習においてしばしば用いられる重要な処理である。本研究では対象が分光反射率データであるため、各チャンネルの値は 0.0～100.0 の範囲に収まっている。特に二次元 CNN(以下、CNN2d)では前処理を行わない場合、学習が収束しないことがあった。そのため本研究ではデータの値を 0.0～1.0 にする正規化を行った。この他、データの平均や標準偏差を利用してデータの平均を 0 に標準偏差を 1 にすることがある。あるいは入力データから平均値を各ピクセルから引くなどの手法も用いられる。MLP では前処理なしで計算ができたが、一般的 CNN2d ではノイズ除去のための平滑化、正規化などの前処理は必要とされている。

3. 2 畳み込み (convolution)

畳み込みは CNN の代表的な手法で核心をなすものである。カーネル(フィルタ)と呼ばれる格子状の数値データをカーネルと同サイズの部分画像に対して要素ごとの積の和を計算して変換する処理である。この変換処理をずらして新しい画像に変換するものである。例えば 32×32 のピクセルを 5×5 のカーネルを 3 ピクセルずつずらしながら畳み込みを行うと 10×10 のピクセルに縮小される。これは下記のようなになる。

入力の高さを H、幅を W、出力の高さを OutH、幅を OutW としフィルタのサイズの高さを FiH、幅を FiW として、パディングを P、ストライドを S とすると

$$\text{OutH}=(H+2P-\text{FiH})/S+1=(32+2*0-5)/3+1=10$$

$$\text{OutW}=(W+2p-\text{FiW})/S+1=(32+2*0-5)/3+1=10$$

となる。

畳み込みでは入力値にフィルタを畳み込む計算を行う。これは一般的なフィルタ処理と変わらない。出力はぼかした画像、エッジの効いた画像などが出力として出てく

る。こうして画像の特徴量を抽出していく。畳み込みは有用な特徴量を画像の位置によって大きく変化させない効果がある。

3. 3 全結合 (full connected, dense)

CNN、MLP では全結合を用いるが、この問題点は 1 次元の入力となるためデータの形状が無視されてしまうことである。入力データが本研究のように分光反射スペクトルの場合、分光スペクトルのデータであるが全結合層に入力するときには 1 次元のデータになる。入力には吸収ピークの形状のデータなどが含まれているが、全結合層は形状に関する情報をわかりにくくしている。深層学習では出力層の手前で使われることが多く、この部分が検出された特徴の組み合わせから、予測結果に分類するための識別部となる。最終的な出力層のユニットの数は分類される数と一致している必要がある。

3. 4 ゼロパディング (zero padding)

カーネルサイズとストライドの設定によってはゼロパディングとよばれる画像の周りに数値の 0 を追加する必要があるが出てくる。パディングをしない場合に比べてパディングを実施した方が画像の端のピクセルまで利用される機会が増え、画像の端の特徴量を把握することができる。畳み込み演算の回数が増えるのでパラメータの更新が多く実行される。カーネルのサイズや、層の数を調整できる上、畳み込み層とプーリング層で出力サイズは次第に小さくなるので、ゼロパディングでサイズを増やすことで層の大きさを増やすことができる。

3. 5 最大プーリング層 (max pooling layer)

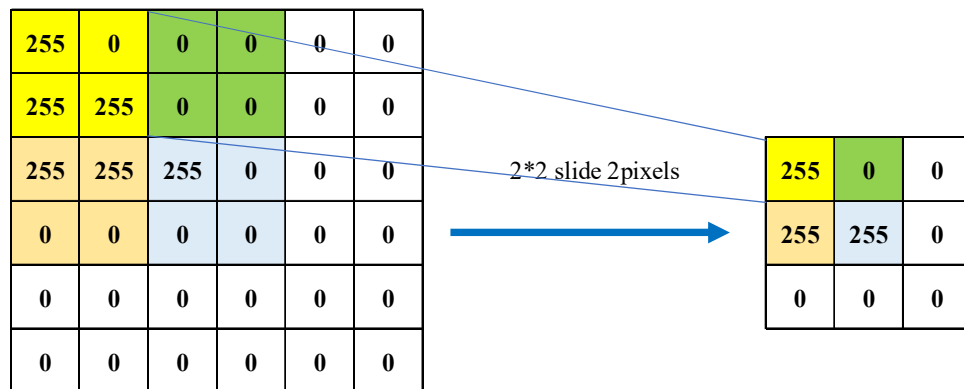


図 3-1 最大プーリング層の概念

プーリング層は複数の入力をまとめる役割を持つ。プーリング層には平均、加算、最大等の方法があり本研究では最大値を用いている。プーリング層は画像の部分パターンから複数の値の代わりに 1 つの値を抽出する。これは学習データに対する過学習の可能性を減少させるものである。最大プーリングは縦・横方向の空間を小さくする演算である。ここでは 2×2 の中での最大値を代表値として要素を集約している。最大プーリング層は微少な位置変化に対しても強く、分光反射スペクトルの吸収位置の誤差にも強いと考えられる。CNN では畳み込み層と最大プーリング層で特徴量を抽出していく。図 3-1 に最大プーリング層の概念を示す。

3. 6 活性化関数 (activation function)

活性化関数とは、ニューラルネットワークにおいて入力信号の総和を出力信号に変換する非線形な関数である。かつては単純パーセプトロンの出力層ではステップ関数が用いられ、中間層では正規化機能のあるシグモイド関数(Han & Moraga, 1995)などが用いられた。現在では誤差逆伝播で勾配が消失しやすいという問題点から中間層には ReLU(Rectified Linear Unit)関数(Nair, V. & Hinton, G. E., 2010)が用いられることが多い。勾配消失問題は、勾配(微分)が数値計算上極めて小さくなってしまい、学習が進まなくなってしまうという現象である。誤差逆伝播法では、積の形で勾配を求めていくため、途中の勾配が小さいと入力層付近の勾配は次第に 0 に近づいていく。

ニューラルネットワークは、分類問題では2値分類ではシグモイド関数、多値問題ではReLU関数、softmax関数(Bouchard, 2007)を用いる。特に出力層ではsoftmax関数は必須であり、下記の式で表される。

$$y_k = \frac{\exp(a_k)}{\sum_{i=0}^n \exp(a_i)}$$

$\exp(x)$ は e^x を表す指数関数であり、出力数が n 個ある場合、 k 番目の出力 y_k を求める計算式を示している。softmax関数の分子は入力信号 a_k の指数関数、分母はすべての入力信号の指数関数の和から構成される。softmax関数の出力は0.0から1.0の間の実数になり、出力の総和は1になる。この総和が1になるということは重要なことで、この性質によりsoftmax関数は確率として取り扱えることになる。出力値が最大のもので示したユニットが選択される。図3-2にsoftmax関数を示す。

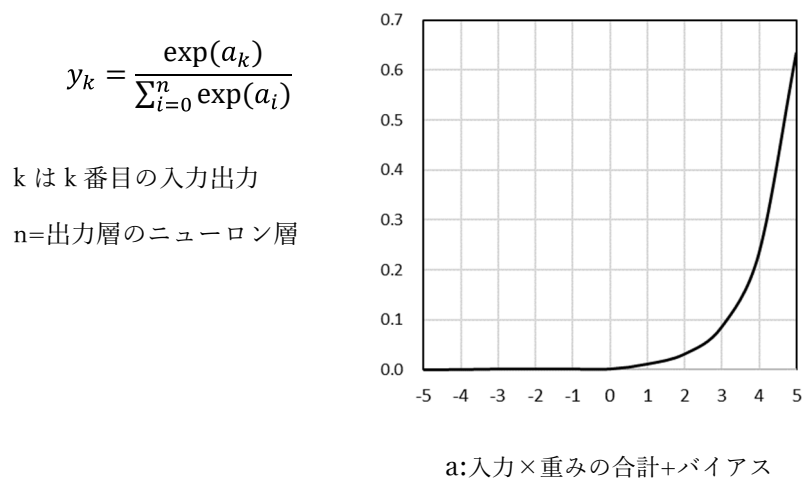
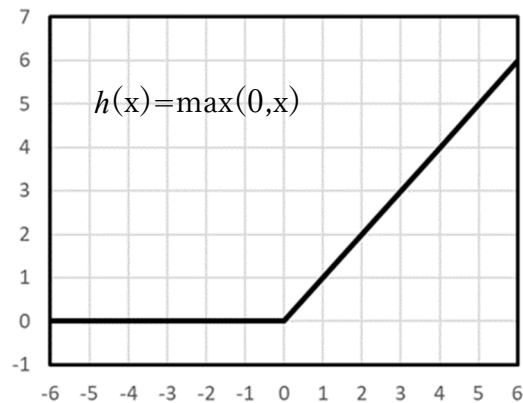


図 3-2 softmax 関数

ニューラルネットワークでは活性化関数に非線形関数を用いる必要がある。もし線形関数を用いると多層の利点を表すことができない。本研究で中間層にReLU関数を用いている。図3-3にReLU関数を示す。ReLU関数を数式で表すと下記のようなになる。

$$h(x) = \begin{cases} x & (x > 0) \\ 0 & (x \leq 0) \end{cases}$$

ReLU 関数を微分するとわかるように x が 0 より大きい場合、微分値は常に最大値である 1 が得られる。すなわち勾配消失しにくいといえる。なお、負の部分はゼロになっている。



x : 入力 \times 重みの合計 + バイアス

図 3-3 ReLU function

ReLU 関数は既存の活性化関数と比べて劇的に精度の向上があったために様々な派生系も考案された。そのひとつは Leaky ReLU 関数(Xu et al.2015)である。

$$h(x) = \begin{cases} x & (x > 0) \\ 0.01x & (x \leq 0) \end{cases}$$

3. 7 誤差逆伝搬法 (Back Propagation)

単純なニューラルネットワークの誤り訂正学習の限界は、出力層のノードへの結合の重みしか修正できないということであった。これに対して出力層よりも前の隠れ層への

ノードへの結合の重みを修正できる学習法として提案されたのが Rumallhart et al.(1986)による誤差逆伝搬法である。

誤差逆伝搬法ではノードの活性化関数を不連続なしきい関数から微分可能な関数に変更することで、出力層での誤差評価 $R(\mathbf{W})$ への任意の結合重み w の寄与を偏微分係数 $\partial R(\mathbf{W})/\partial w$ として計算できるようにした。ここで \mathbf{W} はネットワークのすべての結合の重みをまとめたものである。微分可能な活性化関数としてはシグモイド関数がよく使われる。

$$\text{sig}(x)=1/(1+e^{-x})$$

これによって、勾配降下法などの通常の最適化手法によって、出力と正解との間の誤差が小さくなるようにすべての重みを修正することが可能になった。

ネットワークの層の数を N 、第 n 層の第 k ノードから第 $n+1$ 層の第 j ノードへの結合の重みを $w_{k,j}^{(n+1)}$ とするとき、確率的勾配降下法(stochastic gradient descent method (SGD))による結合の重みの更新は以下のようになる。

$$w_{k,j}^{(n+1)} \leftarrow w_{k,j}^{(n+1)} - \eta \frac{\partial R(\mathbf{W})}{\partial w_{k,j}^{(n+1)}}$$

例えば、出力層での誤差評価 $R(\mathbf{W})$ が二乗誤差 $\frac{1}{2} \sum_j (t_j - y_j^{(N)})^2$ であるとき、具体的な更新則は

$$w_{k,j}^{(n+1)} \leftarrow w_{k,j}^{(n+1)} - \eta \delta_j^{(n+1)} y_k^{(n)}$$

と書ける。

ここで $y_k^{(n)}$ は第 n 層の第 k ノードの出力である。

この式の中の $\delta_j^{(n+1)}$ は、合成関数の微分法に従い、出力層での二乗誤差

$\frac{1}{2} \sum_j (t_j - y_j^{(N)})^2$ から計算される $\delta_j^{(N)}$ から出発して、以下のように再帰的に効率よく計算できる。

$$\delta_j^{(N)} = -(t_j - y_j^{(N)}) y_j^{(N)} (1 - y_j^{(N)})$$

$$\delta_j^{(n)} = \left\{ \sum_{k=1}^{K_{n+1}} \delta_k^{(n+1)} w_{k,j}^{(n+1)} \right\} y_j^{(n)} (1 - y_j^{(n)})$$

ここで、 K_{n+1} は、第 $n+1$ 層のノードの数である。この計算過程が、出力層での誤差を前の層に伝播させてゆく形になっていることが、誤差逆伝播学習という名前の由来である。

多層のニューラルネットワークを勾配降下法によって学習させるという考え方は1960年代にも提案されていたが (Amari, 1967)、この方法には誤差を極小にする重みの値 (local optimum) に収束してしまい必ずしも誤差を最小にする解が求まるわけではないという問題点があるため、あまり実用的とは考えられなかった。

この問題は Local convergence と呼ばれる。また、 $\delta_j^{(n+1)}$ の効率のよい計算法は1970年代に提案された合成関数の高速微分法の一つともいえる。しかし、Rumelhart et al.(1986)がこの手法をいくつかの具体的問題に適用して学習がうまく進み、興味深い内部表現が得られることを示したことがきっかけで、様々な問題への適用が試みられると同時に、勾配降下法以外の最適化手法の適用など様々な学習アルゴリズムが提案されるようになった (神畠 ほか、2015, pp.14~pp.15)。

3. 8 過学習制御 (Over-Fitting Control)

過学習(Over-fitting)は訓練データに対しては良く学習・特化され高い正答率を示すが、未知のデータに対しては汎化できない状態をいう。深層学習において重要なのは与えられたデータを正確に同定できることではなく、これから与えられる未知のデータに対する同定予測精度を上げることである。図 3-4 に過学習の状態を示す。次のテストデータが青い点を通る黄色の曲線に乗らなければ正しく同定できないことになる。ニューラルネットワークは表現能力が高いモデルであるため過学習しやすい。未知のデータすなわちテストデータに対して正答率が高くないと有益なモデルとはいえない。過学習を回避するために、いくつかの手法が考えられており、ドロップアウト、正則化、early stopping 関数などが挙げられる。ドロップアウトは CNN や MLP において中間層の中のノードを半分のノードを消した状態で学習を行う。これにより学習時のニューラルネットワークの自由度を小さくして汎化性能を向上させて過学習が押さえることができる

る。ドロップアウトは Srivastava et al.(2014)によって提案された比較的新しい手法である。

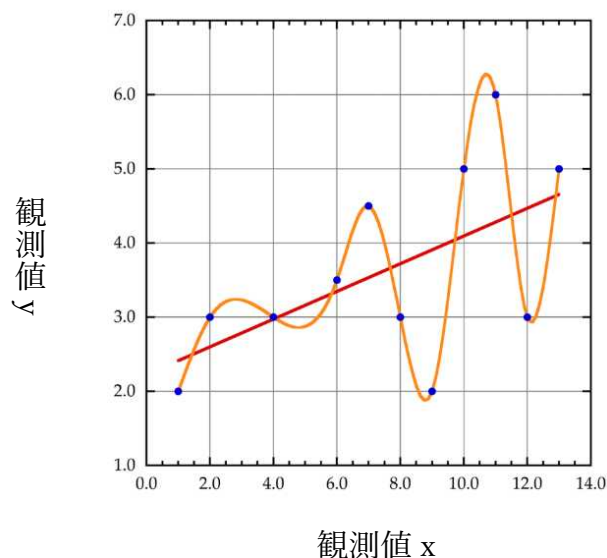


図 3-4 過学習を示す状態

正則化は学習過程でいくつかのパラメータが自動的に機能しないようにする手法である。具体的には損失関数にパラメータの大きさに対するペナルティ項(正則化項)を含めるものである。これによりパラメータを自由に動けなくして、少数のパラメータでフィットさせていくものである。

Early stopping 関数は過学習を防止する機能がある。テストデータに対する損失関数の値に対して減少が停止したときに学習を止めるというものである。

3. 9 勾配降下法 (Gradient Descent method)

勾配降下法には下記の数種類が使われることが多い。最急降下法(Fletcher and Powell, 1963)、バッチ勾配降下法(Bengio et al., 1994)、確率的勾配降下法(Bottou, 2010)、ミニバッチ確率的勾配降下法(Li et al., 2014)などである。

最急降下法は非常に広い範囲の問題の最適な解を見つけ出せる汎用性の高いアルゴリズムである。学習率が重要なパラメータで学習率が小さすぎると収束までの反復数が増えることになり、逆に学習率が大きすぎると発散してしまうことになる。バッチ勾配降下法は欠点があり勾配を計算するためには各ステップで訓練データをすべて使うため、訓練データが巨大であると計算速度が極端に悪くなることである。確率的勾配降下法は訓練セットから無作為にひとつを取り出してそれだけで勾配計算を行う。このため計算は極めて高速に済むが確率的な性質のためかなり不規則な動きをする。時間と共に最小値に近づくが最小値で落ち着く確率は高くなく周辺を探索していることになり最終的なパラメータは十分ではあるが最適なものではない。確率的勾配降下法の場合、一部のデータを用いるため勾配ベクトルは正確には計算されずランダムに移動していくが偶然極小値の谷から出て最小値の深い谷底に落ち着く可能性もある。ミニバッチ勾配降下法では一回あたりのデータ量を減らして最適化処理を何度も繰り返すというものである。本研究では確率的勾配降下法を用いた。

勾配降下法の手順は重みとバイアスの初期化、データをネットワークに読み込ませる、ネットワークの出力と正解ラベルの誤差を計算する、誤差が少なくなるように重みとバイアスを修正する、そして最適な解が出てくるまで繰り返すことになる。

3. 10 勾配降下法の最適化アルゴリズム (Optimization Algorithm of Gradient Descent method)

Momentum(Rumelhart et al. , 1986)、Adam(Kingma & Jimmy, 2014)、Adagrad(Duchi et al., 2011)、Adadelta(Zeiler, 2012)及び RMSProp(Hinton et al., 2014)などいくつかの手法が提案されている。勾配降下法における最適化において学習率の調整は重要である。学習率を高く設定すると訓練は発散してしまう場合があり、学習率を低く設定しすぎると訓練に非常に時間がかかってしまう。最急降下法は原位置での勾配を計算し、その勾配に学習率を掛けたもので位置を更新するものである。Momentum は現在の位置での勾配に学習率を掛け、前回の更新量に一定の減衰率を掛ける。そして両方を用いて位置を更新するものである。慣性が大きくならないように摩擦パラメータを有している。

Adagrad はまれなパラメータに対してはより大きな更新を、頻出のパラメータに対してはより小さな更新を実行するものである。Adagrad の主な弱点は、分母の二乗勾配の蓄積である。どの加算項も正の値なので、訓練の間、累積和は増加し続け、次々に学習率が低下し、最終的に極めて小さくなることである。この問題を解決するために提案されたのが RMSProp や Adadelata である。RMSProp は過去のすべての勾配を均一に加算していくのではなく、過去の勾配を徐々に使わないようにして新しい勾配の情報が大きく反映するように加算していくものである。Adam は Momentum と同じように勾配の減衰平均と RMSProp の過去の勾配の二乗の減衰平均を用いるものである。Wilson et al.(2017)は AdaGrad、RMSProp 及び Adam が一部のデータセットについて汎化性能が十分ではない可能性を示した。

Adadelata では次元量の違うもので計算していた他の手法のミスマッチの問題を解決している。Adadelata は Adagrad の発展形で、学習パラメータが未探索領域に入った時に、探索距離を広げることができる。また累積する際に直近の勾配に重みをおいているため、本研究では Adadelata を用いた。

4 データ

4. 1 対象とする鉱物のデータセット

収集したスペクトルデータは表 4-1 に示す主に熱水鉱床生成に起因する 24 種の変質鉱物等である。また図 4-1 に代表的な変質鉱物の分光反射スペクトルを示す。これらの鉱物が同定できれば主要な変質帯は評価できると考えられた。データセットは Web 上等デジタルで入手したものと、筆者が変質鉱物を分光放射計で測定したものに分けられる。データ検討・収集は、浦井ほか(1989)、坂野・古宇田(2014)、METI Agency for Natural Resources and Energy (2001)、Clark et al. (2007)、Baldridge (2009)等から行った。インターネット上では USGS をはじめとして多様な変質鉱物のスペクトルチャートを見ることができるが、Clark et al.(2007)の USGS の Spectral Library では粒度を変化させ測定されているものがあるものの、可視から短波長赤外、熱赤外へと一般に多様な波長域の鉱物・岩石・物質を測定することに主眼が置かれており短波長赤外域のビッグデータについての情報は限られていた。

深層学習に学習させるビッグデータを生成するために新たに標準的な試料を収集し、独立行政法人 石油天然ガス・金属鉱物資源機構所有の Malvern Panalytical 社 (ASD) FieldSpec4Hi-Res を使用して変質鉱物の分光反射スペクトル測定を行った。機器の仕様を表 4-2 に示す。分光反射スペクトルの測定は参照板の明るさを変化させる、鉱物の様々な向きの測定を行う、粒度を変化させるなどして約 600 個の測定データを追加取得した。オンライン、オフラインで収集・検討したスペクトルデータ、新たに測定したスペクトルデータは最終的に 2,400 個余り 85 種類であった。これらのうち 1 鉱物につき 30 回以上測定があるものを選択した。最終的に 24 鉱物について 30 回の測定を収集・測定したため 720 データがオリジナルのデータ数である。これらのデータは波長方向のサンプリング間隔が多様であり、深層学習を実施する前に波長域 400-2,500 nm で 1 nm おきのデータにそれぞれ変換してデータの基準化を図った。

また、深層学習に用いた波長域は可視域を除く変質鉱物が特徴的な吸収スペクトルを示す短波長赤外域を中心に 900-2,500 nm とした。

表 4-1 ターゲットとした変質鉱物

	mineral name	type	alteration zones	temperature	remark
1	alunite	hydrothermal alteration	acidic alteration zones	low-high	
2	anhydrite	hydrothermal alteration	acidic alteration zones	high	propylite zone
3	calcite	hydrothermal alteration	intermediate zones	medium-high	
4	chlorite	hydrothermal alteration	intermediate zones	high	propylite zone
5	dickite	hydrothermal alteration	acidic alteration zones	medium-high	pyrophyllite zone
6	dolomite	skarmization	-	-	
7	epidote	hydrothermal alteration	intermediate zones	high	propylite zone
8	gibbsite	bauxite	-	-	
9	goethite	hydrothermal alteration	mineralization, gossan	-	
10	gypsum	hydrothermal alteration	kuroko type	-	
11	halloysite	hydrothermal alteration	acidic alteration zones	low	
12	hectorite	lithium	-	-	
13	illite	hydrothermal alteration	intermediate zones	medium-high	
14	jarosite	hydrothermal alteration	-	low	
15	kaolinite	hydrothermal alteration	acidic alteration zones	medium	
16	montmorillonite	hydrothermal alteration	intermediate zones	low	
17	muscovite	hydrothermal alteration	intermediate zones		
18	natrolite	hydrothermal alteration	alkali zones	medium	
19	nontronite	hydrothermal alteration	basic alteration	low	
20	phlogopite	hydrothermal alteration	-	-	skarmization
21	pyrophyllite	hydrothermal alteration	acidic alteration zones	high	
22	rhodochrosite	hydrothermal alteration	-	-	
23	talc	hydrothermal alteration	-	-	skarmization
24	vermiculite	hydrothermal alteration	-	-	granite

表 4-2 Fieldspec4 Hi-Res の仕様

Spectral Range	350-2500 nm
Spectral Resolution	3 nm @ 700 nm 8nm @ 1400/2100 nm
Sampling Interval	1.4 nm @ 350-1000 nm 1.1nm @ 1001-2500 nm
Scanning Time	100 milliseconds
Stray light specification	VNIR 0.02%, SWIR 1 & 2 0.01%
Wavelength reproducibility	0.1 nm
Wavelength accuracy	0.5 nm
Maximum radiance	VNIR 2X Solar, SWIR 10x Solar
Channels	2151
Detectors	VNIR detector (350-1000 nm): 512 element silicon array SWIR 1 detector (1001-1800 nm): Graded Index InGaAs Photodiode, Two-Stage, TE Cooled SWIR 2 detector (1801-2500 nm): Graded Index InGaAs Photodiode, Two-Stage, TE Cooled
Input	1.5 m fiber optic (25° field of view). Optional narrower field of view fiber optics available.
Noise Equivalent Radiance (NEdL)	VNIR 1.0×10^{-9} W/cm ² /nm/sr @ 700 nm SWIR 1 1.2×10^{-9} W/cm ² /nm/sr @ 1400nm SWIR 2 1.9×10^{-9} W/cm ² /nm/sr @ 2100 nm

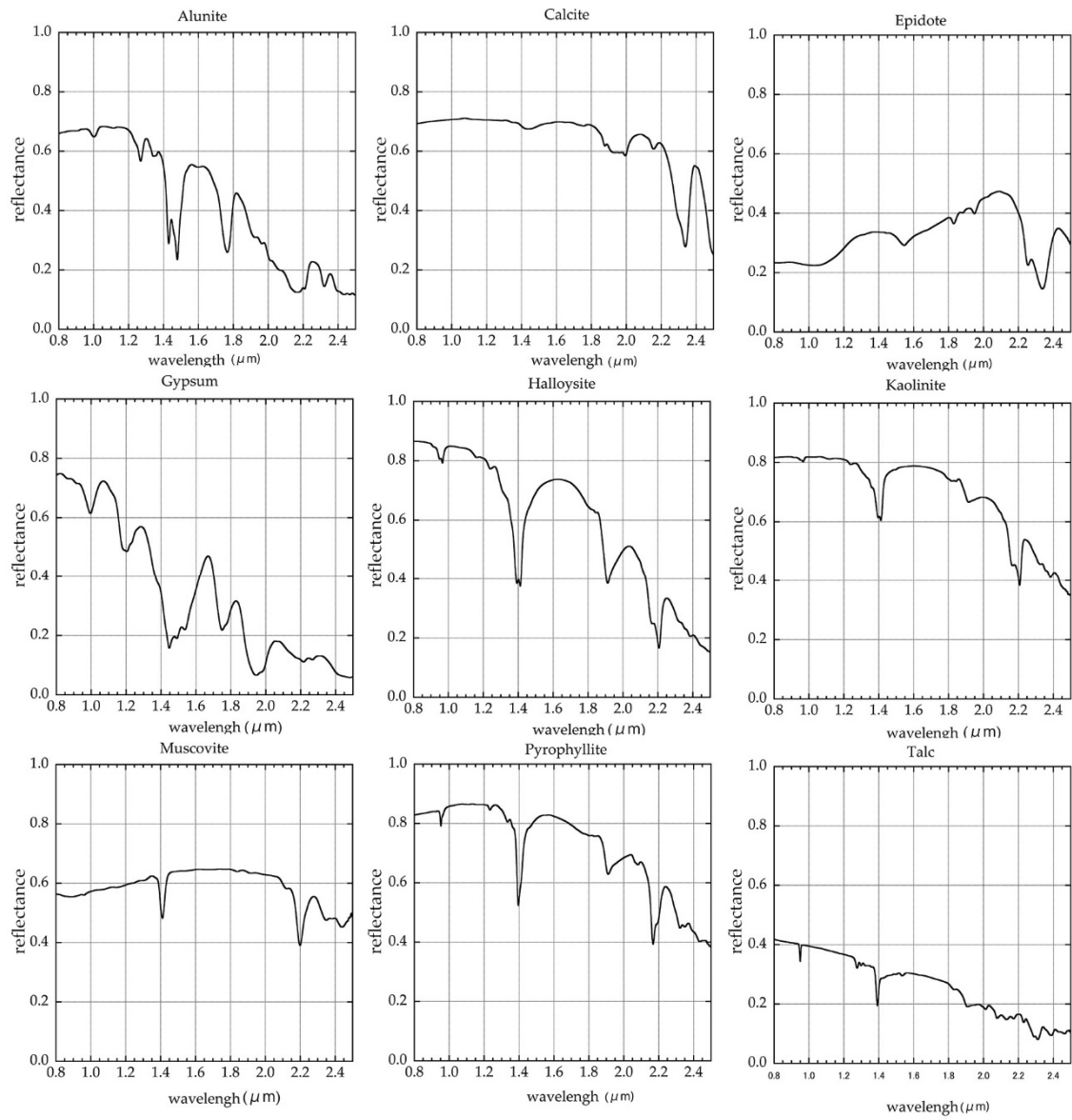


図 4-1 代表的な変質鉱物の分光反射スペクトル

収集データを次に示す。

- 浦井ほか (1989) の可視～短波長赤外における岩石・鉱物の反射スペクトル・カタログは GER 社 IRIS で測定されており約 80 個の粘土鉱物と約 30 個の炭酸塩岩の測定値がある。500-2,500 nm の分光反射スペクトル特性と化学分析値、X 線回折チャートが記述されている。このうち約 30 個の変質鉱物のデータを用いた。
- 坂野・古宇田 (2014) は国立研究開発法人 産業技術総合研究所 地質調査総合センター地質標本館所蔵鉱物の高波長分解能反射スペクトルを Field Spec Pro と FT-NIR で測定している。84 試料のサンプル測定しており、その中から約 20 個の変質鉱物のデータを用いた。
- Clark et al. (2007) の USGS の Spectral Library は多岐に渡り、X 線同定結果がないものもあるが、約 1,800 個余りの測定がある。このうち約 30 種類の変質鉱物のデータを用いた。
- METI Agency for Natural Resources and Energy (2001) の ASTER data usage manual にはオプトリサーチ社の MSR7000 で測定されてデータが付属していた。59 サンプルの岩石・鉱物の測定があったがその中から約 20 種類の変質鉱物のデータを用いた。

検討したが用いなかったデータ

- Baldridge(2009)の JPL ASTER Spectral Library には 120 個のデータにはほとんど変質鉱物が含まれていなかった。
- Brown 大学の RELAB (Reflectance Experiment Laboratory): A NASA Multiuser Spectroscopy Facility は波長分解能が 5 nm と大きく、変質鉱物が少なかった。
- RRUFF(Liu et al., 2017)には鉱物の赤外域のスペクトルが収録されているが FTIR 変換赤外分光光度計他で測定されていた。

4. 2 ハルコーシェント処理 (Hull quotient 処理(HQ 処理))

収集した全鉱物種のデータの全測定回数、及び全データ拡張処理に対して HQ 処理 (Green and Craig, 1985)を実施した。HQ 処理は吸収スペクトルの波長や強さを知るための波形データ処理方法のひとつで、一般には分光反射率の高い側の包絡線をいくつかの直線で近似し、これと測定値との商で表される。これは、分光反射スペクトルを明瞭にするとともに小さな吸収特徴を強調したい場合に有用である。この HQ 処理は半自動のスペクトル同定ソフトではしばしば行われている処理であり Continuum Removal 処理(Clark et al., 2003)とも呼ばれる。弱点としてはノイズがあった場合にそれを強調してしまうことが上げられる。図 4-2 にハロイサイトのオリジナル分光反射スペクトルと HQ 処理済みの分光反射スペクトルを示す。1,400 nm 付近のダブルット吸収、1,900 nm の鋭い吸収及び 2,200 nm のダブルット吸収などが特徴であるが HQ 処理により各吸収スペクトルが強調され、小さな吸収も強調されている。

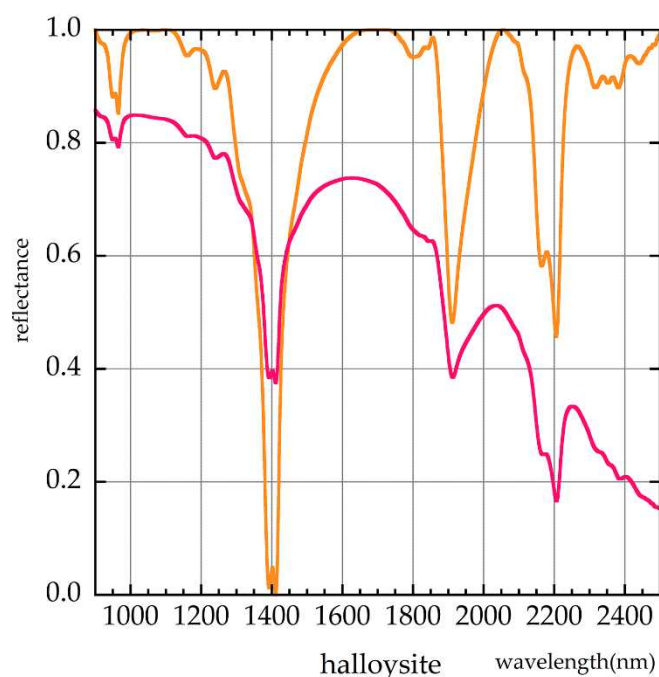


図 4-2 HQ 処理の有無(900-2,500 nm)
赤色：オリジナルの分光反射スペクトル(halloysite)
オレンジ色：HQ 処理を施した分光反射スペクトル(halloysite)

4. 3 データ拡張処理 (Augmentation (AUG 処理))

ニューラルネットワークに学習させるデータ数を増やすために AUG 処理を実施した。収集されたスペクトルデータ 1 サンプルにつき 50 回の AUG 処理を実施した。通常 1 鉱物につき 30 回の測定を行っているので、全 24 鉱物に対して 1set、36,000 個のデータで HQ 処理なし、HQ 処理済みで計 72,000 個の学習用データセットとなった。

AUG 処理は下記のように反射率方向、波長方向に乱数を与えてシフトさせた。各チャンネル方向も乱数を与えてシフトさせた。AUG 処理はこの 3 種類を同時に行っている。通常の画像に対する水平方向の反転、鉛直方向の反転等は不可であり、シフト量も通常考えられる範囲内に抑えなくてはならなかった。

- a) 分光計の波長方向のばらつきを考慮すること
- b) サンプルとの距離、光源のばらつきを考慮すること
- c) 分光計の各チャンネルのばらつきを考慮すること

HQ 処理なしデータセットでは浦井ほか(1989)の GER 社のスペクトルメーターのノイズを参考にノイズを付加させ、波長方向は実際のデバイスのズレ(数 nm 程度)を参考にして AUG データを作成した。図 4-3 データ拡張処理による HQ 処理のアルーナイトの各スペクトルを示す。ほぼ同じ吸収位置にピークを示すがノイズの乗り方が異なりデータ拡張されたものとなっており学習サンプルとして用いることができた。

データ拡張処理のパラメータの乱数は確率変数の定義域が切断正規分布を用いている。これは通常の正規分布では確率的に大きな値が入るためである。

・乱数の発生;

切断正規分布に従う乱数を生成

・波長方向

C0： 最小値 -1.5 最大値 1.0 平均値 0.0 標準偏差 1.5

C1： 最小値 $1.0 - 1.e-4$ 最大値 $1.0 + 1.e-4$ 平均値 1.0 標準偏差 $5.e-3$

C2： 最小値 $-5.e-6$ 最大値 $5.e-6$ 平均値 0.0 標準偏差 $5.e-7$

C3： 最小値 $-5.e-9$ 最大値 $5.e-9$ 平均値 0.0 標準偏差 $-5.e-10$

$$WL' = C0 + C1*WL + C2*WL*WL + C3*WL*WL*WL$$

波長方向に最大 5nm 程度ズレを発生させた。

・ 全体の感度変化

a： 最小値 0.95 最大値 1.05 平均値 1.0 標準偏差 0.025

b： 最小値 -0.05 最大値 0.05 平均値 0.0 標準偏差 0.025

$$Val' = a * Val + b$$

・ ノイズ成分

LN： 最小値 -0.005 最大値 0.005 平均値 0.0 標準偏差 0.0015

$$Val' = Val + \sum LNi$$

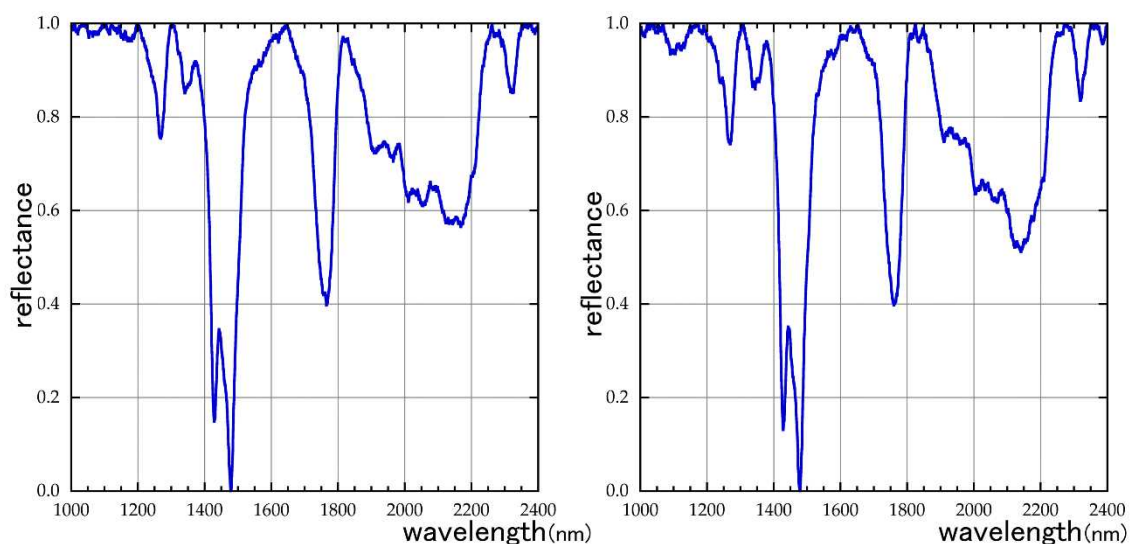


図 4-3 データ拡張処理による HQ 処理の alunite のスペクトル(オリジナルの異なる 2 種類の例)

5 手法

5. 1 深層学習 (Deep Learning)

深層学習(Goodfellow et al.,2016)の歴史は古く、パーセプトロンが Rosenblatt によって考案されたのは 1958 年であるが、線形非分離の問題が解けないなど多くの欠点を有し、研究が長く続くことはなかった。1980 年代には Rumelhart et al.(1986)により誤差逆伝播法が考案されたが非効率的なメカニズムやマシンのスペックのために再びブームは沈静化していった。2012 年には物体の認識率を競う ILSVRC 国際コンペティションにおいてトロント大学チームが深層学習を用いて劇的な進歩を示し(Krizhevsky et al., 2012)、その後の国際コンペティションである Common Objects in Context(Lin, et al., 2014)等でも深層学習の流れが主流となっている。こうして深層学習の現在の潮流が作られていった。

深層学習は人間の脳回路をモデルにした人工ニューロンを多層化し、入力データに基づいて各ニューロンの重み付けを変化させながら学習させ、解析を繰り返す毎に学習を重ねて賢くなっていくのが特徴である。

しかしながら効果的な学習を行うためには膨大なデータが必要で、それらに対して多様な深層学習を選択し適用する必要がある、強力なコンピューティングのパワーと深い経験が要求される。深層学習にチャレンジしてみたいと思っても、容易には踏み出せない。

深層学習のプロセスは、ビッグデータをもとに学習モデルを生成するトレーニングと、新しく入力されたデータに学習済みのニューラルネットワークを適用して分類・分析を行うスコアリングの 2 つのフェーズに大きく分けられる。

このうち特に処理時間が問題となるのがトレーニングである。データが大規模化するに従い学習モデルの生成時間（前処理を含む）が長期化し、数日から数週間を要する場合がある。すなわちディープラーニングに本格的に取り組み、研究で限られた時間内に成果を上げるためには、このトレーニングに要する時間を短縮できるかが重要な鍵とな

っている。本研究でも CNN2d はハイエンド GPU 搭載マシンで 1 日程度学習時間が必要であった。

従来の機械学習においても、これまで Scale Invariant Feature Transform (SIFT) (Lowe D.G., 2004)、Speed Up Robust Features (SURF) (Bay H. et al., 2006)、Histograms of Oriented Gradients (HOG) (Dalal & Triggs, 2005) などさまざまな特徴抽出アルゴリズムが提案されてきており、例えば SIFT は特徴点周りの輝度値の変化を調べて特徴量を生成するものである。これらの特徴抽出による精度やスピードと比較しても、データから特徴抽出に人手を介さず自動的に抽出させる手法である深層学習が注目された。

深層学習は優れたアルゴリズムであり、今までできなかった分野にブレークスルーをもたらしたことは間違いなく、主に自動運転、画像解析、言語翻訳など幅広い範囲で多くの成果をあげており、今後も発展していくと考えられる。現在、深層学習がこれほど注目されているのはハードウェアの進歩と、大量のデータへのアクセスが可能になってきたこと、ニューラルネットワークに関する技術的な進歩などにより実用化が進んできたことによる。

深層学習の手法は多岐にわたるが本研究では一次元に用いることができ二次元の形態認識にも実績のある MLP と画像認識に定評があり図としての分光反射スペクトルパターンを学習させることができる CNN を用いた。

5. 2 多層パーセプトロン (Multi-Layer Perceptron (MLP))

MLP は 4 層以上のニューラルネットワークを有するものをいうが層構造は入力層と出力層の間には何層もの全結合層とドロップアウト層を含むものである。初期のものは誤り訂正学習が入力データの教師信号と出力との関係から結合重みを更新するが層を超えて誤りが訂正できないので多層の場合は学習ができなかった。したがって、初期の MLP は入力層と中間層の間の結合重みは乱数で決めて、中間層と出力層間の重みのみ誤り訂正学習で学習していた。そのためうまく機能しないことが多かったため、解決策として考案されたのが誤差逆伝播法である。この逆伝搬を行う上で重要なのが出力と正解との差である損失関数であり、多分類問題では交差エントロピー誤差を、二値分類は平均二乗誤差を用いる。交差エントロピーは次の式で表される。N は学習データのサン

プル数、 K は出力素子数、 y はニューラルネットワークモデルの出力である。 d_{nk} は教師データである。

$$E = -\frac{1}{N} \sum_{n=1}^N \sum_{k=1}^K d_{nk} \log y_{nk}$$

平均二乗誤差は下記のようなになる。 x_i はネットワークの i 番目の出力、 y_i は i 番目の出力に対する目標値(教師データ)でこの値を最小になるように重みが更新されていく。

$$f(x) = \frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2$$

$$\frac{\partial f(x)}{\partial x} = \frac{2}{n} (y_i - x_i)$$

MLP で学習させるときに各パラメータの更新量を算出する必要がある。更新するパラメータは各層の重みとバイアスである。MLP では学習誤差を用いて各パラメータの更新量を求める。この各パラメータの更新量を求める方法として誤差逆伝播法を用いる。次に、誤差逆伝播法により求めた更新量を勾配降下最適化法を使いパラメータを更新する。MLP の弱点は単純な層の繰り返して学習時間は比較的短い、パラメータ数が多く調整が容易ではないことが挙げられる。

損失関数とは学習時に入力データに対する出力データとモデルの出力のの違いの度合いを表して評価する関数のことである。

5.3 畳み込みニューラル ネットワーク (Convolution Neural Network (CNN))

CNN は畳み込み層、プーリング層そしてドロップアウト層を積み重ねたフィードフォワード型のニューラルネットワークである。原型は Fukushima(1980)のネオコグニトロンにあるとされている。ネオコグニトロンは学習方法に誤差逆伝播法を用いていないのが CNN と異なるのみとされている。CNN は画像認識に強いとされており、最近では学習データの大規模化が行われているが、なぜ高い性能を示すかは証明されていない。

CNN は出力層では全結合層を有し、最終層には softmax 関数を活性化関数として用い分類させることができる。

通常の 2 次元 CNN(CNN2d) は畳み込み層とプーリング層を多段に設定して最後に全結合を行い出力するものである。実際には過学習が起きるためにドロップアウト層を多用した。活性化関数にはシグモイド関数に比べて軽量で使い易い ReLU 関数を用いている。通常、畳み込みで得られる画像はもとの画像より小さくなり、この後のプーリング層で処理をうけ、さらに特徴量は圧縮されていく。フィルタは複数設定するのでチャンネル層は増加していく。

1 次元 CNN(CNN1d)は縦 1×横 3 のサイズのフィルタリングを分光反射率に直接畳み込み演算をしていく手法である。

5. 4 研究の流れ

本研究の流れを図 5-1 に示す。

HQ 処理を施すかどうかという前処理の相違はあるが深層学習のステップは概ね下記通りである。

- ①変質鉱物の吸収ピーク位置の検討
- ②分光反射率データの収集
- ③波長分解能を 1nm おきデータに変換
- ④波長域を 900-2,500nm に統一
- ⑤オーギュメンテーション処理の実施
- ⑥各ネットワーク手法・構造の決定、パラメータの検討
- ⑦HQ 処理
- ⑧各ネットワークの学習とテスト
- ⑨交差検証法での各ネットワークの評価

学習とはモデルの持つ特徴を捉えることであるが、言い換えると入力データを用いてモデルが同じデータを再現できるように訓練することである。初期の段階では誤差が生じるのでそれを評価する関数(すなわち、損失関数)をいかに小さくするかが主な作業と

なる。一般に深層学習は解析的には答えを得ることができないので誤差逆伝播法を用いて解を求めていく。

前述のようにネットワーク構造は MLP と CNN を用いた。ともに近年までに、いくつかのブレイクスルーを経て現在も高い評価を得ている手法である。MLP に関しては層構造を変化させて MLP_S(全結合層 4 層, ドロップアウト層 3 層)、MLP_M(全結合層 9 層, ドロップアウト層 8 層)、MLP_L(全結合層 13 層, ドロップアウト層 12 層)の層数の異なる 3 タイプを評価した。これはニューラルネットワークの層数が変質鉱物同定にどれだけ影響を与えるかを検討するためである。S は small を M は medium を L は large をそれぞれ層構造の厚さを示す。

CNN に関しては 1 次元 CNN(CNN1d)と 2 次元 CNN(CNN2d、CNN2d_Hi)のものを評価した。1 次元は 1×3 のフィルタを分光反射率に畳み込みを行ったものである。反射率は 1 nm ごとに並んでおり 1 次元のデータと見なせる。2 次元 CNN は画像識別などに用いられているものでありスペクトルチャートを作成して深層学習を実施した。分解能は二段階(430×286 、 1773×229)に変化させて評価した。

研究は HQ 処理なしのものと HQ 処理を施したものに区分しそれぞれ学習させて変質鉱物同定を実施した。f 値と正答率により HQ 処理の能力を評価した。各処理に対して正答率・損失関数とエポック数のグラフを示し、過学習の有無を確認した。一部は混合行列を示した。損失関数はどれだけ正しい値から離れているかの指標である。エポック数とはその学習を繰り返した回数のことである。

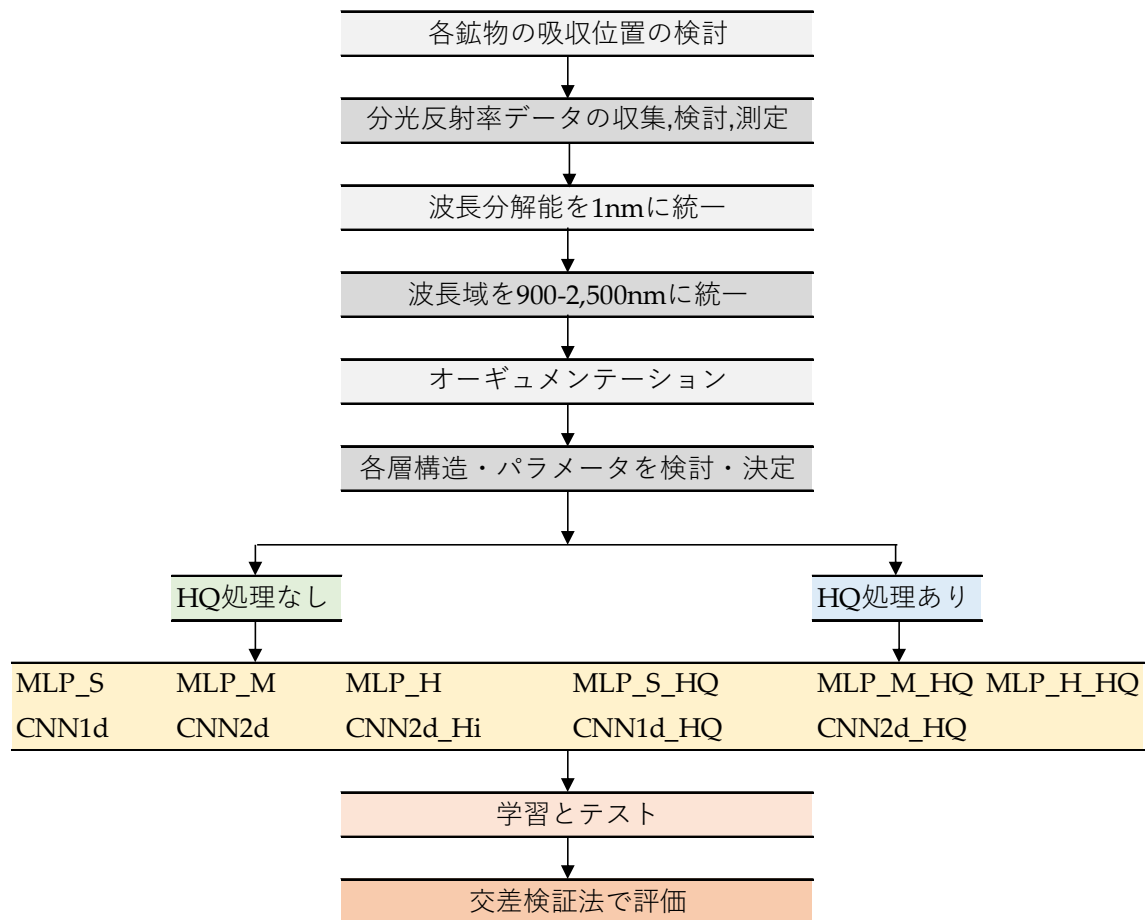


図 5-1 Flow of study

用いたソフトウェアとハードウェアの主な諸元を示す。

- ソフトウェア:
Ubuntu 16.04, Windows10
NVIDIA Docker
Python 3.6.3, Keras 2.1.0, Tensorflow 1.4.0
- ハードウェア:
CPU: Intel Core i7-6700 3.4GHz
Memory: 64GB
NVIDIA GeForce GTX 1080 Ti 11GB

5. 5 モデルの評価法

深層学習のモデルの評価は主なものとしてホールドアウト法と交差検証法がある。今回は交差検証を各ネットワークにそれぞれ適用した。表 5-1 に混合行列を図 5-2 に深層学習の評価の図を示す。

ホールドアウト法はモデルを作る学習データとモデルを評価するテストデータに分類して評価する。これは学習データでテストしてもモデルの汎化能力は評価できないためである。データセットを学習データとテストデータに分割する。学習データを使ってモデルを開発し、テストデータで f 値・正答率等を評価する。

交差検証はデータ群を k 等分（今回は 5 等分）して、そのうち 1 つをテスト事例として用い、 k 回を学習と変質鉱物同定を行うものである。得られた k 回の結果を平均して 1 つの推定を得る。データ量が少ない場合には、テストデータの選び方によって、推定精度に大きな誤差が生じる可能性が高くなる。そのため、交差検証法が有効である。本研究では交差検証における学習時には学習データに対応した AUG データを含めてい。評価時にはテストデータとテストデータに対応した AUG データを含めた精度評価もそれぞれ実施した。交差検証法とホールドアウト法を比べた場合、交差検証の方がより評価結果の信頼性が高くなる。

表 5-1 混合行列

	Actual	
	True Positive	False Positive
Predicted	False Negative	True Negative

適合率 (Precision)、再現率 (Recall) は「各クラス」に対して計算される。以下、2 値分類の例で説明をする。適合率、再現率を計算するためには、以下の 4 項目 (TP, FP, TN, FN) を計算すればよい。

真陽性：True Positive (TP) = "正しく" calcite と分類

偽陽性：False Positive (FP) = "誤って" calcite と分類

真陰性：True Negative (TN) = "正しく" calcite ではないと分類

偽陰性：False Negative (FN) = "誤って" calcite ではないと分類

このとき、正答率、適合率、再現率を以下のように定義する。

$$\text{正答率} = \frac{\text{真陽性} + \text{真陰性}}{\text{真陽性} + \text{偽陽性} + \text{真陰性} + \text{偽陰性}}$$

変質鉱物同定結果の予想された結果が、実際にそうであるものの割合

$$\text{適合率} = \frac{\text{真陽性}}{\text{真陽性} + \text{偽陽性}} \quad (\text{Precision})$$

変質鉱物同定で予測が正で実際に正であったものの割合

$$\text{再現率} = \frac{\text{真陽性}}{\text{真陽性} + \text{偽陰性}} \quad (\text{Recall})$$

変質鉱物同定で実際に正であるものの中で正だと予測できた割合

適合率と再現率のバランスを定量化したものがf値 (f score)である。分類精度は、再現率、正解率、およびこれらの調和平均であるf値を算出することで評価する。

$$f = 2 \times \frac{\text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}}$$

複数のアルゴリズムのうち、どの手法を用いるか決定する際、正答率を参考にしながらf値が大きいものを選ぶことで偏りのあるデータにも対応できると考えられる。

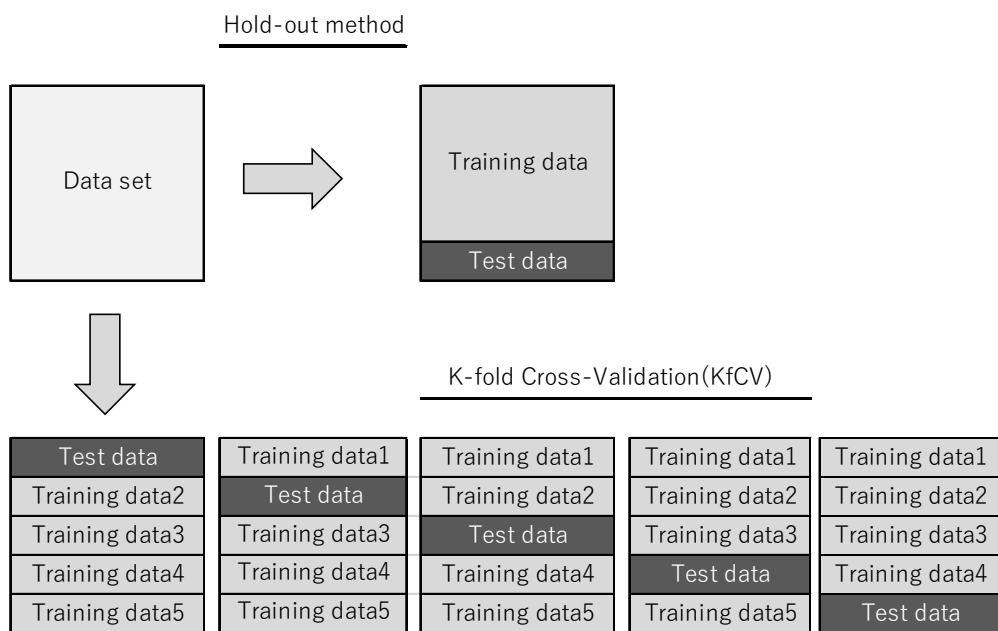


図 5-2 深層学習の評価法(ホールドアウト法と交差検証法)

6 結果

6. 1 では HQ 処理を行わず、分光反射率データをそのまま用いたデータセットで変質鉱物の学習・同定を行い、これに対して 6. 2 では HQ 処理を施したデータセットで変質鉱物の学習・同定を行った。

6. 1 HQ 処理なしの深層学習の結果

この 6. 1 ではすべてのデータに対して HQ 処理を行わず、AUG 処理 (4. 3 参照) を行って深層学習を実施した。層構造の厚さの異なる MLP_S、MLP_M、および MLP_L が変質鉱物の同定に対してどの程度の f 値・正答率を示すのか、またどの程度の層構造が変質鉱物同定に適しているかの検討を行った。また、合わせて CNN1 d と CNN2 d も評価を実施した。各変質鉱物の測定回数は 30 回であり、解析結果は交差検証法で評価した。

表 6-1 に MLP_S の層構造を示す。

表 6-1 MLP_S の層構造

Layer Type	Output	Parameter#
dense_1	(None, 512)	820224
dropout_1	(None, 512)	0
dense_2	(None, 512)	262656
dropout_2	(None, 512)	0
dense_3	(None, 256)	131328
dropout_3	(None, 256)	0
dense_4	(None, 24)	6168

Dense 層は全結合層を表す。ニューラルネットワークは表現力が高いモデルなので過学習しやすい。そのためより精緻な過学習制御が必要となる。ドロップアウト層は過学習を防ぐ強力な学習方法であり、近年の大規模なニューラルネットワークの成功の主要因の一つとされている。ドロップアウト層は毎回ランダムに一定割合のノードを消した状態で学習を行う。本研究ではドロップアウトのパラメータは 0.5 を用いた(神畠ほか,

2015)。なお、中間層の活性化関数は ReLU 関数を用いている。最終段では多分類問題であるので softmax 関数を用いている。

図 6-1 に MLP_S の学習時とテスト時の正答率(acc)と損失関数(loss)のグラフを示す。横軸はエポック数である。

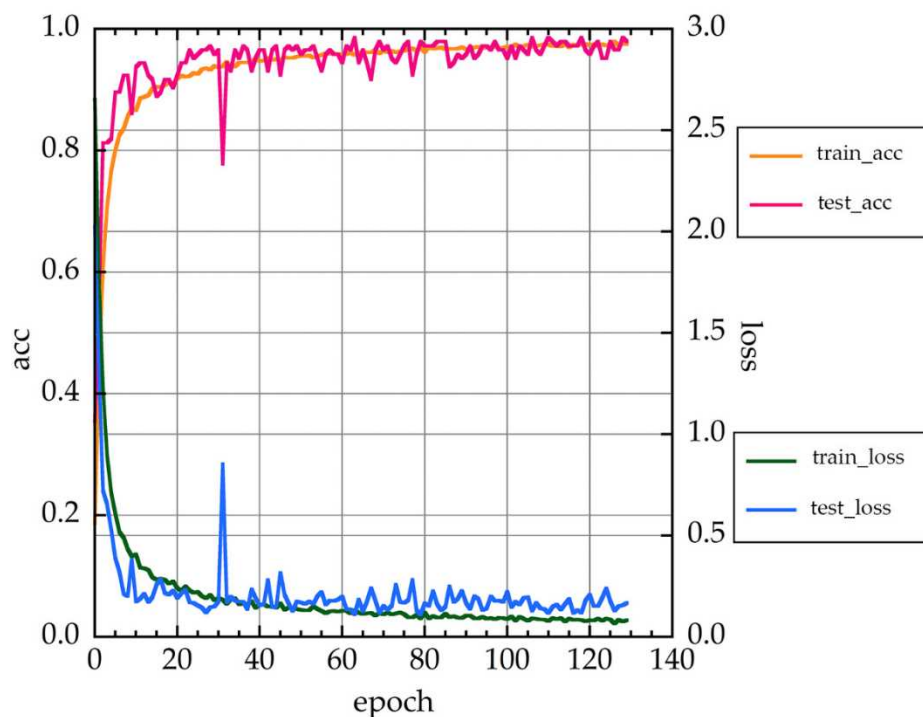


図 6-1 MLP_S の学習時とテスト時の正答率と損失関数
(train_acc;学習時の正答率、test_acc;テスト時の正答率、
train_loss;学習時の損失関数、test_loss;テスト時の損失関数)

図 6-1 ではテスト時の正答率とテスト時の損失関数はやや値にばらつきが認められる。特にエポックの 30 付近の誤差関数は大きく乱れている。その後は過学習の傾向も明瞭には認められないが学習時の損失関数とテスト時の損失関数はエポックが増すにつれてわずかに離れて行っている。最良正答率は 98.6%を示し（表 6-6）、交差検証の f 値は 96.1%を示す（表 6-7）。AUG データを含む all_f 値は 94.7%を示す（表 6-7）。また損失関数は 0.5 を下回るまでに 5 エポック程度が必要であった（図 6-1）。変質鉱物自動同定のためのモデル化は検討を要するが大きな損失関数の乱れは 30 エポックのみで後半は安定している。60 エポック付近で可能性はあると考えられる。

表 6-2 に MLP_S に比較して層構造を厚くした MLP_M の層構造を示す。図 6-2 に MLP_M の学習時とテスト時の正答率と損失関数のグラフを、表 6-3 にさらに層構造を厚くした MLP_L の層構造を示す。図 6-3 に MLP_L の学習時とテスト時の正答率と損失関数のグラフをそれぞれ示す。層構造は MLP_S よりも厚くなるわけであるが、厚くなるに従いテスト時の損失関数は乱れ、上下に振れ始めている。MLP_M と MLP_L の最良正答率はそれぞれ 98.6%、99.3%を示し(表 6-6)、交差検証の f 値は 96.4%と 96.2%をそれぞれ示す(表 6-7)。また、MLP_M と MLP_L の AUG データを含む all_f 値は 95.3%と 95.2%をそれぞれ示す(表 6-7)。これらの結果からは MLP_M と MLP_L はいずれも高い f 値等を示すが損失関数が乱れ、テスト時の正答率も安定しておらず、わずかに過学習し、ノイズが発生していることから変質鉱物自動同定のためのモデル化は難しいと考えられる。

表 6-2 MLP_M の層構造

Layer Type	Output	Parameter#
dense_1	(None, 512)	820224
dropout_1	(None, 512)	0
dense_2	(None, 512)	262656
dropout_2	(None, 512)	0
dense_3	(None, 256)	131328
dropout_3	(None, 256)	0
dense_4	(None, 256)	65792
dropout_4	(None, 256)	0
dense_5	(None, 128)	32896
dropout_5	(None, 128)	0
dense_6	(None, 128)	16512
dropout_6	(None, 128)	0
dense_7	(None, 64)	8256
dropout_7	(None, 64)	0
dense_8	(None, 64)	4160
dropout_8	(None, 64)	0
dense_9	(None, 24)	1560

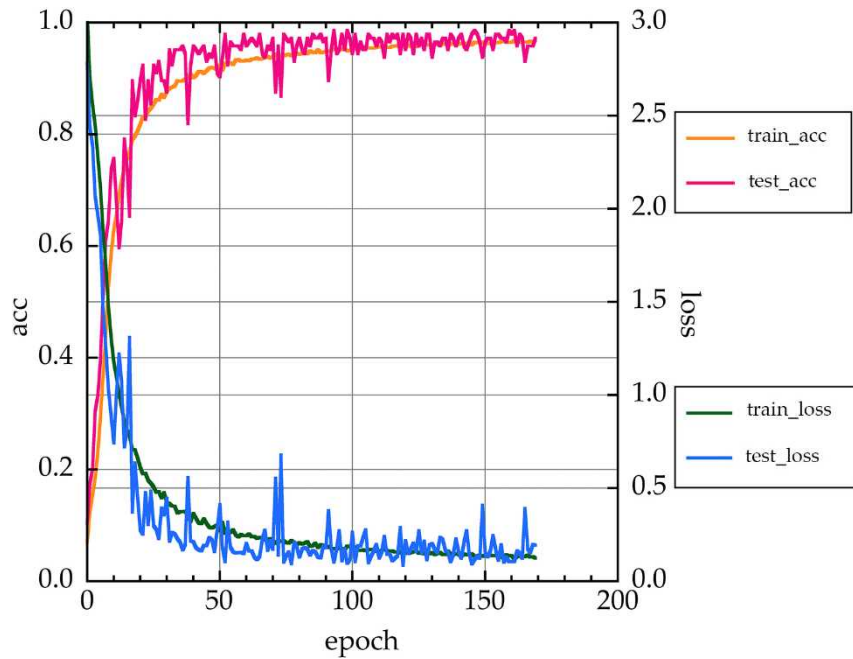


図 6-2 MLP_M の学習時とテスト時の正答率と損失関数

表 6-3 MLP_L の層構造

Layer Type	Output	Parameter#
dense_1	(None, 512)	820224
dropout_1	(None, 512)	0
dense_2	(None, 512)	262656
dropout_2	(None, 512)	0
dense_3	(None, 512)	262656
dropout_3	(None, 256)	0
dense_4	(None, 256)	131328
dropout_4	(None, 256)	0
dense_5	(None, 256)	65792
dropout_5	(None, 256)	0
dense_6	(None, 256)	65792
dropout_6	(None, 256)	0
dense_7	(None, 128)	32896
dropout_7	(None, 128)	0
dense_8	(None, 128)	16512
dropout_8	(None, 128)	0
dense_9	(None, 128)	16512
dropout_9	(None, 128)	0
dense_10	(None, 64)	8256
dropout_10	(None, 64)	0
dense_11	(None, 64)	4160
dropout_11	(None, 64)	0
dense_12	(None, 64)	4160
dropout_12	(None, 64)	0
dense_13	(None, 24)	1560

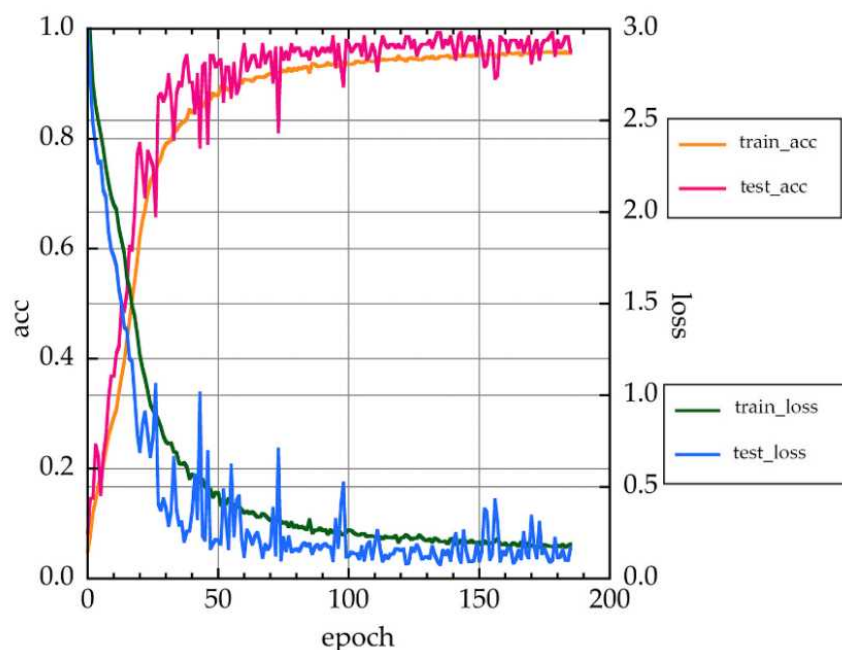


図 6-3 MLP_L の学習時とテスト時の正答率と損失関数

表 6-4 に CNN1d の層構造、図 6-4 に CNN1d の学習時とテスト時の正答率と損失関数のグラフを示す。テスト時の正答率と損失関数のグラフは乱れており、損失関数は過学習の傾向を示し、エポック 50 以降は発散の傾向をしている。最良正答率は 95.1% と低く（表 6-6）、交差検証の f 値も 89.8% と低い（表 6-7）。AUG データを含む all_f 値も 87.3% と低い（表 6-7）。またテスト時の正答率も安定していない。図 6-4 の結果を見る限り、HQ 処理なしのオリジナルデータの分光反射率は CNN1d の変質鉱物同定解析に適していないと考えられる。1×3 の畳み込みが分光反射スペクトルの特徴を抽出できていないと考えられ、変質鉱物自動同定のための CNN1d のモデル化は困難である。

表 6-4 CNN1d の層構造

Layer Type	Output	Parameter#
conv2d_1	(None, 1, 1599, 32)	128
max_pooling2d_1	(None, 1, 799, 32)	0
dropout_1	(None, 1, 799, 32)	0
conv2d_2	(None, 1, 797, 64)	6208
max_pooling2d_2	(None, 1, 398, 64)	0
dropout_2	(None, 1, 398, 64)	0
conv2d_3	(None, 1, 396, 128)	24704
max_pooling2d_3	(None, 1, 198, 128)	0
dropout_3	(None, 1, 198, 128)	0
flatten_1	(None, 25344)	0
dense_1	(None, 64)	1622080
dropout_4	(None, 64)	0
dense_2	(None, 32)	2080
dropout_5	(None, 32)	0
dense_3	(None, 24)	792

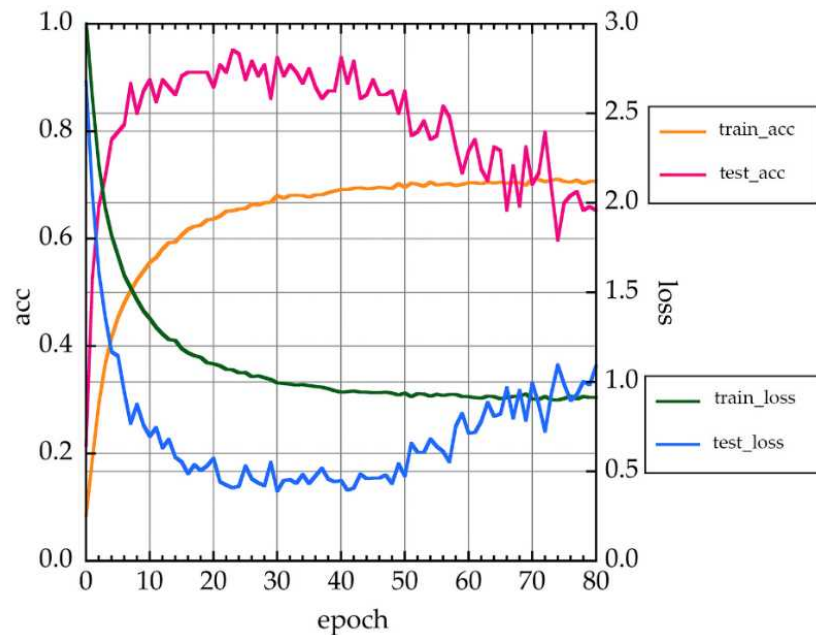


図 6-4 CNN1d の学習時とテスト時の正答率と損失関数

表 6-5 に CNN2d の層構造を示す。図 6-5 に CNN2d の学習時とテスト時の正答率と損失関数のグラフを示す。

表 6-5 CNN2d の層構造

Layer Type	Output	Parameter#
conv2d_1	(None, 286, 430, 32)	896
max_pooling2d_1	(None, 95,143,32)	0
dropout_1	(None, 95,143, 32)	0
conv2d_2	(None, 93 ,141, 64)	18496
max_pooling2d_2	(None, 31, 47, 64)	0
dropout_2	(None, 31, 47, 64)	0
conv2d_3	(None, 29, 45, 128)	73856
max_pooling2d_3	(None, 9, 15, 128)	0
dropout_3	(None, 9, 15, 128)	0
conv2d_4	(None, 7, 13, 128)	147584
max_pooling2d_4	(None, 2, 4, 128)	0
dropout_4	(None, 2, 4, 128)	0
flatten_1	(None, 1024)	0
dense_1	(None, 64)	65600
dropout_5	(None, 64)	0
dense_2	(None, 32)	2080
dropout_6	(None, 32)	0
dense_3	(None, 24)	792

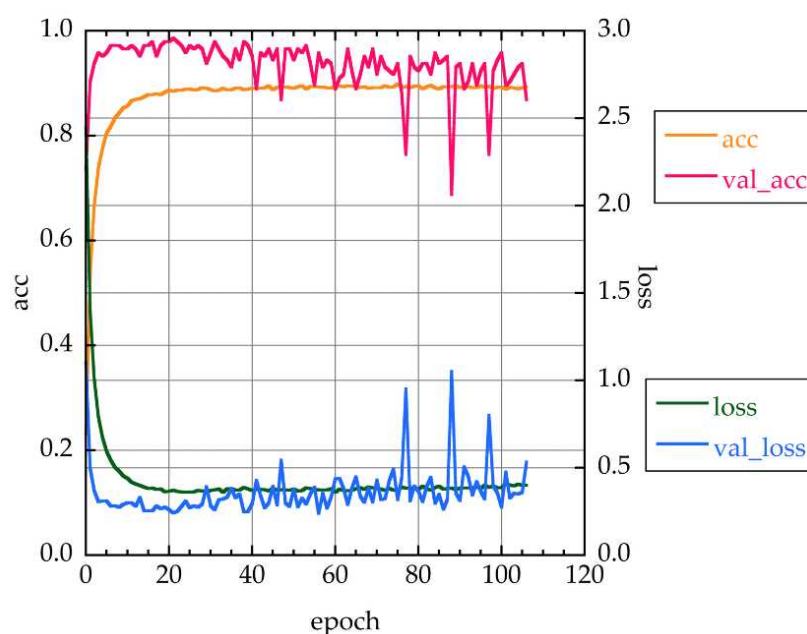


図 6-5 CNN2d の学習時とテスト時の正答率と損失関数

CNN2d のテスト時の正答率と損失関数のグラフはやや乱れている。損失関数はエポック 40 以降で乱れてやや過学習の傾向を示しめす。(図 6-5)。最良正答率は 98.6%

と高く（表 6-6）、交差検証の f 値は 96.7%を示す（表 6-7）。AUG データを含む all_f 値は 95.5%を示す(表 6-7)。テスト時の正答率もエポックが進むほど安定していない。損失関数が 0.5 を下回るまでにエポック数は 5 回程度である。CNN2d の特徴はテスト時の正答率が学習時の正答率を上回っていることである。モデル化を行う場合は前半部分の高い正答率の部分で行うことになるが学習の正答率が 90%程度とあまり進んでおらず、変質鉱物自動同定のための CNN2d によるモデル化は困難である。

表 6-6 ネットワークごとの最良正答率%

number of measurements	mineral species	epoch	type of layer	test_accuracy
30 times	24	63	MLP_S	98.6
		93	MLP_M	98.6
		134	MLP_L	99.3
		23	CNN1d	95.1
		21	CNN2d	98.6
		4	CNN2d_Hi	98.6

表 6-7 に交差検証の値を示す。複数のネットワークのうちどれを用いるか決めるとき、f 値が大きいものを選べば、偏りのあるデータにも対応できると考えられる。f 値はテストデータのみの値である。all_f 値は AUG 処理データも含んだ値である。表 6-7 のネットワーク毎の交差検証の値をみると CNN1d 以外は f 値で 96～97%台でほぼ高い同じ値を示しており大きな差は認められない。しかし、CNN は損失関数にスパイク状のノイズが載って損失関数の悪化が認められる。

表 6-7 ネットワーク毎の交差検証の平均結果%(HQ 処理なし)

nw	acc	pr	rc	f	all_acc	all_pr	all_rc	all_f
MLP_S	96.1	96.7	96.1	96.1	94.7	95.2	94.7	94.7
MLP_M	96.4	96.9	96.4	96.4	95.3	95.6	95.3	95.3
MLP_L	96.2	96.6	96.3	96.2	95.2	95.6	95.2	95.2
CNN1d	89.7	91.9	89.7	89.8	87.1	89.2	87.1	87.3
CNN2d	96.7	97.2	96.7	96.7	95.5	95.8	95.5	95.5
CNN2d_Hi	97.1	97.6	97.1	97.1	96.4	96.7	96.4	96.4

nw : network, acc : accuracy, pr : precision, re : recall

MLP では層構造が薄いほど損失関数が 0.5 を切るのが早い傾向がある。CNN2d が比較的小さなエポック数となっている。CNN1d は過学習が顕著であり誤差関数の発散の傾向を示す。

表 6-7 に CNN2d と CNN2d_Hi の交差検証の各値を示す。分解能の高い CNN2d_Hi が 0.4~0.9% 上昇しており f 値も 97.1 と高いが図 6-6 に示すように学習時の正答率は低く、損失関数が大きく乱れており過学習の傾向が出ている。変質鉱物自動同定のための CNN2d_Hi のモデル化は必要な正答率が得られない可能性があり、困難である。

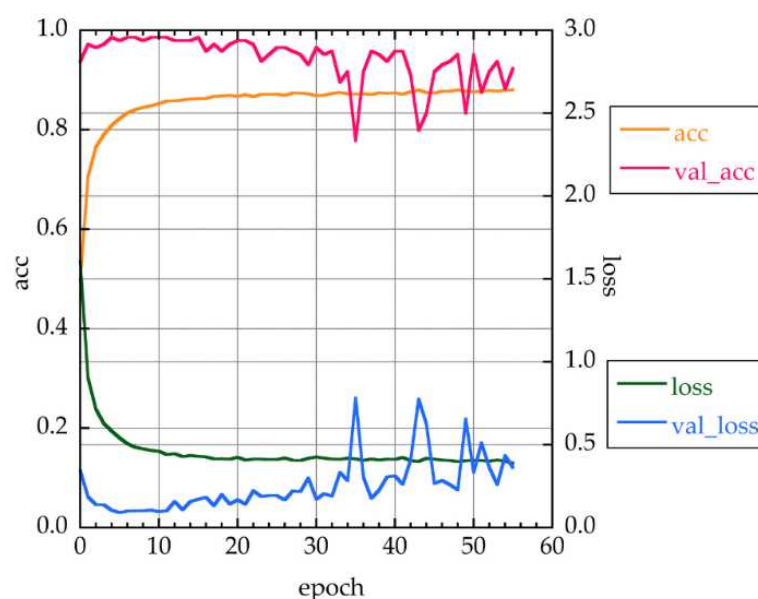


図 6-6 CNN2d_Hi の学習時とテスト時の正答率と損失関数

HQ 処理なしの各ネットワークではモデル化ができる可能性があるのは MLP_S のみで他は過学習が認められたり、大きなノイズが現れたりして変質鉱物自動同定のためのモデル化は困難と考えられた。

6. 2 HQ 処理済みの深層学習の結果

4. 3 で述べたように収集した変質鉱物 24 種類のデータを 30 回測定したものに AUG 処理を施し、これらに対して HQ 処理を実施して学習と同定の各処理を行った。HQ 処理を施すことにより分光反射スペクトルが明瞭になり同定結果が向上することを期待

した。各ネットワークとも層構造・条件は HQ 処理なしと同じである。波長域も 900～2,500nm で HQ 処理なしと同じ波長域である。

図 6-7 に HQ 処理を施した MLP_S_HQ の学習時とテスト時の正答率と損失関数、表 6-8 に MLP_S_HQ の混合行列をそれぞれ示す。最良正答率は 98.6%(表 6-9)である。また交差検証の f 値は 97.5%で、AUG 処理データも含んだ all_f 値は 96.4%を示した(表 6-10)。これらの f 値及び all_f 値は 3 層の MLP の中では最大値のものである。

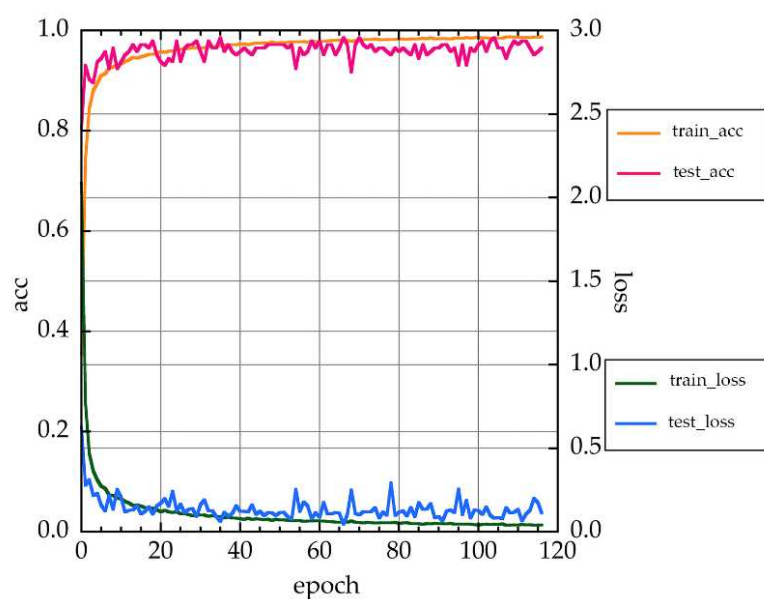


図 6-7 MLP_S_HQ の学習時とテスト時の正答率と損失関数

表 6-8 MLP_S_HQ の混合行列

MLP_S_HQ	Alu	Anh	Cal	Chl	Dic	Dol	Epi	Gib	Goe	Gyp	Hal	Hec	Ill	Jar	Kao	Mon	Mus	Nat	Non	Phl	Pyr	Rho	Tal	Ver
Alunite	300																							
Anhydrite		293				1			1					1				3		1				
Calcite		8	289	1			1															1		
Chlorite				298				1												1				
Dickite					296										4									
Dolomite						300																		
Epidote							300																	
Gibbsite		1		1		5	1	261	25		1						4	1						
Goethite	1	2		1	1	3	3	3	270					2			2	1		8	2	1		
Gypsum										300														
Halloysite								2			298													
Hectorite												300												
Illite		1											297		2									
Jarosite		2							15					283										
Kaolinite													3		297									
Montmorillonite																300								
Muscovite													1				299							
Natrolite																		300						
Nontronite																			300					
Phlogopite		1		1																298				
Pyrophyllite								11			3										285		1	
Rhodochrosite																						300		
Talc																							300	
Vermiculite		1																						299

図 6-7 はテスト時の正答率と損失関数の動きが比較的落ち着いている。損失関数の乱れも 50 エポック以降にわずかに見られるのみである。これらのことから変質鉱物自動同定のためのモデル化は可能と考えられる。また表 6-8 の混合行列では jarosite が goethite に、pyrophyllite が gibbsite にそれぞれ誤同定された部分がある。図 6-8 に類似したスペクトル例(ヘクトライトとモンモリロナイト)を示すがこれらは同定率 100% である。

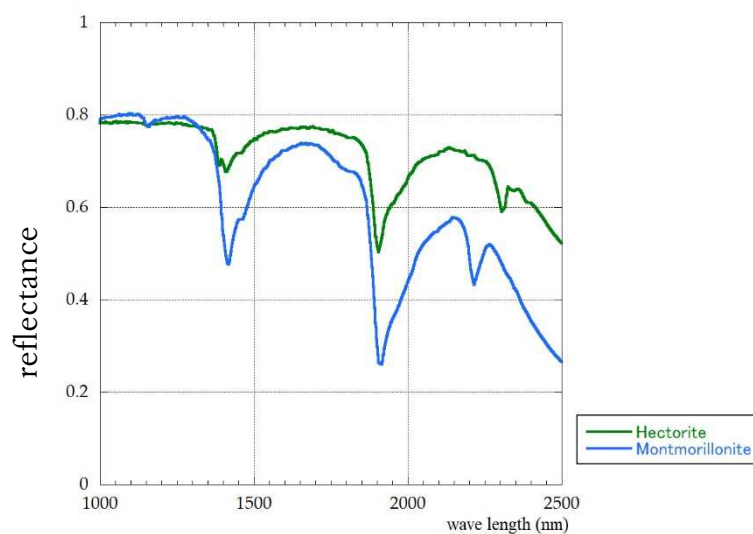


図 6-8 Hectorite と Montmorillonite の分光反射スペクトルチャート

図 6-9 に MLP_M_HQ の学習時とテスト時の正答率と損失関数を示す。MLP_S_HQ 同様にテスト時の正答率と損失関数はほぼ安定した動きを示す。損失関数が 0.5 を下回るのは 10 エポック程度である。テスト時の正答率は比較的安定している。しかし損失関数はわずかに上下動を示している。最良正答率は 97.9%(表 6-9)を、交差検証の f 値は 96.8%(表 6-10)を、AUG データも含んだ all_f 値は 95.5%(表 6-10)をそれぞれ示した。MLP_M_HQ は変質鉱物自動同定のためのモデル化が可能と考えられる。

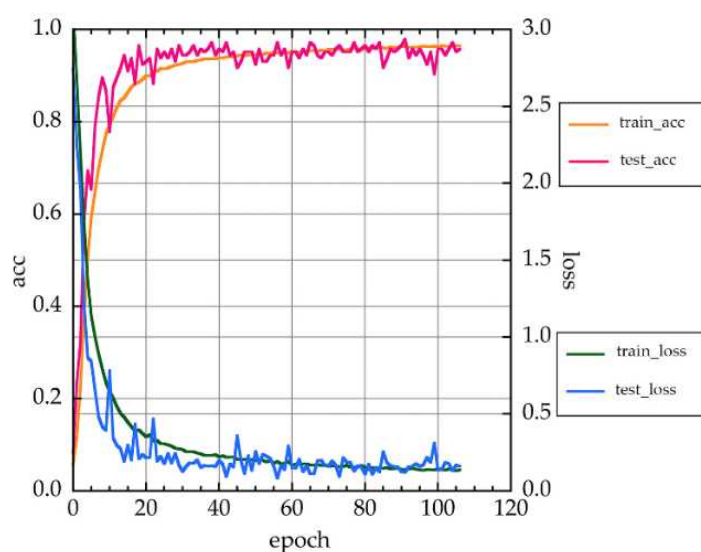


図 6-9 MLP_M_HQ の学習時とテスト時の正答率と損失関数

図 6-10 に MLP_L_HQ の学習時とテスト時の正答率と損失関数を示す。MLP_S_HQ、MLP_M_HQ と比較してテスト時の正答率と損失関数はやや大きな上下動が認められる。損失関数が 0.5 を下回るのも 25 エポック程度である。損失関数はエポックが大きくなるにつれやや乱れだしている。HQ 処理なしの MLP_L に認められたスパイク状の損失関数の乱れ(図 6-3)は認められない。テスト時の正答率は比較的安定している。最良正答率は 99.3%(表 6-9)で、交差検定の f 値は 97.1%(表 6-10)、all_f 値は 94.8%(表 6-10)でモデル化は 150 エポックより前で行うことになるが検討を要する。

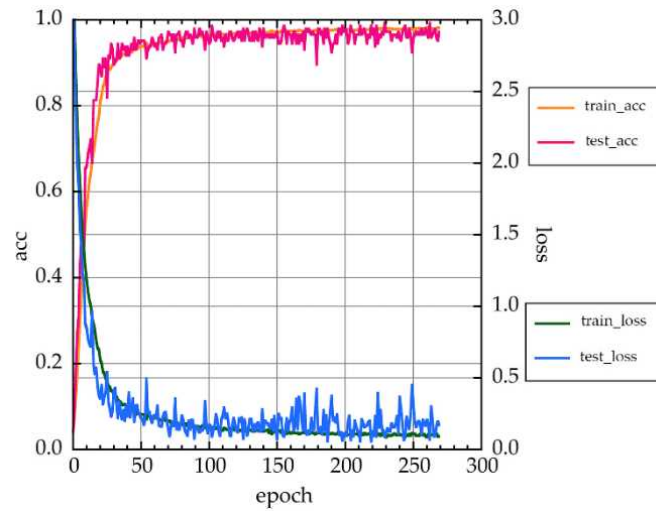


図 6-10 MLP_L_HQ の学習時とテスト時の正答率と損失関数

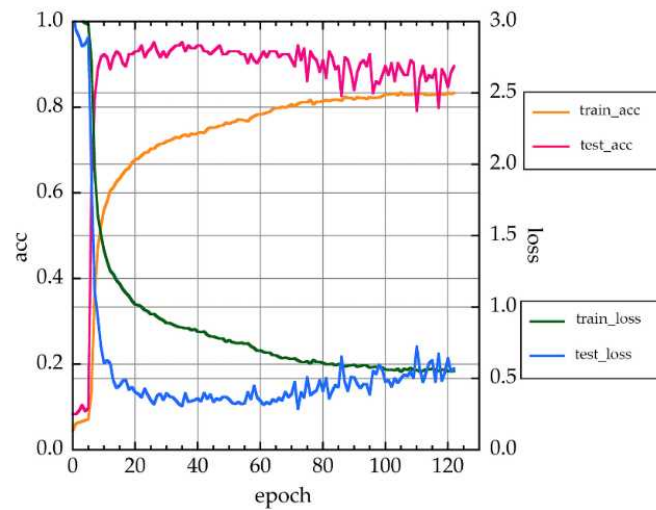


図 6-11 CNN1d_HQ の学習時とテスト時の正答率と損失関数

図 6-11 の CNN1d_HQ は図 6-4 の HQ 処理なしの結果に比べてやや良好な結果を示している。しかし 70 エポック程度から過学習が認められ、正答率も最後は 80%程度まで落ちている。最良正答率は 95.1(表 6-9)で、交差検証の f 値は 92.3、AUG データを含む all_f は 91.6 である(表 6-10)をそれぞれ示す。学習時の正答率が 120 エポックを過ぎても 80%程度である。変質鉱物自動同定のための CNN1d_HQ のモデル化は難しいと考えられる。

図 6-12 に CNN2d_HQ の学習時とテスト時の正答率と損失関数を示す。テスト時の正答率、誤差関数の動きが大きい。エポック 15 付近から次第に損失関数が大きく上下動を始める。学習時の正答率が 60 エポックを過ぎても 90 程度である。変質鉱物自動同定のための CNN2d_HQ のモデル化は難しいと考えられる。

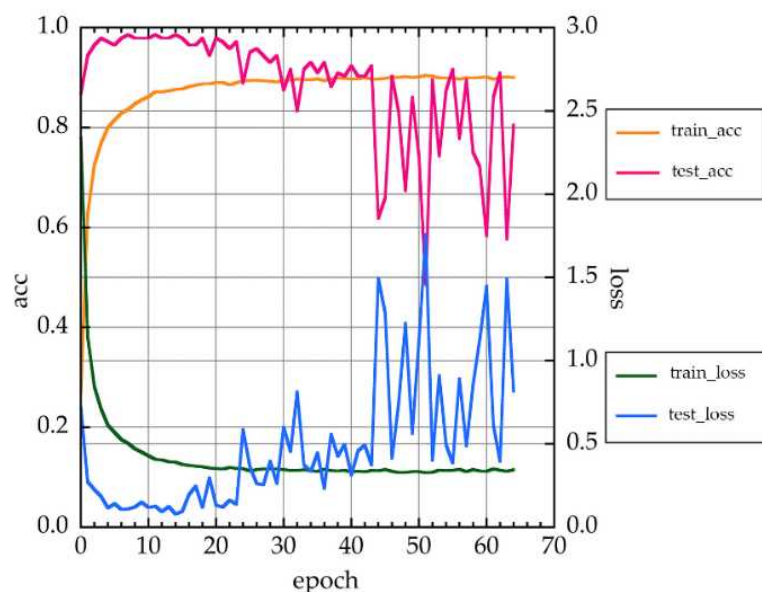


図 6-12 CNN2d_HQ の学習時とテスト時の正答率と損失関数

HQ 処理を行ったデータに対して各ネットワークを適用して学習をさせ変質鉱物同定を行った。その結果、HQ 処理なしに比べて全般的に正答率・f 値は上昇した。大きく正答率が向上した CNN1d を除くと f 値で 0.4~1.4%の向上である。

テスト時の最良正答率はいずれも高い値となっており、CNN1d を除けば最低でも 97.9%である（表 6-9）。表 6-10 は交差検証を行ったもので、CNN1d を除けば各ニューラルネットワークは 97%程度の f 値を示している。

各ネットワークの学習時とテスト時の正答率と損失関数の動きと f 値を総合的に考えると変質鉱物自動同定のための CNN のモデル化はほぼ困難であり、MLP の 3 モデルでは MLP_S_HQ と MLP_M_HQ はモデル化が可能と考えられる。また、各ネットワークの正答率の高さから考えて AUG 処理によって波長方向の誤差を有している分光反射スペクトルも正しく同定していると考えられる。

表 6-9 ネットワーク毎の最良テスト正答率%(HQ 処理済み)

number of measurements	mineral species	epoch	type of layer	test_accuracy
30 times	24	35	MLP_S	98.6
		91	MLP_M	97.9
		229	MLP_L	99.3
		26	CNN1d	95.1
		7	CNN2d	98.6

表 6-10 ネットワーク毎の交差検証の平均結果%(HQ 処理済み)

nw	acc	pr	rc	f	all_acc	all_pr	all_rc	all_f
MLP_S	97.5	97.8	97.5	97.5	96.4	96.5	96.4	96.4
MLP_M	96.8	97.1	96.8	96.8	95.5	95.8	95.5	95.5
MLP_L	97.1	97.5	97.1	97.1	94.8	95.2	94.8	94.8
CNN1d	92.5	94.1	92.5	92.3	92.0	93.1	92.0	91.6
CNN2d	97.4	97.7	97.4	97.4	96.1	96.4	96.1	96.1

nw : network, acc : accuracy, pr : precision, re : recall

7 議論

本研究では 24 種類の鉱物同定に CNN と MLP の 2 つの手法を用いたが共に HQ 処理済みでは高い f 値・正答率を示した。

しかし CNN2d では学習時の正答率が 90%弱で足踏みをし、エポックの後半でスパイク状の損失関数の悪化が認められ、過学習を示した(図 6-5、図 6-12 など)。原因の特定はできていないが、勾配降下法において最小値付近で一旦は計算が収束傾向にあるものが、その後、突然損失関数が大きなところに移動している。特に HQ 処理ありの方が顕著である。

CNN1d は HQ 処理の有無に関わらず学習時の正答率が 70%弱(HQ 処理なし)、80%弱(HQ 処理あり)と MLP の学習時の正答率に比較して低かった。CNN1d では反射率を 1 次元のデータとして 1×3 の畳み込みを行ったが、分光反射スペクトルの特徴量抽出は十分ではない。テスト時の損失関数が発散傾向を示している。正答率は CNN 全般で比較的高いが、変質鉱物自動同定のためのモデル化を行えるような同定結果はない。

本研究では MLP の方が CNN に比べてより安定した損失関数の動きであり、高い正答率・f 値を示した。

本研究では変質鉱物同定には MLP_S_HQ および MLP_M_HQ のモデル化が可能という結果であった。MLP_S_HQ の層構造は表 6-1 同様には全結合層が 4 枚(ドロップアウト層が 3 枚)であるが今後の実装化のためにそれをさらに薄くして最適な層構造を検討した。その結果は表 7-1 のとおりで正答率(acc)・f 値を見ると 2 層で f 値が 97.6%を示すが、図 7-1 に見られるように 2 層構造ではテスト時の損失関数が過学習を示しており、3 層構造(出力層を含まず)が最適であるという結果になった。

表 7-1 層構造と正答率の関係

層数	層1	層2	層3	層4	acc	all_rc	f	all_f
4	512	512	256	24	97.6	97.9	97.6	97.6
3	512	256	24		97.8	98.0	97.8	97.8
2	512	24			97.6	98.0	97.6	97.7
2	256	24			97.6	97.9	97.6	97.6
2	128	24			97.4	97.7	97.4	97.4
2	64	24			93.1	94.0	93.1	93.0
2	32	24			53.5	53.4	53.5	50.3

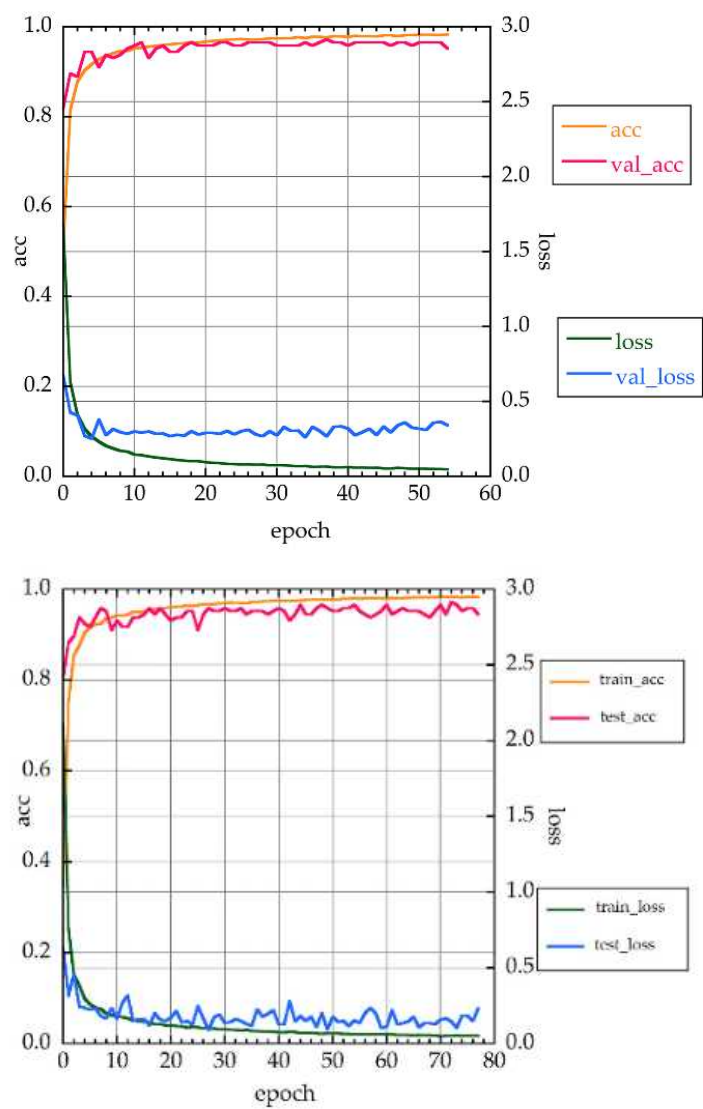


図 7-1 2 層構造(上)と 3 層構造(下)の学習時とテスト時の正答率と損失関数

表 7-2 に各ネットワーク構造と f 値、過学習の状態、変質鉱物自動同定のためのモデル化の可否などの一覧を示す。この表からも MLP_HQ の 3 層構造(MLP_3L_HQ)が高い f 値を示し 1、過学習がほとんどなく、大きなノイズもなくモデル化に適していることがわかる。MLP_S_HQ、MLP_M_HQ もモデル化が可能である。この他、MLP_S と MLP_L_HQ はやや過学習の傾向が認められるもののモデル化の可能性はある。

表 7-2 各ネットワークとモデル化の可能性

ネットワークモデル	全結合	ドロップ アウト	畳み込み 層	プーリン グ	f 値	過学習	大きな ノイズ	学習時の正答率	モデル化
CNN1d	3	5	3	3	89.8	強	あり	70%	×
CNN2d	4	6	4	4	96.7	中	あり	85%	×
CNN2d_Hi	4	6	4	4	97.1	中	あり	85%	×
CNN1d_HQ	3	5	3	3	92.3	強	あり	80%	×
CNN2d_HQ	4	6	4	4	97.4	強	顕著	85%	×
MLP_S	4	3	-	-	96.1	弱	あり	90%以上	△
MLP_M	9	8	-	-	96.4	中	顕著	90%以上	×
MLP_L	13	12	-	-	96.2	中	顕著	90%以上	×
MLP_2L_HQ	2	1	-	-	97.6	強	なし	90%以上	×
MLP_3L_HQ	3	2	-	-	97.8	微弱	なし	90%以上	◎
MLP_S_HQ	4	3	-	-	97.5	微弱	なし	90%以上	○
MLP_M_HQ	9	8	-	-	96.8	弱	なし	90%以上	○
MLP_L_HQ	13	12	-	-	97.1	弱-中	あり	90%以上	△

学習回数が正答率に与える影響をみるために鉱物数を同じ条件で学習回数を変化させて、学習回数が正答率に与える影響の確認を行った。検証に用いたネットワークは MLP_S_HQ である。条件は HQ 処理済みの AUG データで学習をさせておりテストデータは 10 鉱物と 24 鉱物の 2 種類である。学習回数は 20 回と 30 回である。その結果は表 7-3 に示すように同定鉱物数と測定回数の関係は測定回数と同じなら同定鉱物数が少ないほど正答率が高くなり、また変質鉱物数が同じならば測定回数が多いほど正答率があがった。測定回数が 30 回で同定鉱物数が 10 の場合は f 値が 99.3%となっている。図 7-2 に測定回数 30 回鉱物数 10 の学習時とテスト時の正答率と損失関数を示す。高いテスト時の正答率を示しているが損失関数にやや乱れがある。

表 7-3 測定回数と同定鉱物数の関係(%)

条件	acc	pr	rc	f	all_acc	all_pr	all_rc	all_f
測定回数20 鉱物数10	98.1	98.5	98.2	98.1	97.2	97.6	97.4	97.3
測定回数30 鉱物数10	99.3	99.4	99.3	99.3	97.2	97.4	97.2	97.2
測定回数20 鉱物数24	96.9	97.6	96.9	96.9	95.8	96.1	95.8	95.7
測定回数30 鉱物数24	97.5	97.8	97.5	97.5	96.4	96.5	96.4	96.4

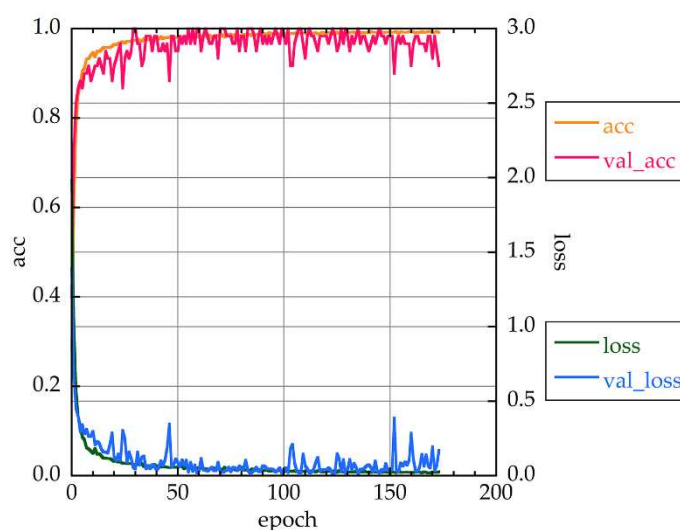


図 7-2 測定回数 30 回鉱物数 10 の学習時とテスト時の正答率と損失関数

本研究では HQ 処理を行っていない各ネットワークではやや過学習の傾向を示した。これを解消するためには AUG 処理時のパラメータを大きくして変化に富んだデータ群を作成する必要がある。但しこのパラメータが大きすぎると変化には富むがオリジナルデータとの相違が出てくる可能性があり、最終的に正答率が下がってくる。Tanaka et al.(2019)では AUG 処理時に大きめの乱数の初期値を与えており、MLP_S_HQ で f 値は 97.1%と本研究(97.5%)に比べてやや低い。このように AUG 処理のパラメータの強さと正答率・f 値はトレードオフの関係にある。HQ 処理は過学習抑制の機能があるので効果的な HQ 処理を行えばより有効な結果が得られると考えられるが、試行錯誤の側面がある。本研究では結果の評価は交差検証で行った。交差検証はテストデータを k 分割して k 回検証を行うものでデータに偏在性などがあるときに有効である。グラフ化

された損失関数とエポックの関係なども合わせて検討する必要がある。ホールドアウト検証などで最良正答率のみを評価に用いると出来たモデルの評価を過大評価する可能性がある。なおモデル化するネットワークを決定した後は学習結果とその時のパラメータは各エポック毎に保存されており、最良正答率を示すエポックの回数で精度の高い学習モデルを生成することになる。MLP_S_HQ でモデリングを行う場合は表 6-9 の最良テスト正答率の 98.6%の能力が出るように 35 エポック目に保存されているパラメータを復元させることになる。

深層学習用のツールは多種あり有効であるが、f 値・正答率を比較する場合、python、tensorflow、keras 等のツールのマイナーバージョンアップが頻繁に行われており、時期が違えば結果もわずかながら異なった。その意味で手法の比較検討、パラメータの検討は同時期に行う必要がある。本研究でも一連の計算を再度、最後にまとめて実行して比較検討を行った。

8 結論

本研究は、深層学習を用いて分光反射スペクトルデータから変質鉱物同定を自動的に
行う方法の開発と、その有効性の検討を目的とした。

MLP では学習する層構造の厚さを変化させて変質鉱物同定に最適な層構造を検討し
た。CNN では 1 次元と 2 次元で同様の検討を行った。さらに HQ 処理が変質鉱物同定
に及ぼす有効性についても検討を行った。

MLP では学習する層構造を異なる 3 つのモデル(MLP_S, MLP_M, MLP_L)を設定し
た。層構造の増減に伴う f 値・正答率への影響を検討した。MLP 法では学習する層構造
が薄い場合において変質鉱物同定が十分に行えた。逆に層構造の厚いものでは過学習の
影響が出てきた。また HQ 処理を施さないものは過学習が発生した。HQ 処理を施した
ものには過学習は発生しなかった。一般的に HQ 処理を施した方が高い f 値を示す。
HQ 処理を施したいずれの MLP も f 値は 97%前後(表 6-10)であり、MLP_S_HQ が過
学習も示さずモデル化可能と考えられる。さらに MLP_S_HQ(4 層構造)よりも層構造
を少なくして薄い層構造が変質鉱物同定に適しているかを検討した。その結果は 3 層構
造が高い f 値(97.8%)を示し、過学習も微弱で最もモデル化に適していると考えられる。
実装するモデルを選択する時は正答率や f 値などの数値だけに頼らず、正答率・損失関
数-エポック数のグラフの各パラメータの動きをみて、最適なモデルを決めるべきであ
る。なお同定モデルはオーギュメンテーション処理によって 5nm 程度の波長方向の誤
差も許容して同定できるものになっている。

CNN は CNN1d と CNN2d の検討を行った。CNN1d は分光反射率に 1×3 のフィル
タを畳み込んでいくものであるが形態を認識するプロセスが弱いと考えられた。
CNN2d は通常の画像で高い最良正答率(98.6%)を示している。交差検証では、 430×286
の画像に対しても 96.7 という高い f 値(表 6-10)を示す。画素数を落とした画像でも同
定が可能ということは、CNN2d が特徴として捉えているものが単に吸収の位置だけで
はなく全体の形状と吸収の位置の関係も含んでいると考えられる。CNN2d_Hi は交差
検証の f 値が CNN2d よりやや値が高い(表 6-7)。しかし、CNN は学習時の正答率が
80~90%と低くスプック後半にスパイク状のノイズが損失関数に表れた。また、実際の

学習時間が長く、また原画像に対して適切な前処理が必要である。本研究では CNN で変質鉱物自動同定のためのモデル化できるものではなく、MLP に比較して精度が低い結果となった。

用いる波長帯によっては分光反射スペクトルのノイズが強調される現象が予想されるため、AUG 処理の波長域の設定には注意が必要である。今回は 900-2,500nm に統一したがノイズの影響もなく、可視域の影響も受けず妥当だったと考えられる。

各鉱物の測定回数を増やしデータ数を増やすことは変質鉱物の同定に大きく寄与するが、現実的な測定回数は 1 試料につき数十回程度であると考えられる。量的には良質なデータが多いほどよいが、学習に多くの時間を要する。質に関しては実際のテスト入力データがあまり良質でないと予測されるならば、それに合わせたデータを学習データとして用意する必要がある。また、学習データの偏在性を確認する上でも交差検証は重要な評価法である。AUG 処理の乱数パラメータ調整などビッグデータの生成のあり方も正答率に大きく関係する。ドロップアウト層は極めて優秀な層で、ニューラルネットワークを分断することにより過学習をかなり押さえることができた。EarlyStopping 関数はエポック数の打ち切りを決定できる関数であり、損失関数の動きが安定せずに収束しない場合に有効であった。

短波長赤外域でのスペクトルの深層学習を用いた研究では田辺ほか(1994)があるが専用ハードウェアを用いて同定させるものであった。また三輪ほか(2017)はメーカー製の機械学習ソフトウェア・ハードウェアを用いて火山灰粒子解析を行っていた。共に高額ソフトウェア・ハードウェアである。しかし近年、GPU が深層学習に転用できるようになり、デスクトップパソコン+ α で深層学習環境が手に入る時代になった。また Google Colaboratory のように深層学習の普及を目的としたプロジェクトもあり、ウェブ上でかつ無料で深層学習のための環境を提供してくれるサービスも始まっている。深層学習をウェブ上にある画像データ集などを使用し新たなネットワーク構造を提案して学習させ、正答率の向上を検討することが容易にできる環境がウェブ上に整ってきた。このように研究者の環境は劇的に変化してきている。しかし本研究のように変質鉱物同定というテーマに沿って深層学習を適用するとデータの入手、データの規格化など前準備に長い時間を必要とする。データ拡張などでも長い時間を要し試行錯誤であった。

同定対象変質鋳物は 24 鋳物であるが、深層学習の能力を考えると、今後増やしていくと考えられる。鋳物に不純物がわずかでも含まれていると同定結果に影響を示すため混入する不純物のタイプ別にラベルも作成すべきである。地域ごと、鋳床タイプ別、酸化銅鋳などの他の鋳物グループへの適用も今後の重要な課題である。またさらなる正答率向上のために MLP ではニューラルネットワークの層構造の検討、ニューロン数の検討、CNN では畳み込み、プーリングの回数の検討など多数の検討項目がある。

今後は変質鋳物の混合比などの連続値を予想する回帰問題でも深層学習の有効性を確認したい。作成された変質鋳物自動同定のための学習済みモデルは、市販の携帯型変質鋳物同定システムに深層学習プログラムとして組み込むことが可能でコンピュータに実装することにより、現実的な変質鋳物同定システムが製作できると考えられる。

謝辞

本研究を進めるに当たり、多くの方々にお世話になりました。記して感謝の意を表します。終始熱心なご指導を頂いた指導教官の山口 靖教授、ゼミで終始適切な助言を頂きました高野雅夫教授、平野恭弘准教授、宮坂隆文助教に感謝の意を表します。また講座の皆様からも適切な助言をいただきました。

独立行政法人 石油天然ガス・金属鉱物資源機構資源探査部の増田一夫課長、川上 裕課長代理、野田周帆氏、浅野友紀瑛氏の皆様には FieldSpec4Hi-Res の借用をはじめひとかたならぬお世話になりました。株式会社システム計画研究所の鶴英雄氏、染野和昭氏からは最新の深層学習理論と実践をご教示いただきました。ジオテクノス株式会社の藤川真治氏には終始アドバイスを頂きました。深海資源開発株式会社の西圭介氏には終始論文作成のコメントを頂きました。以上の方々に深くお礼いたします。

本研究の一部は、独立行政法人石油天然ガス・金属鉱物資源機構との共同研究「深層学習を用いた分光反射スペクトルによる変質鉱物同定手法の開発」（代表：山口 靖）および JSPS 科学研究費補助金基盤研究(B)一般「ハイパースペクトル及び熱赤外データの統合処理による鉱物・岩相マッピング法開発（15H04225）」（代表：山口 靖）による補助を受けて行われました。

引用文献

Amari, S. (1998): Natural gradient works efficiently in learning. *Neural Computation*, Vol. 10, Issue 2, pp. 251-276.

Baldrige, A. M.; Hook, S. J.; Grove, C. I.; Rivera, G. (2009): The ASTER Spectral Library Version 2.0, *Remote Sensing of Environment*, Vol. 113, pp. 711-715.

坂野靖行, 古宇田亮一 (2014): 地質標本館所蔵鉱物の高波長分解能反射スペクトルデータ. 地質調査総合センター研究資料集. 産業技術総合研究所地質調査総合センター, No. 611.

Bay, H.; Tuytelaars, T.; Van Gool, L. (2006): Surf: Speeded up robust features. In: *European conference on computer vision*. Springer, Berlin, Heidelberg, pp. 404-417.

Bengio, Y.; Simard, P; Frasconi, P. (1994): "Learning long-term dependencies with gradient descent is difficult." *IEEE transactions on neural networks*, Vol. 5, No. 2, pp. 157-166.

Bottou, L. (2010): Large-scale machine learning with stochastic gradient descent, In: *Proceedings of COMPSTAT'2010*, Physica-Verlag HD, pp. 177-186.

Bouchard, G. (2007): Efficient bounds for the softmax function, applications to inference in hybrid models, In: *Presentation at the Workshop for Approximate Bayesian Inference in Continuous/Hybrid Systems at NIPS-07*, pp. 9.

Cheng, G.; and Guo, W. (2017): Rock images classification by using deep convolution neural network, *Journal of Physics: Conference Series*, Vol. 887, No. 1, pp. 012089.
doi:10.1088/1742-6596/887/1/012089

Clark, R. N.; Swayze, G. A.; Livo, K. E.; Kokaly, R. F.; Sutley, S. J.; Dalton, J. B.; Gent, C. A. (2003): Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *Journal of Geophysical Research: Planets*, Vol. 108, No. E12.
doi:10.1029/2002JE001847

Clark, R. N.; Swayze, G. A.; Wise, R.; Livo, E.; Hoefen, T.; Kokaly, R.; Sutley, S. J. (2007): *USGS digital spectral library splib06a*: U.S. Geological Survey. Digital Data Series 231. <https://speclab.cr.usgs.gov/spectral.lib06/ds231/index.html> (アクセス日, 2017/02/09)

COCO(Common Objects in Context) <http://cocodataset.org/index.htm#home> (アクセス日, 2019/03/30)

Dalal, N.; Triggs, B. (2005): Histograms of oriented gradients for human detection. *In: International Conference on computer vision & Pattern Recognition (CVPR'05)*. IEEE Computer Society, pp. 886-893.

Duchi, J.; Hazan, E.; Singer, Y. (2011): Adaptive subgradient methods for online learning and stochastic optimization, *Journal of Machine Learning Research*, pp. 2121-2159.

Fletcher, R.; Powell, M. J. D. (1963): A rapidly convergent descent method for minimization. *The computer journal*, Vol. 6, No. 2, pp. 163-168.

Fukushima, K. (1980): "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position." *Biological cybernetics*, Vol. 36, No. 4, pp. 193-202.

Goodfellow, I.; Bengio, Y.; Courville, A. (2016): *Deep Learning*. MIT press: Cambridge MA USA, 775p.

Green, A. A.; Craig, M. D. (1985): Analysis of aircraft spectrometer data with logarithmic residuals. *In: Proceedings of the Airborne Imaging Spectrometer Data Analysis Workshop*. JPL Publication, pp. 111-119.

Han, J.; Moraga, C. (1995): The influence of the sigmoid function parameters on the speed of backpropagation learning. *In: International Workshop on Artificial Neural Networks*. Springer, Berlin, Heidelberg, pp. 95-201.

Hinton, G.; Srivastava, N.; Swersky, K. (2014): Lecture 6e, Rmsprop, Divide the gradient by a running average of its recent magnitude.

https://www.cs.toronto.edu/%7Etijmen/csc321/slides/lecture_slides_lec6.pdf (アクセス日, 2019/03/30)

Ishikawa, S. T.; Gulick, V. C. (2013): An Automated Classification of Mineral Spectra. *44th Lunar and Planetary Science Conference, Woodlands, Texas, USA, 18–22 March, 1719, 3085pdf*.

石川洋平, 沢口俊美, 岩谷伸一, 堀内正俊 (1976): 黒鉱鉱床の下盤石英安山岩溶岩による探査標的へのアプローチ -その活動様式と深沢鉱床の変質鉱化ハローの追跡-, 鉱山地質, Vol. 26, No. 2, pp. 105-117.

神嶋敏宏, 麻生秀樹, 安田宗樹, 前田新一, 岡野原大輔, 岡谷貴之, 久保陽太郎, ポレガラダヌシカ (2015): 深層学習-Deep Learning-. 人工知能学会監修, 267p.

Kingma, D. P.; Jimmy, B. (2014): *Adam A method for stochastic optimization*, arXiv preprint arXiv:1412.6980.

Krizhevsky, A.; Sutskever, I.; Hinton, G. E. (2012): Imagenet classification with deep convolutional neural networks, *In: Advances in neural information processing systems*, pp. 1097-1105.

Li, M.; Zhang, T.; Chen, Y.; Smola, A. J. (2014): Efficient mini-batch training for stochastic optimization. *In: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 661-670.

Lin, T. Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollar, P.; Zitnick, C. L. (2014): Microsoft coco: Common objects in context. *In: European conference on computer vision*. Springer, pp. 740-755.

Liu, J.; Osadchy, M.; Ashton, L.; Foster, M.; Solomone, C. J.; Gibson, S. J. (2017): Deep convolutional neural networks for Raman spectrum recognition: a unified solution. *Journal of Analyst*, Vol. 142, No 21, pp. 4067-4074.

Lowe, D. G. (2004): Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, Vol. 60, No 2, pp. 91-110.

METI Agency for Natural Resources and Energy (2001): *ASTER data usage manual, Development of exploration technology by remote sensing*, Survey report on mineral exploration technology development.

三輪学央, 長井雅史, 入山 宙, 上田英樹 (2017): ディープラーニングを用いた火山灰分類に関する試行分析. 日本火山学会講演予稿集 日本火山学会. 特定非営利活動法人 日本火山学会, pp. 110-110.

Montero, S.; Irene, C.; Brimhall, G.H. (2001): Semi-automated Mineral Identification Algorithm for Ultraviolet, Visible and Near-Infrared Reflectance Spectroscopy. *In Proceedings of the 6th Annual Conference of the International Association for Mathematical Geology, Cancun, Mexico*, 14pp.

Nair, V.; Hinton, G. E. (2010): Rectified linear units improve restricted boltzmann machines. *In: Proceedings of the 27th international conference on machine learning, ICML-10*, pp. 807-814.

岡谷貴之 (2015): 深層学習 Deep learning. 講談社, 165p.

RELAB; NASA Reflectance Experiment Laboratory

http://www.planetary.brown.edu/relabdocs/relab_related_data.htm (アクセス日, 2019/03/30)

Rosenblatt, F. (1958): The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review*, Vol. 65, pp. 386-408.
doi:10.1037/h0042519.

RRUFF Project

<http://rruff.info> (アクセス日, 2019/03/30)

Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. (1986): Learning representations by back-propagating errors. *Nature*, Vol. 323, No 9, pp. 533-536. doi:10.1038/323533a0.

斎藤康毅 (2016): ゼロから作る Deep Learning: Python で学ぶディープラーニングの理論と実装. オライリー・ジャパン, 298p.

Solar, M.; Perez, P.; Watkins, F. (2008): Neural Recognition of Minerals, *Artificial Intelligence in Theory and Practice II*, pp 433-437.

Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. (2014): Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, Vol. 15, Issue 1, pp. 1929-1958.

田辺和俊, 上坂博亨, 井上常史, 高橋博之, 田中壮一郎 (1994): ニューラルネットワークによる近赤外スペクトルからの鉱物成分の同定, *分析化学*, Vol. 43, No 10, pp. 765-769.
doi:10.2116/bunsekikagaku.43.765

Tanaka, S.; Tsuru, H.; Someno, K.; Yamaguchi, Y. (2019): Identification of Alteration Minerals from Unstable Reflectance Spectra Using a Deep Learning Method, *Geosciences*, Vol. 9, No 5, pp. 195-213.
doi:10.3390/geosciences9050195

通商産業省資源エネルギー庁(1992): 平成 4 年度鉱物資源探査技術開発調査報告書, 変質鉱物簡易同定技術開発, 204p.

浦井 稔, 佐藤 功, 二宮芳樹, 古宇田亮一, 宮崎芳徳, 山口 靖 (1989): 可視～短波長赤外域における岩石・鉱物の反射スペクトル・カタログ, 地質リモートセンシングの研究 (昭和 62・63 年度大型プロジェクト研究開発成果報告書), 地質調査所, pp. 129-135.

Wilson, A. C.; Roelofs, R.; Stern, M.; Srebro, N.; Recht, B. (2017): The marginal value of adaptive gradient methods in machine learning. *In: Advances in Neural Information Processing Systems*, pp. 4148-4158.

Xu, B.; Wang, N.; Chen, T.; Li, M. (2015): *Empirical evaluation of rectified activations in convolutional network*. arXiv preprint arXiv:1505.00853, 5 p.

吉原 篤, 滝口哲也, 有木康雄 (2017): 衛星画像における分類精度の向上を目的とした CNN の導入および非学習領域への応用検証, 都市安全研究センター研究報告、神戸大学, Vol. 21, pp. 151-156.

Zeiler, M. D. (2012): "ADADELTA: an adaptive learning rate method." arXiv preprint arXiv:1212.5701, 6 p.