

地質粒子画像分類を支援する深層学習を活用したソフトウェア

宮川歩夢¹・松本恵子²・板木拓也¹・下司信夫^{2,3*}

1. 産業技術総合研究所 地質調査総合センター 地質情報研究部門
2. 産業技術総合研究所 地質調査総合センター 活断層・火山研究部門
3. *現 九州大学 大学院理学研究院 地球惑星科学部門

MIYAKAWA Ayumu, MATSUMOTO Keiko, ITAKI Takuya, GESHI Nobuo (2024) Software using deep learning to support geological particle image classification. *Open-File Report of the Geological Survey of Japan, AIST*, no. 750, 7p.

Abstract: Recent advances in artificial intelligence (AI) technologies have revolutionized several fields, including geoscience, by enabling the processing of massive amounts of data. The ability of AI to perform image-based classification, especially enhanced by deep learning, has improved the accuracy of image classification tasks. However, the rapid development of AI technologies has also brought challenges, in particular that high-level programming skills, complex data preprocessing and parameter tuning, and the need for specialized knowledge can be barriers to the use of AI. We have developed the GSJ Particle Analyzer (GSJ PA). This application allows classification of geological sample images without extensive programming or AI expertise. It focuses on geological particles consisting of granular materials such as rocks and microfossils. The application emphasizes ease of use and efficiency by dividing the workload of AI-based image processing into three software components for different stages of processing: Clipping Tool for image pre-processing, Training Tool for building models using transfer learning, and Classification Tool for classifying images using trained models.

Keywords: geological sample, image-based classification, deep learning, software

1. 概要

近年、AIを用いた各種技術は様々な分野で急速に発展しており、多くの産業や研究分野で活用されており、競争力の維持や技術者の不足を補うためにも、AI技術の適切な活用が重要である。地球科学分野においても、AI技術を導入することによりこれまで困難であった大量の情報処理が可能になってきている（例えば、Itaki et al. 2020; Uchide 2020）。AI技術の得意とする分野の一つとして、画像の特徴に基づく分類が挙げられる。特に、ディープラーニングをはじめとするAI技術の発展により、画像分類タスクの精度が向上している。

地球科学においても多くの視覚情報に基づいて様々な分析・判断がなされることから、AI技術による画像分類を適切に活用することで、試料の同定といった専門家の知見を要するような作業のAI技術による代替が見込まれる。これにより、作業の高速化、作業工程の省力化、人材減少の懸念の解消が期待される。

AI技術の普及に伴い、研究者や技術者が直面する課題も増加している。特に、AI技術の急速な進展により、これらの技術を活用するための知識不足が顕著な問題となっている。たとえばAI技術が進化しても、その利用に高度なプログラミングスキルが求められる場合、これまでプログラミングに携わってこなかった人々にとっては、その技術の利用が困難である。さらに、データの前処理やパラメータ調整が複雑であったり、専門的な知識が必要であったりすることも、AI技術を活用する際の障壁となる。これらの課題を克服するために、使いやすいソフトウェアやツールの開発がますます重要となっている。

AI技術の地球科学分野への活用の重要性に対し、地球科学系の研究者が必ずしもプログラミングを得意としているわけではないことを考慮し、プログラミングやAI技術に精通していなくても地質試料の画像分類が可能となるよう、GSJ Particle Analyzer (GSJ PA)として本ソフトウェアを開発した。特に、岩石や微化石などの粒状物質で構成される多くの地質粒子に焦点を当て、大量の粒子が撮影された画像から粒子の種類に応じて分類することを目的としている。さらに、本ソフトウェアは使いやすさと効率性を重視し、AIによる画像処理にかかる作業を分担して行えるよう、処理の段階別に以下の3つのソフトウェアで構成されている。

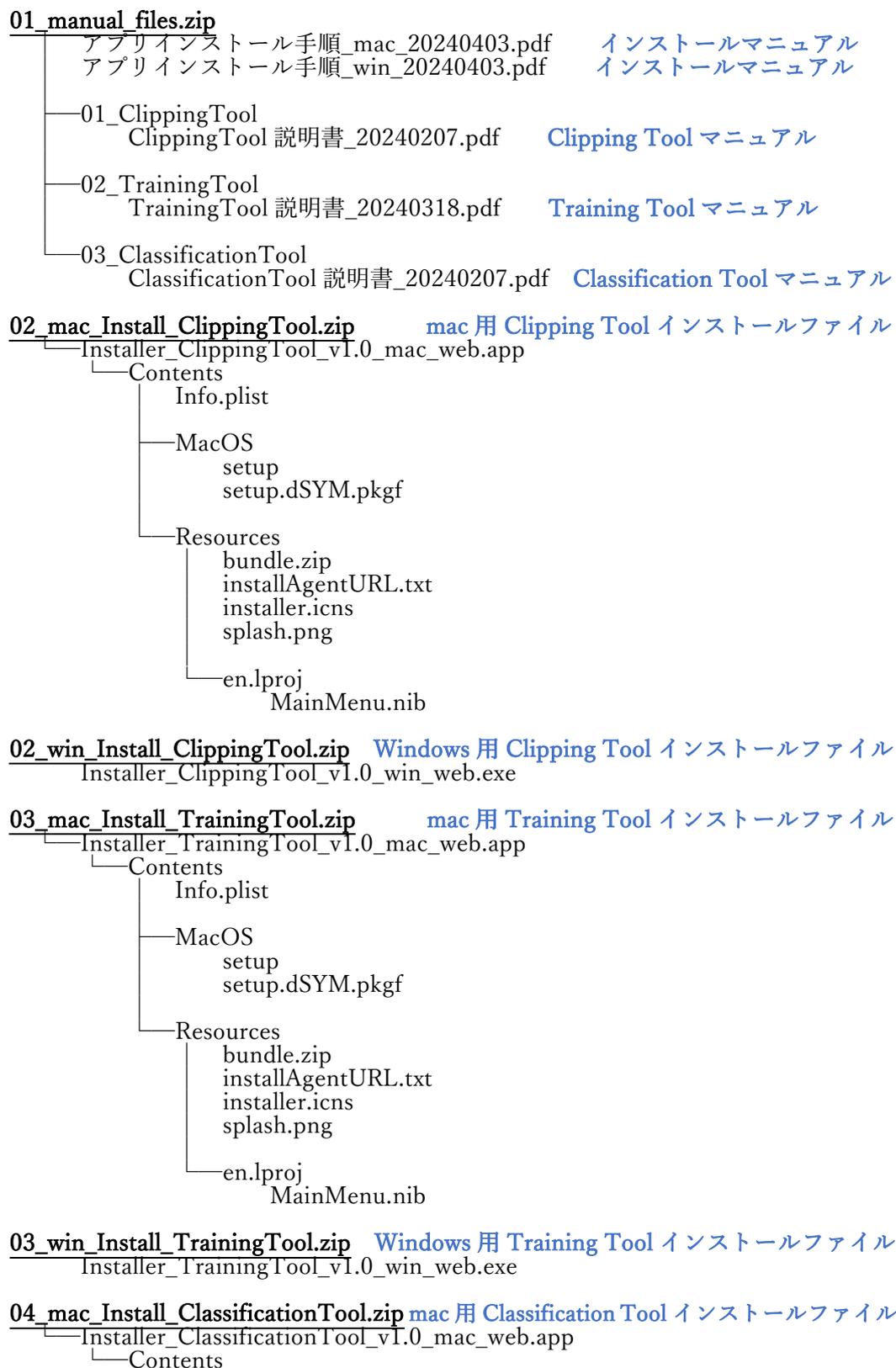
1. 粒子画像切り出しソフト (Clipping Tool) : 画像から粒子を切り出すためのソフトウェアである。これにより、データの前処理を行い、学習に用いる教師画像および分類するための画像を作成する。
2. 画像分類転移学習ソフト (Training Tool) : 転移学習を利用して、既存の学習済みのモデルをベースに新しい教師データを学習させることで、新しいタスクを実行可能なモデルを作成するソフトウェアである。これにより、学習時間の短縮と精度向上が期待できる。
3. 画像分類ソフト (Classification Tool) : 学習済みのモデルを利用して、画像の分類を行うソフトウェアである。これにより、高速に画像分類が可能である。

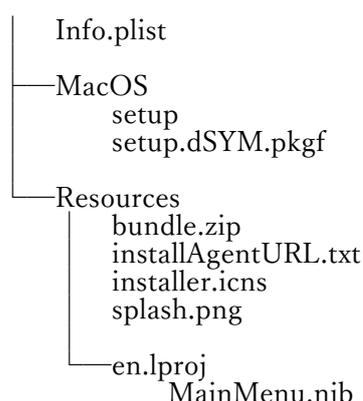
以下では、上記3ソフトウェアの概要について紹介する。具体的なインストールおよび利用方法については、別添のマニュアルを参照されたい。

2. 本資料に含まれるファイル

本資料に含まれるファイル群は、ソフトウェアインストールファイル、マニュアル、例題

用のテストファイルで構成される。青字にて主要なフォルダ・ファイルを説明する。





04_win_Install_ClassificationTool.zip Windows用 Classification Tool インストールファイル
 Installer_ClassificationTool_v1.0_win_web.exe

05_sample_data_ClippingTool.zip Clipping Tool サンプルデータ
 SampleImage.jpg サンプル粒子顕微鏡画像

06_sample_data_TrainingTool.zip Training Tool サンプルデータ
 TestTrainData.zip 分類済みサンプル画像群 (要展開)

07_sample_data_ClassificationTool.zip Classification Tool サンプルデータ
 netTransfer.mat 学習済みモデル
 TestData.zip 分類用サンプル画像群 (要展開)

3. 本資料に含まれるプログラムの概要

3.1. 粒子画像切り出しソフト：Clipping Tool

粒子画像切り出しソフト (Clipping Tool) は、画像から粒子を抽出し、粒子ごとに画像およびアスペクト比などの形態情報を保存するためのツールである。本ソフトウェアは、画像の読み込み、背景除去、粒子抽出、画像および粒子情報の出力、そしてピクセル値から長さへの変換といった一連の機能を提供する。これにより、ユーザーは画像中の粒子を個別に特徴を抽出し、後の作業で実施する学習や分類のためのデータセットを生成することができる。

粒子画像切り出しソフトの主な機能は、以下の3点である。

1. 粒子の抽出: 背景を除去し、粒子を矩形範囲で切り出すことで、個別の粒子画像を取得することができる。また、重なった粒子の分割設定や最小粒子径の設定することで、抽出する粒子を柔軟に調整することができる。
2. 粒子特性の分析: 抽出された粒子ごとに、長径、短径、径平均、アスペクト比、伸長度、円形度、周囲長包絡度、面積包絡度などの情報を取得する。これにより、粒子の形状やサイズなどの特徴をから、さらなる分析を行える。
3. 学習と分類のためのデータセット生成: 切り出された粒子画像を保存し、Training Tool や Classification Tool で使用する形式でデータセットを生成することができる。これに

より、粒子に基づく画像分類タスクを効果的に実行することが可能である。

3.2. 画像分類転移学習ソフト：Training Tool

画像分類転移学習ソフト（Training Tool）は、既存の学習済みモデルをベースにして、新たに教師データを与えることで、特定の画像分類タスクのために学習させるためのソフトウェアである。このソフトウェアは、深層学習を利用するために、学習パラメータを設定し、画像数の調整、画像データ拡張機能、学習過程の可視化、および学習結果の検証などの機能を提供する。本ソフトで用いる学習では、転移学習と呼ばれる技術を利用する。転移学習は、既存の深層学習用のネットワークに対して事前に学習を済ませたモデルを用いることで、初めから学習させるよりも短時間で効果的に学習させることができる技術である。

画像分類転移学習ソフトの主な機能は、以下の3点である。

1. 高速な画像分類学習: 転移学習を利用することで、既存の学習済みモデルをベースに、新たな画像分類タスクに対して効率的に学習を進めることができる。これにより、ユーザーは短時間でタスクに適したモデルを構築できる。
2. カスタマイズ可能な学習パラメータとデータ拡張: 学習パラメータを設定し、画像数の調整や画像データ拡張機能を利用することで、学習データセットに適した学習条件を整えることができる。これにより、教師画像が少ない場合でも効果的に学習を行うことができる。
3. 学習過程の可視化と結果の検証: モデルの学習過程をリアルタイムで可視化し、学習の進捗状況や評価指標を確認することができる。また、学習結果の検証機能を用いて、モデルの性能を評価することができる。これらの機能により、ユーザーはモデルの調整や改善を行える。

3.3. 画像分類ソフト：Classification Tool

画像分類ソフト（Classification Tool）は、学習によって得られた学習済みモデルを利用して、新たな画像データを迅速に分類するためのソフトウェアである。このソフトウェアは、学習済みモデルを用いた画像分類、分類の確度指定、各クラスに分類された画像の枚数表示、画像のクラス別保存、および各画像に対する分類結果とスコアのテキスト保存などの機能を有している。

学習済みネットワークからの画像分類ソフトの主な機能は、以下の3点である。

1. 迅速な画像分類: 学習済みモデルを利用することで、新たな画像データを迅速に分類する。これにより、ユーザーは短時間で分類結果を得ることができる。
2. 確度指定による分類結果の信頼度の調整: 信頼する確度を指定することで、指定された確度以上の画像のみを分類することができる。これにより、分類結果の信頼性を高める

ことができる。

3. 分類結果の整理と保存: 各クラスに分類された画像の枚数を表示し、画像をクラス別に保存することができる。また、各画像に対する分類結果とスコアをテキスト形式で保存することができる。これにより、ユーザーは分類結果を容易に解析し、さらなる研究に活用することができる。

4. インストールと利用マニュアル

本ソフトウェアは、Matlab®を使用して開発されている。本ソフトウェアは、Windows および Mac の両方のオペレーティングシステムに対応している。本ソフトウェアは、コンパイル済みのバイナリ形式で配布され、ユーザーはソースコードをコンパイルすることなく、ソフトウェアをインストールして利用できる。ただし、インストールの際には、別途 MATLAB Runtime のインストールが必要である。MATLAB Runtime は、Matlab®がインストールされていない環境でも、Matlab®で開発されたソフトウェアを実行できるようにするためのランタイムライブラリである。MATLAB Runtime の詳細については、Mathworks 社の当該ウェブページを参照されたい。具体的なインストール方法および各ソフトウェアの利用マニュアルは別添のファイルを参照されたい。

MATLAB Runtime (Mathworks 社ウェブページ)

<https://jp.mathworks.com/products/compiler/matlab-runtime.html>

(アクセス日 2023 年 5 月 1 日)

5. ライセンスと利用規約:

本ソフトウェアは Mathworks 社の Software License Agreement に基づき使用される。また、著作者はソフトウェアの動作を保証することではなく、結果について責任を一切負わない。

MATLAB® (c) 1984 -2023 The MathWorks, Inc.

6. 免責

本プログラム・データ・マニュアルを使用することによって生じるいかなる損害も、国立研究開発法人産業技術総合研究所はその責を負わない。

7. 引用例

宮川歩夢・松本恵子・板木拓也・下司信夫 (2024) 粒子画像分類を支援する深層学習を活用したソフトウェア. 地質調査総合センター研究資料集, no. 750, 産業技術総合研究所地質調査総合センター, 7p.

8. 参考文献

Itaki, T., Taira, Y., Kuwamori, N., Maebayashi, T., Takeshima, S., & Toya, K. (2020). Automated collection of single species of microfossils using a deep learning–micromanipulator system. *Progress in Earth and Planetary Science*, 7, 1-7.

Uchide, T. (2020). Focal mechanisms of small earthquakes beneath the Japanese islands based on first-motion polarities picked using deep learning. *Geophysical Journal International*, 223(3), 1658-1671.