

卒業研究報告書

機械学習を用いたエンセラドスのクレーター 分布解析ツールの開発とその地形的理解

報告者

学籍番号: 1250019

氏名: 井上 昂輝

指導教員

山本 真行 教授

令和7年2月17日

高知工科大学 システム工学群 航空宇宙工学専攻

目次

1. 序論	1
1.1. 背景	1
1.2. エンセラドスという星	3
1.3. クレーター・カウンティング	3
1.4. 先行研究	4
1.5. 目的	5
2. 実験手法	6
2.1. YOLO	6
2.2. 学習データ	7
2.2.1. 画像データの確保	7
2.2.2. labeling によるアノテーション	7
2.2.3. 学習データの増し	10
2.2.4. 学習データの振り分け	10
2.3. モデルの評価指標	10
2.3.1. 損失関数	11
2.3.2. クラス分類・回帰	11
2.3.3. IoU (Intersection over Unit)	13
2.3.4. PR 曲線と mAP	14
2.3.5. 手動による評価	16
2.4. クレーター検出の手順	17
3. 結果	19
3.1. 3種類の損失値の比較	19
3.2. PR 曲線の比較	20
3.3. mAP の比較	23
3.4. モデルの決定と物体検出の結果	24
3.5. 氷の亀裂に対するモデル精度	26
3.6. クレーター密度とサイズ	27
4. 考察・評価	29
4.1. モデルの性能	29
4.1.1. 損失における性能差	29
4.1.2. 氷の亀裂に対する信頼度の閾値	30
4.1.3. 解像度が及ぼす物体検出への影響	30
4.1.4. 先行研究との比較	31
4.2. クレーターマップとエンセラドスの地形	34

5. 結論	37
謝辭	38
参考文献	39

1. 序論

1.1 背景

大学3年次の講義「宇宙探査工学」の課題にて宇宙探査に関して各々の興味のあることについて調べた。著者は土星に興味があったので土星について調べていた。著者が土星に興味を抱いたのは単純で、土星が持つリングがインパクトとして大きかったからだ。土星を調べていると、土星のリングは7つあり、その中に様々な衛星が存在しておりそれらの衛星が相互に影響し合っリングを形成していることを学んだ。講義が終わってしばらくしてから卒業研究のテーマを模索しているときにこの講義のことを思い出し、再び土星に関する情報を調べた。その時私が出会ったのがエンセラドス（図 1.1）という星（土星の衛星の一つ）である。



図 1.1 エンセラドスの写真[1]。英表記で Enceladus と書くが日本での呼び名は「エンケラドス」、「エンセラドス」、「エンセラダス」など表記ゆれしている。本論文ではエンセラドスに統一している。

我々が地球の外へ目を向け、宇宙を探求しようとする理由は様々だがその理由の一つに「生命の起源・地球外生命体の存在」への関心が挙げられる。

現在全宇宙内で生命体を確認されている星は我々の住むこの地球のみである。太陽系内では生命の居住可能領域として液体の水が存在できる。ハビタブルゾーンは地球全域と火星の一部のみとされてきた。火星表面では過去に水の流れた痕跡が見られたものの生命が存在した直接的な証拠とはなっていない。

ハビタブルゾーン外では生命が居住できる環境ではないものの小惑星や水星の極域では水の氷が確認されており、生命が存在していないとは必ずしも言えない。例えば、太陽系内で第2の大きさを誇る土星の第6衛星タイタンは極寒の地であるが地表に豊富な液体のメタンやエタンが存在することが確認されている。もしかしたら、そこには我々の常識とは全く異なる生命体が存在している可能性があるのだ。ところで、同じ土星の衛星の中にエンセラドスという星が存在する。土星の一番外側のリング内に存在する直径 500 km のちっぽけな氷衛星である。表面が氷で覆われた衛星ではあるもののその実態は 2007 年まで謎に包まれており、これと言って注目されることもなかった。しかし、2007 年に米国 NASA の土星探査機カッシーニの搭載カメラはこの状況を一変させる光景を偶然捉えてしまった。エンセラドスの南極付近から広範囲に間欠泉が噴出している様子 (図 1.2) が確認されたのだ。それは、内部に大規模な液体の水の海洋と熱水活動が存在することを示唆し、太陽からの距離による区分ではハビタブルゾーンでなくとも生命が居住できる可能性のある天体が存在するという直接的な証拠となった。

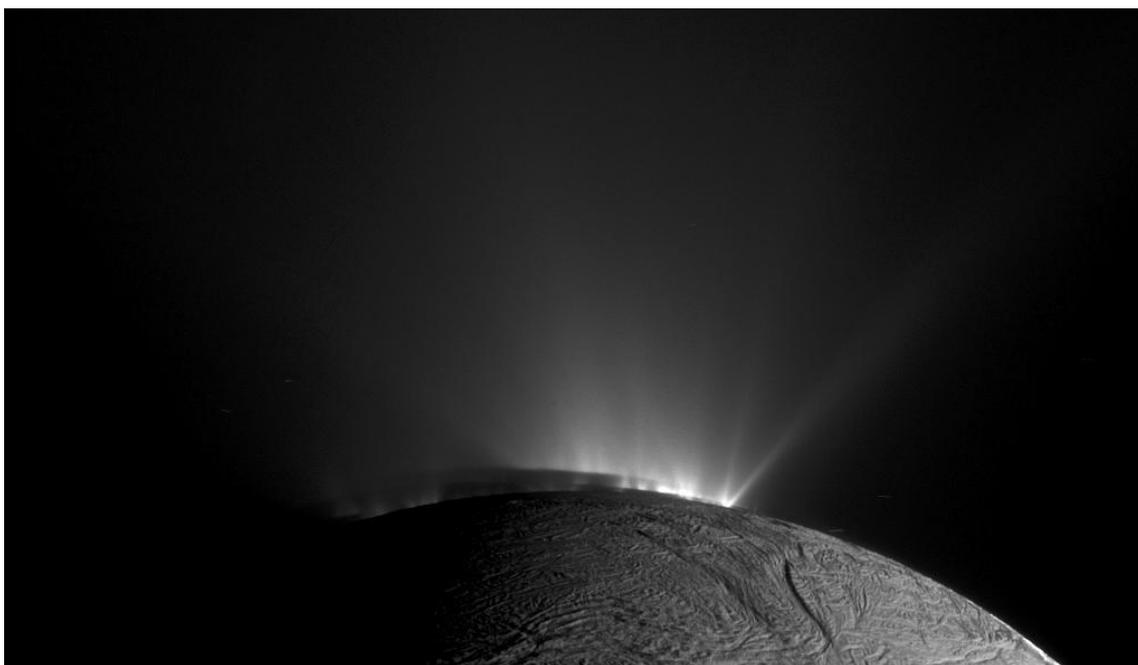


図 1.2 探査機カッシーニが捉えたプルーム噴出の写真[2]。エンセラドスの南極域から噴き出すプルーム (惑星内部の液体、水や氷) をとらえた写真。

著者はそこでエンセラドスの実態を探るべく、まずは天体の表層に視点を当てた。元来、宇宙の観測というものは天体の表面 (望遠鏡観測など) から始まり、やがて探査機によるフライバイ観測、周回軌道観測、着陸機によるその場観測やサンプルリターン、地震計などによる内部観測という順に天体は解析されていく。探査機を

むやみやたらに飛ばすということではできないので、天体の表面から得られる情報は非常に重要である。天体の表面には何があるのかというと、主に大気やクレーターなどの地形や模様である。特に、大気のない地球型惑星・岩石衛星・氷衛星ではクレーターが目につくものである。クレーターというのはある一つの天体に対して別の小天体が衝突した際にできる痕跡である。天体のクレーター数とサイズを測定することをクレーター・カウンティングといい、天体の表層構造や天体周辺の環境を考察する材料の一つとして用いられる。本研究は、手動で数えるには手間のかかるクレーター・カウンティングを機械学習モデルに自動で行ってもらうことを目的に行った。

1.2 エンセラドスという星

エンセラドスは直径 500 km 程度の小さな土星の衛星である。外側から氷殻、海洋、岩石のコアで構成されており、密度は $1,609 \text{ kg/m}^3$ である。エンセラドスは内部に液体の水を有しているという点で注目を集めている。その規模は地球の海洋と同程度であり、生命体存在の可能性が示唆されている。内部海の存在は 2007 年に探査機カッシーニがエンセラドスの南極域から液体の水が噴出する間欠泉現象（図 1.2）を捉えたことにより発覚した。この現象はプルーム噴出と呼ばれ、氷殻の地下貯水域に貯まったとされる水が内部の熱による圧力によって放出されたものであるというのが現在では有力な説となっている。全くの偶然で発見されたこの現象からエンセラドスは俄然注目を浴びることになり様々な研究が行われ探査計画が議論されることとなった。最近の研究では生命に必要な 3 要素のひとつである有機物（リン酸）も発見された (Frank Postberg et al., 2023) [3]。エンセラドスに対する生命の起源、生命体の存在の可能性は高まるばかりである。

1.3 クレーター・カウンティング

クレーター・カウンティングとは惑星や衛星の表面のクレーターの数を数えることである。クレーターの数や大きさがその星の表層年代や形成の歴史を推定する指標として使われる。例えば、地球と月では隕石が大量に衝突したといわれる後期重爆撃がある。この時期のインパクトの影響でそれ以前のクレーターがすべて消失してしまった。なぜそのような重爆撃が起こったのかというと、その時期に木星と土星の軌道が変化し小惑星帯が乱されてしまい、軌道から外れた小惑星が大量に降り注いだからといわれている。このような太陽系の歴史を推定する根拠とな

るのがクレーターなのである。

また、月の裏側は表側よりもクレーターの数が多い。これは潮汐ロックが関係している。惑星の自転周期と衛星の公転周期が一致することを潮汐ロックといい、このとき衛星は惑星に対して同じ面を向いて公転する。月が地球にいつも同じ面を向けているため月の表側は地球に守られる形となり必然的に裏側への衝突は多くなる。こういった理論は月の探査機が月の裏側の衛星写真を撮影したことによってより確実なものとなった。

クレーター・カウンティングを行う元来の方法は手動で数える、または画像処理による輪郭抽出などであった。手動での計測は時間がかかるが正確性は高いといえる。画像処理での計測は早いがクレーターを検出させるためのプログラムが難しかったりクレーターでないものを検出してしまったりとやや正確性に欠けるものであった。しかし、近年発展した機械学習の登場によりクレーターのカウントに適用可能な技術が向上した。大量の学習データを取り込むことにより、人間と同様に判断をしてくれる学習モデルを作成できる。学習済みモデルは対応したラベルを含んだ画像などであれば応用することができる。

しかし、機械学習にも欠点はある。それは、データが少なすぎると精度が全く得られないことや例えばデータが大量にあったとしても 100%の精度で検出するのは不可能ということだ。特に、クレーターのような「何をもってクレーターとするか」という定義が曖昧なものに対する検出は難しい。そもそも手動でのカウントもカウントした人間によってクレーターの判別が変わるため、機械学習においては様々な検出引数を用いて平均化する手法をとることとする。幸いクレータリングによる惑星の表層考察は正確な個数よりも全体としてみた各地での集密度やサイズを重視するため個々のクレーターについては 100%の正確でこだわらなくてもよいところがある。

1.4 先行研究

先行研究ではクレーターを認識する条件としてクレーターにあたる太陽の光とそれによってできる影をセットとして定義した (Blanco-Rojas et al., 2024) [4]。これにより、クレーター特有の光と影の当たり方をモデルに学習させ、クレーターを正確に認識できるようにした。また、学習データは 7,256 枚あり、そのうち 7 割を訓練データに、3 割を教師データに使用していた。学習には Google の開発した機械学習のためのソフトライブラリである TensorFlow 上で実行される Python ライブ

ラリのうち Keras というモデルを使用していた。

1.5 目的

本研究では、機械学習によるクレーターカウントツールを作成し、そのツールの性能を評価する。さらに、使用済み学習モデルを用いてエンセラドスのクレーターマップを作成しクレーターマップから読み取れることを考察することを目的とする。

2. 実験手法

本研究では機械学習を用いて、まず「クレーター」の特徴を学習させ、その学習済みモデルを適用して天体の全球マップに対してクレーターの検出を行いクレーターの数をカウントする手法を用いた。本研究では YOLOv8 という学習モデルを用いた。

2.1 YOLO

機械学習を行うモデルとして YOLO と呼ばれる手法がある。YOLO (You Look Only Once) とは ultralytics 社が無償で提供している物体検出、及び画像分類を得意とする機械学習モデルである。高速な検出と精度の高さが評価されており、何より物体検出を手軽に行えることが大きな利点である。デフォルトの学習済みモデルのみでも自転車や犬などの物体検出が行えるが、学習データさえあれば誰でも任意の物体検出を行うモデルを作ることができる。本研究で用いた YOLOv8 は 2023 年にリリースされた YOLO シリーズの 8 番目のモデルである。

YOLOv8 にはいくつかのモデルサイズがあり、モデルの最後に n, s, m, l, x と付く。それぞれのパフォーマンス指標は表 2.1 のようになっている。mAP (mean Average Precision) については後述するがモデルを幅広く評価するために用いられる指標であり数字が高いほど良いとされる。モデルのサイズが小さいほど性能は落ちるが CPU の処理速度は速くなっており、モデルサイズが大きいほど性能は高くなるが処理速度は遅くなってしまふ。精度を高めたいなら大きいサイズのモデルを、処理を早く済ませたいなら小さいサイズのモデルを使用すればよい。ここでは YOLOv8l を用いた。

表 2.1 YOLOv8 のパフォーマンス指標[5]

モデル	サイズ (ピクセル)	mAPval 50-95	速度 CPU ONNX (ms)	速度 A100 TensorRT (ms)	params (M)	FLOPs (B)
YOLOv8n	640	37.3	80.4	0.99	3.2	8.7
YOLOv8s	640	44.9	128.4	1.20	11.2	28.6
YOLOv8m	640	50.2	234.7	1.83	25.9	78.9
YOLOv8l	640	52.9	375.2	2.39	43.7	165.2
YOLOv8x	640	53.9	479.1	3.53	68.2	257.8

2.2 学習データ

2.2.1 画像データの確保

米国 NASA のホームページ上には膨大な数の探査データアーカイブ (PHOTOJOURNAL Image Galleries) [6]があり、月とエンセラドスについても多くの画像があるので、本研究ではそこから全球マップや天体の4方向画像などをダウンロードして用いる。各画像はサイズが大きすぎるので任意の大きさに画像を分割して、月の表面画像 (PIA13506、PIA13059、PIA13998、PIA14011、PIA14024、PIA00405) を 205 枚、エンセラドスの表面画像 (PIA14937、PIA18435) を 143 枚用意した。なお、分割した結果表面の情報が全くないもの (画像の端や完全に真っ黒な部分) は除外した。

2.2.2 labeling によるアノテーション

YOLO の学習データを作るために labeling というソフトを使用した。通常、学習データには対象のラベルと座標、大きさを決定するアノテーションという作業が必要で、それを行うためのソフトが labeling である。

まず、学習データ用の画像を一つのフォルダあるいは train (訓練データ) と val (教師データ) のフォルダに分けて置いておく。labelimg を開いたら、open directory をクリックし画像を置いたフォルダを選択し、次に save directory をクリックしてアノテーションファイルを保存するフォルダを選択する。アノテーションファイルについては後に記述する。これらのフォルダは本研究では後にデータの増し作業を行うために同じフォルダに保存しているが、別のフォルダに保存することが一般には推奨されている。そし

て、保存形式を YOLO 形式に変更しておく (図 2.1)。

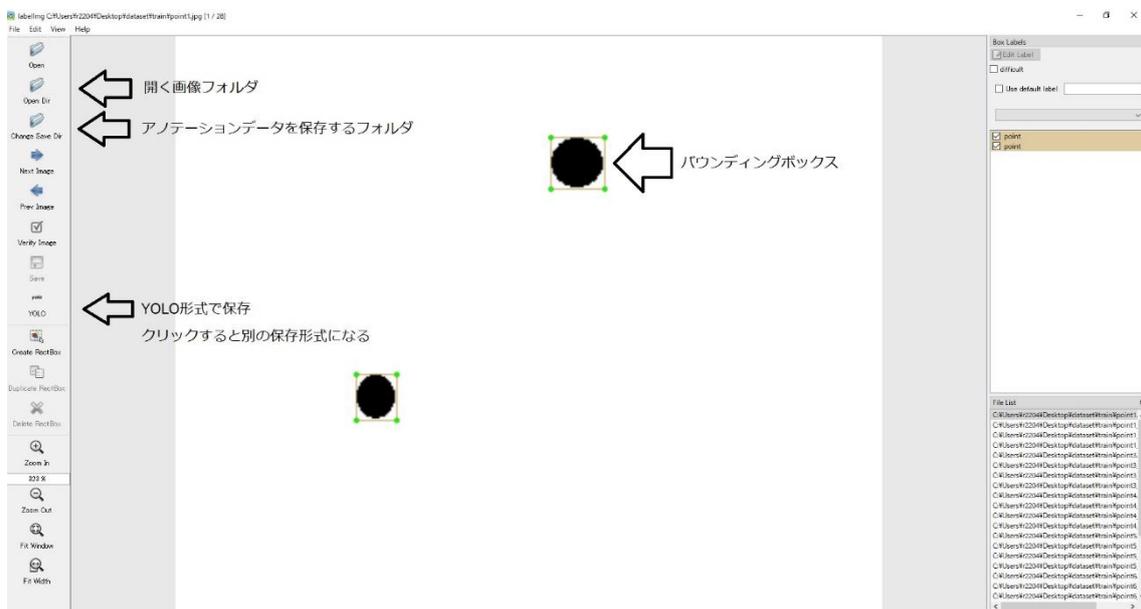


図 2.1 labeling で主に使用する機能の例

open directory でフォルダを選択すると、フォルダ内の画像が表示されるのでこれらの画像の対象物にアノテーションを行なっていく。Create Box または W キーを押すと Bounding Box (BB) を生成できる状態になるのでドラッグ&ドローをして対象物を囲む。囲んだ後にラベルの名前を付けられるので対象に対する任意の名前を付け、OK を押す。以降は既にあるラベルは項目として追加されるのでそれらをクリックしてアノテーションを行っていく。一つの画像に対してすべてのアノテーション作業が終わったら保存を押し、次の画像に切り替える。保存を押し、save directory にアノテーションファイルが保存され、アノテーションファイルには対象の BB のラベル、中心座標、幅、高さが記録される。このとき View 項目で Auto save モードにしておくとも画像を切り替えるたびにアノテーションファイルが保存されるので便利である。

今回は crater と old_crater という二つのラベルに分けてラベルを付けた。ラベルの定義は図 2.2 および図 2.3 に示すようになっている。

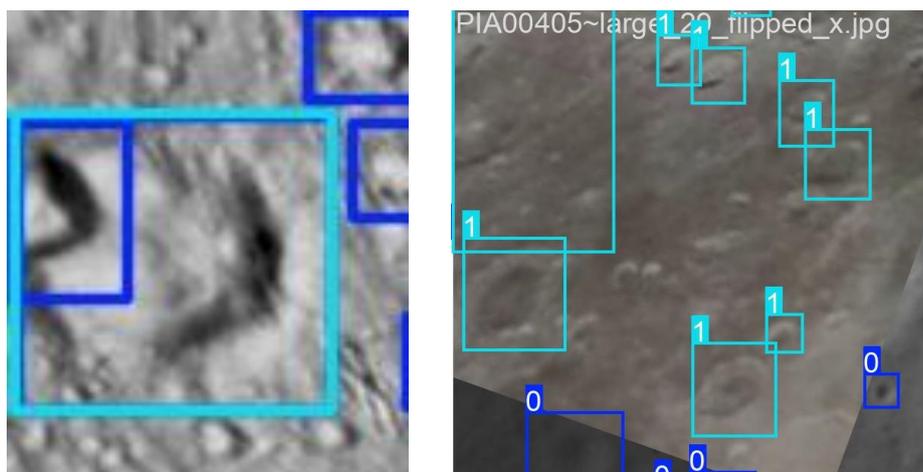


図 2.2 ラベル「old_crater」を認識している画像。「old_crater」は二つ以上重なったクレーター（左）、もしくは風化して識別しづらくなったクレーター（右）と定義した。大きなクレーターの上に小さなクレーターができた場合には天体の表面にその様子が残り地形的な歴史を残すが、その逆である場合は小さなクレーターが存在した歴史は残らないため天体表層の時間の経過を推定するのに重要である。また、風化したクレーターも天体表面での経過した時間を推定するのに重要であるため、一つのラベルとして定義とした。BB の色は水色である。

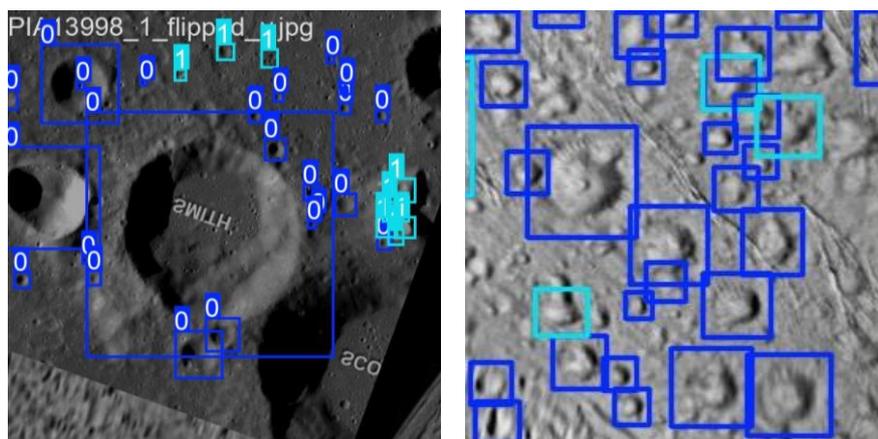


図 2.3 ラベル「crater」を認識している画像。「crater」は一般的にクレーターと認識できるものと定義しており、モデル内部では0のクラスとなっている。BB の色は青である。左の画像は学習モデルがテストとして物体検出を行っている様子。右の画像の中央左上付近のクレーターは「crater」として認識されているが、実際には2つ重なっており、今回の区分的には「old_crater」である。

以上の流れで全ての画像 348 枚（月：143 枚、エンセラドス：205 枚）に対してアノテーション作業を行った。

2.2.3 学習データの水増し

アノテーションを行って作成した学習データは全部で 348 枚であった。学習データは多すぎると過学習を起こしてしまうが多いほうが良いので、プログラムを用いて学習データを 4 倍にした。手法は簡単で、元の画像データに対して左右反転、上下反転、上下左右反転の処理を行いそれらに対応したアノテーションファイルも同時にした。Web 上で参照可能なプログラムコード [7] を引用してデータを水増しした。

水増しした結果、学習データ数は 1392 枚となった。この内、データとしての情報がほとんどないものを除外した結果、合計の学習データ数は 1265 枚となった。

2.2.4 学習データの振り分け

作成した 1265 枚の学習データを訓練データと教師データに分けた。この際の振り分けは一般的に 7 : 3 の振り分けが良いとされている。今回作成した学習モデルは 3 つあり、それらは教師データに設定する画像をエンセラドスのみに固定し、訓練データに設定する画像は以下の 3 つのケースに分けた。

一つ目はエンセラドスと月両方の画像データを訓練データに 849 枚、教師データにエンセラドスのみの画像データ 416 枚を設定した。これを「エンセラドス・月モデル」と定義する。

二つ目はエンセラドスのみの画像データを訓練データに 468 枚、教師データにエンセラドスの画像データ 200 枚をそれぞれ設定した。これを「エンセラドスモデル」と定義する。

三つめは月のみの画像データを訓練データに 724 枚、教師データにエンセラドスの画像データ 324 枚をそれぞれ設定した。これを「月モデル」と定義する。

2.3 モデルの評価指標

学習済みモデルの性能を評価する指標として重要なのは対象の物体の位置や正確な分類である。また、学習モデルが正しく学習できているかの指標として損失関数などが挙げられる。これらの指標を参考にすることでより良い性能のモデルを作

成することができる。

2.3.1 損失関数

機械学習における最終的な目標は最適解にたどり着くことである。それは物体検出においては対象の物体の位置とラベル（種類）を特定することである。対象の位置やラベルを特定するために学習を行うわけであるが、当然、学習段階で実際の値とモデルの予測した値が完全に合致することは難しい。このモデルの予測値と実際の値のズレを計算するのが損失関数であり、このズレをどのように効率よく小さくしていくかが最適化関数である。

機械学習の原点となっているニューラルネットワークの仕組みは人間の脳内にある神経細胞である「ニューロン」に基づく。ニューロンは入力を解釈して何らかの出力を行うが、機械学習におけるニューラルネットワークとは「入力層」、「中間層」、「出力層」からなり、それらは「パーセプトロン」という単純な構造がいくつも集まって構成されている。パーセプトロンもいくつかの入力を受けて処理を行い出力を行うものである。

ここで行われる処理がズレの計算で、損失関数が使われる。この出力には「重み」の値が付加されるが、これは次の出力結果がより良いものになるために付加される。この重みを決める役割を担うが最適化関数で、これらの過程を経て学習を終えた結果、予測値と実測値のズレが可能な限り小さくなっていけばよい、というのが機械学習の原理である。

逆に学習過程でズレが大きくなっていった場合は過学習やデータ不足が疑われる。同じ学習データを学習させ続けると同じものしか学習しないので早い段階で学習が完了した状態となる。この状態で学習を続けさせても既知のデータに対しては良い精度を出すが無知のデータに対しては同様の精度を出せない過学習という状態に陥ってしまう。そのため、学習データが少ないときはデータを増やすか学習回数を少なくする工夫が必要である。

なお、YOLO モデルに関しては Early stopping という機能があり、性能が頭打ちになった状態で一定回数の学習が行われると自動で学習が終了するので同じ学習回数で比較するとき以外は特に気にする必要はない。

2.3.2 クラス分類・回帰

いくつかのクラスを正確に分類、回帰（予測）するために YOLO では分類の損失関数として交差エントロピー誤差（CEE：Cross Entropy Error）、回帰の損失関数として平均二乗誤差（MSE：Mean Squared Error）を使用して

いる。それぞれの式を以下に記す。

$$CEE = - \sum_k t_k \ln y_k \quad (2.1)$$

$$MSE(y_i, \hat{y}_i) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.2)$$

CEE の変数について、 k 個のデータに対し、 t_k はクラス分類が正解なら 1、不正解なら 0 となる。 y_k はモデルの予測確率である。

MSE の変数について、 n 枚あるデータの i 番目のデータに対して y_i はクラスの実測値、 \hat{y}_i はモデルにおけるクラスの予測値である。

ところで、重み更新の最適化手法の一つに最急降下法がある。これはニューラルネットワークに入力されたすべてのデータに対し予測値を計算した後、損失関数を重みで微分し、微分して出た値から重みを更新する。これを重みの更新量が微小量になるまで繰り返す手法である。これらの手順を式で表すと以下のようなになる。

$$w_{t+1} = w_t - \alpha \nabla_w L(w) \quad (2.3)$$

w_t は更新 t 回目の重み、 α は学習率、 ∇_w は重み w での微分、 $L(w)$ は損失関数である。学習率というのは重みパラメータをどのくらい変化させるかを表すパラメータであり、値が大きいほど学習は早くなり、小さいほど学習は遅くなる。ただし、値が大きいと極小値に陥ったり過学習になってしまうなど最適解にたどり着かなくなる可能性が大きくなる。

一方で確率的勾配降下法 (SGD : Stochastic Gradient Descent) というものがあり、これは一回の重みの更新にすべてのデータを使うのではなくランダムに一つ選んだデータのみに対して行う手法である。これは最急降下法の場合では直近の極小値 (微分しても常に 0 になってしまう値) に陥ってしまう可能性があるという欠点を回避する方法である。しかし、通常の SGD では重みを更新するのに一つのデータしか使わないので非常に効率が悪くなってしまう。

この問題を解決するのにミニバッチ学習 SGD というものが使われる。全データをバッチといい、これをいくつかのデータセットに分けたものがミニ

バッチである。ミニバッチに分けることにより計算の並列化が行えるようになり学習が効率よく行えるようになる。例えば8つのミニバッチに分けたとすると、8つのミニバッチの中でそれぞれ一つずつデータを選び微分値を出す。8つの微分値から新たに重みを更新するので、1つのデータからのみ重みを更新した場合と比べて単純に8倍の効率で重みを更新できることになる。

YOLOの最適化方法はSGDが使われており、この方法はCEEとの相性がいいとされる。その理由は、CEEが自然対数から求める誤差であるからだ。最適化の手法として微分を用いるので微分しても計算式が複雑にならない自然対数関数($\ln x$)や自然指数関数(e^x)は最適化手法として用いやすいということになる。

2.3.3 IoU (Intersection over Union)

IoUとはBBの予測値と実測値の重なり具合を評価する指標であり対象の正確な位置を特定するのに重要である。図2.4のようにモデルの予測したBBと実際のBBがあるとき、分母はBBすべての領域となり、分子はBBの重なった領域のみとなる。重なった領域が全くなければ分子は0となるのでIoU=0となり、完全に重なっていればIoU=1となる。

物体検出で学習済みモデルを用いる際にIoUは閾値として使われる。IoU=0.5であれば予測値と実測値が5割重なっていれば検出対象となる。閾値が高くなるほど検出の条件は厳しくなるということである。

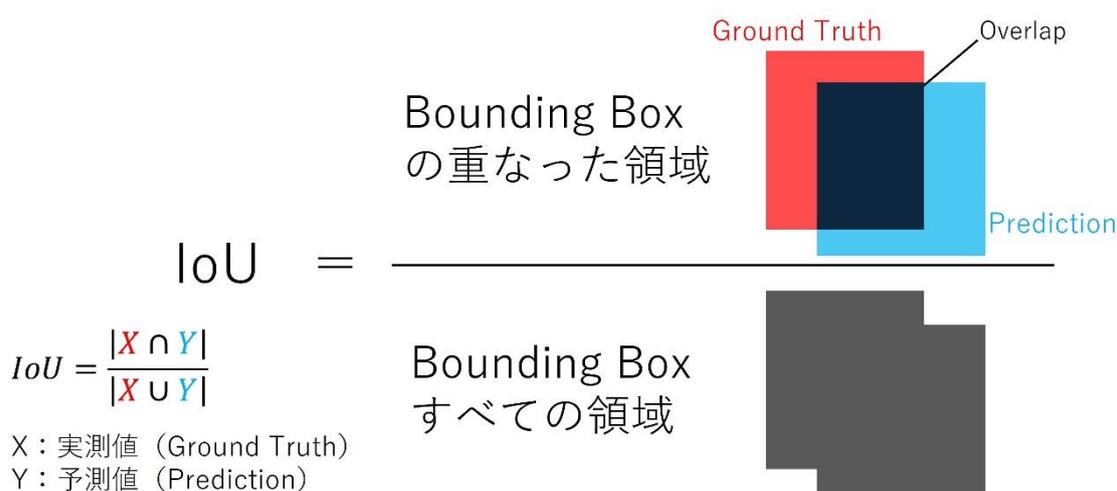


図 2.4 IoU の模式図

YOLO のモデル評価指標として `box_loss` というものがあり、これは IoU の損失を表している。一般的な IoU 損失は以下のような式で表される。

$$L_{IoU} = 1 - IoU \quad (2.4)$$

$$L_{IoU} = -\ln IoU \quad (2.5)$$

上記の式の L_{IoU} はどちらも IoU 損失であり予測値が実測値に近いほどこの値は小さくなる。しかし、この IoU 損失には欠点があり、それは予測と実測の BB が全く重なっていないときに BB 同士が近くであろうと遠くであろうと損失は 1 で一定になり、それ以上損失を小さくすることができない点である。このままでは学習が続かないという問題が生じてしまう。

YOLOv8 で使われている IoU 損失は「Smooth L1 損失関数」である。 i 番目の BB の中心座標あるいは四隅の座標の予測値を \hat{y}_i 、実測値を y_i とするとこれらの誤差は以下のようにおける。

$$a = \hat{y}_i - y_i \quad (2.6)$$

このとき、Smooth L1 損失関数は以下のように書ける。

$$\text{SmoothL1Loss}(a) = \begin{cases} \frac{1}{2}a^2, & |a| \leq 1 \\ |a| - \frac{1}{2}, & |a| > 1 \end{cases} \quad (2.7)$$

この式の表す意味は誤差が 1 以内の時は誤差の二乗に 0.5 を掛け、誤差が 1 より大きいときは誤差の絶対値から 0.5 を引いていくので、BB が重なっていないときでも誤差を小さくしていくことのできる関数である。

2.3.4 PR 曲線と mAP

予測と実際の結果を表す指標として適合率 P (Precision) と再現率 R (Recall) がある。適合率 P は (正しく検出できた数) / (全予測 BB 数) と表され、再現率 R は (正しく検出できた数) / (全正解 BB 数) と表される。これらの指標を表すパラメータとして TP (True Positive)、TN (True Negative)、FP (False Positive)、FN (False Negative) がある。予測の Positive か Negative に対し、実際には正解 (True) だったか不正解 (False) だった

かを表すパラメータである。先ほどの適合率 P と再現率 R を表す式に表すと以下のようなになる。

$$Precision = \frac{TP}{TP + FP} \quad (2.8)$$

$$Recall = \frac{TP}{TP + FN} \quad (2.9)$$

適合率の分母の全予測 BB とは、ある任意のラベルと認識した全ての BB の数のことで、正解であるか否かは問わない。一方、再現率の分母の全正解 BB とは、ある任意のラベルを正解としたときの全ての BB の数であり、予想したラベルが選ばれたラベルでなかったと判断されても実際は正しいラベルであれば全正解 BB に含まれる。

適合率 P はどのくらい正しく対象の物体を検出できているかを表す指標となる。TP の数が多いければ正しく検出できていることになり、逆に FP が多いときは誤検出が多いということになる。

再現率 R はどのくらい対象の物体を見逃さずに検出できているかの指標になる。FN が多いときはそれだけ正解のラベルを検出できずに見逃しているということになる。

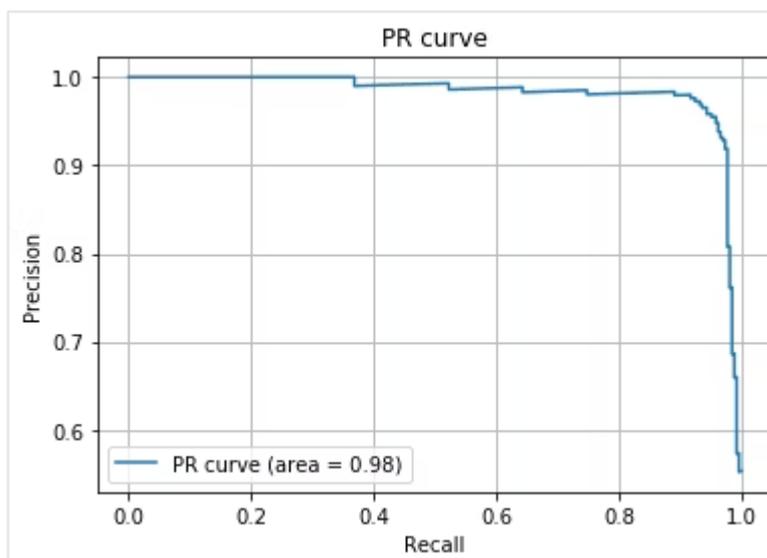


図 2.5 PR 曲線の例[8]。縦軸と横軸の 0 以上から PR 曲線で囲まれた部分が平均適合率 AP の値となる。

再現率を横軸 P に、適合率 R を縦軸にした時のグラフを PR 曲線という。PR 曲線が縦軸、横軸ともに 1 に近づくほどモデルとしての性能は良くなる。そして、PR 曲線の積分値をとった値が平均適合率 (AP: Average Precision) である (図 2.5)。PR 曲線は一つのクラスに対して一つなので AP も同様である。AP が大きいほどそのクラスに対する正しい検出が行われているかの指標になる。さらに、各クラスの AP を平均したものが mAP (mean Average Precision) である。この値によって全クラスを平均した値を表せるので学習モデルの性能を幅広く評価する指標として使われる。

2.3.5 手動による評価

土星を周回する氷衛星であるエンセラドスにおいてクレーターを識別するために厄介なのは氷上の亀裂の誤認である。クレーターと形は違えど亀裂によって生まれる光と影はクレーターのそれと誤認される可能性がある。よって、学習モデルが亀裂をどれくらい誤認するのかを信頼度を用いて評価する。信頼度 (confidence) とは、学習モデルが検出した物体に対して 2.4 節に後述する各項などから得られる評価に基づいて出力する予測確率である。信頼度が高ければその予測は正確である可能性が高く、低ければその予測は不正確である可能性が高いと言える。しかし、元々のモデルの性能が低ければ信頼度が高くとも正確ではない検出が増える可能性が高くなる。

評価手法としては信頼度に応じた亀裂の数と全物体検出数を比較する方法とした。学習モデルが亀裂をクレーターと誤認したかどうかは検出後に出力される画像を見なければわからないので亀裂のカウントは手動となった。

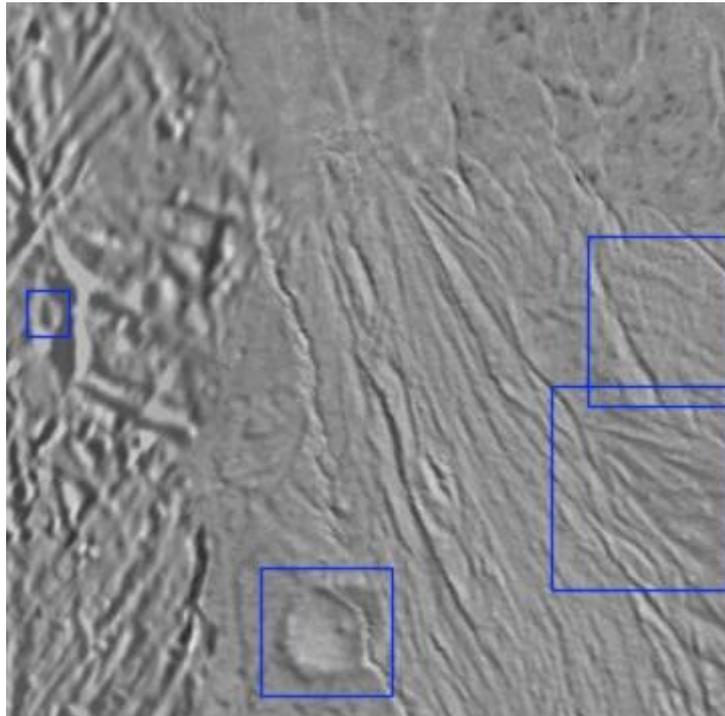


図 2.6 氷の模様をクレーターと誤認している例。右側二つのバウンディングボックス (BB) で囲まれた領域には明らかにクレーターが存在していない。中央下付近の物体はクレーターであると言える。左側の物体はどちらともいえない。

2.4 クレーター検出の手順

本項から実際に行ったクレーター検出の手順を記す。まず、準備段階として Google Drive 内のフォルダに dataset フォルダを作り、その中に訓練データ用のフォルダと教師データ用のフォルダを追加する。訓練データ用フォルダには訓練データとそれに付随したアノテーションファイルを入れておく。教師データも同様とする。これらのファイルをプログラムから読み込むために .yaml ファイル (図 2.7) が必要になるのでテキストファイルなどから作成しておく。このファイルは dataset フォルダ上に置いておく。次に Google Colab を開き任意のファイル (ソースコードを書くためのファイル) を作り Google Drive と同期した後、YOLOv8 を読み込む。そして、学習する際のパラメータを調整し、学習を開始する。学習が終わると損失や mAP などの学習結果をまとめたファイルが train フォルダ (/runs/detect/train) に作成されるので保存しておく。Google Colab との接続が切れると作成したファイルや出力結果は残らないので注意する。

```
1 #dataset.yamlファイルの中身
2
3 #トレーニングデータと検証用データの場所
4 train: /content/drive/MyDrive/Colab Notebooks/dataset/craters/train
5 val: /content/drive/MyDrive/Colab Notebooks/dataset/craters/val
6
7 #クラスの数と、各クラス名
8 names:
9     0: crater
10    1: old_crater
```

図 2.7 datase.yaml ファイルの設定

次に、Jupyter Notebook にて物体検出を行った。YOLOv8 を読み込み、学習済みモデルを読み込むプログラムを書く。学習結果をまとめたフォルダの中に weights というフォルダがあり、その中に last.pt と best.pt という学習済みモデルが保存されている。last.pt は最後の学習結果の重みを保存しており、best.pt は学習状況の中で最も性能の良かった時の重みを保存しているため基本的には best.pt を使って物体検出を行う。物体検出を行うと検出された画像と検出した物体の位置やラベルの情報が保存されたファイルが出力される。プログラム等を使いこれらのファイルから必要情報を抽出しクレーター数のカウントを行う。

3. 結果

3.1 3種類の損失値の比較

学習データをエンセラドスと月、エンセラドスのみ、月のみの3種類に分けた場合のボックス損失(box_loss)、クラス損失(cls_loss)、および畳み込み損失(dfl_loss)の収束状況を図 3.1 に示す。

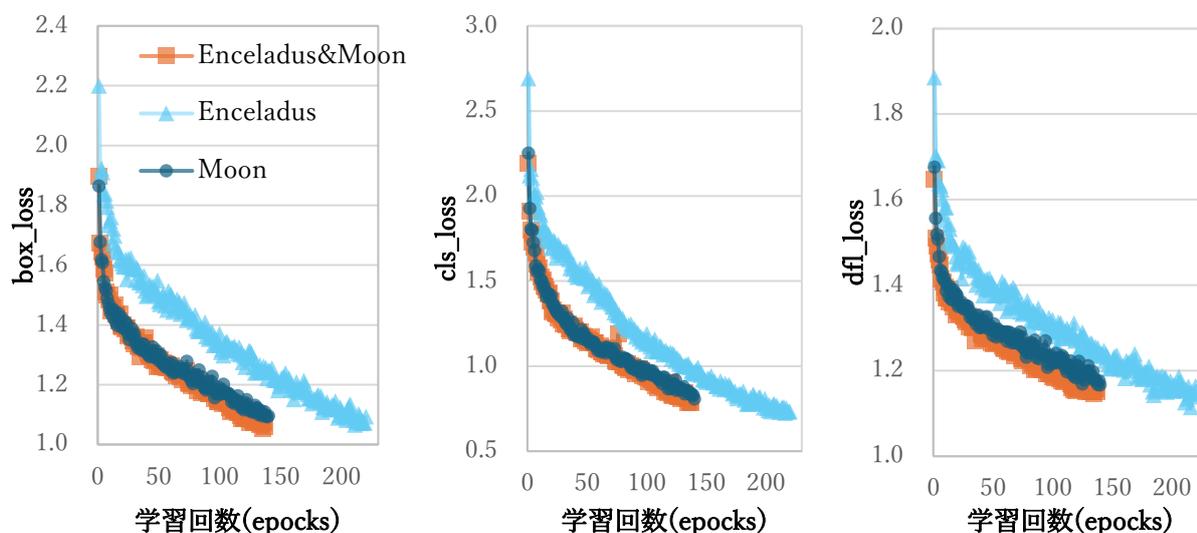


図 3.1 訓練データのボックス損失 (左)、クラス損失 (中)、畳み込み損失 (右)。エンセラドス・月モデル、エンセラドスモデル、月モデルによる損失の経過をそれぞれ表す。

図 3.1 を見るといづれも、数値の大きい方からエンセラドスモデル、月モデル、エンセラドス・月モデルという結果になった。また、学習回数ではエンセラドスのみ早期終了に達する回数が遅く、他の訓練データより 90 回ほど学習回数が多くなった。一方、各損失において、エンセラドスモデルが一時的に損失が他のものより最も小さくなる瞬間が訪れた。

最終的な損失は以下の表 3.1 のようになった。

表 3.1 最終的な各損失値

	box_loss	cls_loss	dfl_loss
Enceladus&Moon	1.06101	0.78681	1.15017
Enceladus	1.09354	0.73269	1.16391
Moon	1.09523	0.80706	1.16681

各損失値はいずれも月モデルが最も大きく、エンセラドスモデルの結果ではクラス損失だけが最も小さい損失値となった。月モデル及びエンセラドス・月モデルの場合で早期終了が起きた 140 回地点ではエンセラドスモデルの損失値がいずれも最も大きかったが、学習回数を多く繰り返したことにより収束し他の学習結果に近づく結果となった。

3.2 PR 曲線の比較

エンセラドス・月モデルの学習データでの PR 曲線は図 3.2 に示すようになった。

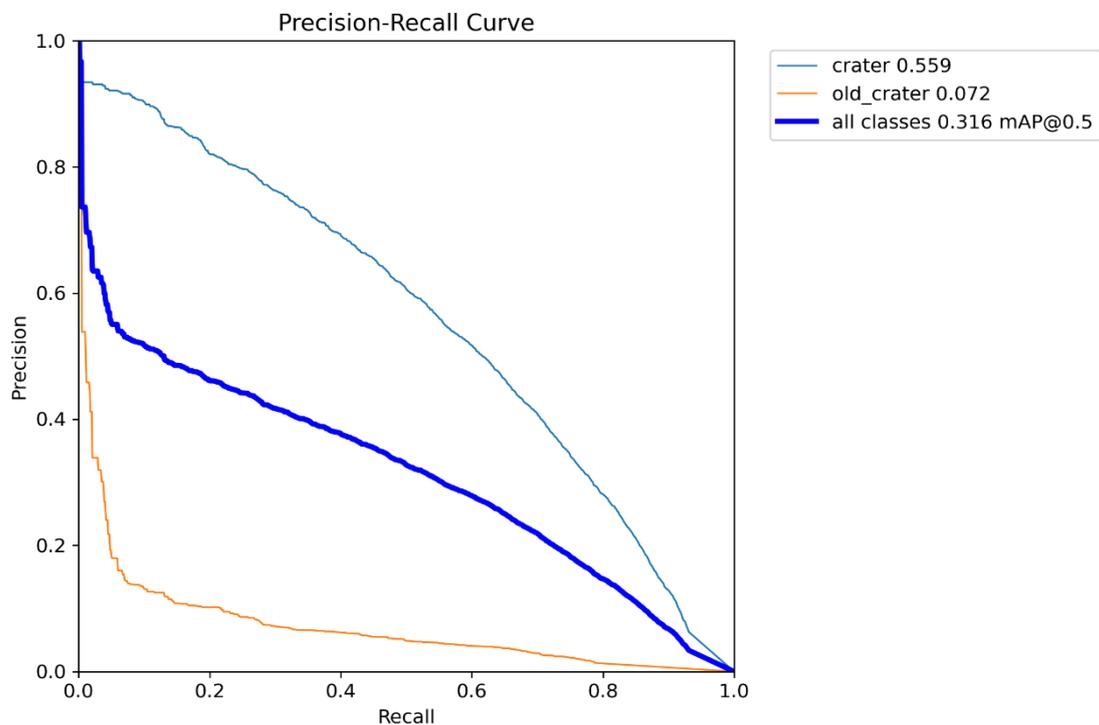


図 3.2 エンセラドス・月モデルの PR 曲線。クラス「crater」と「old_crater」と全クラス平均の PR 曲線。凡例の右側に示す数値は平均適合率であり all classes は IoU=0.5 のときの mAP である。

エンセラドスモデル、および月モデルの訓練データの PR 曲線はそれぞれ図 3.3、図 3.4 のようになった。

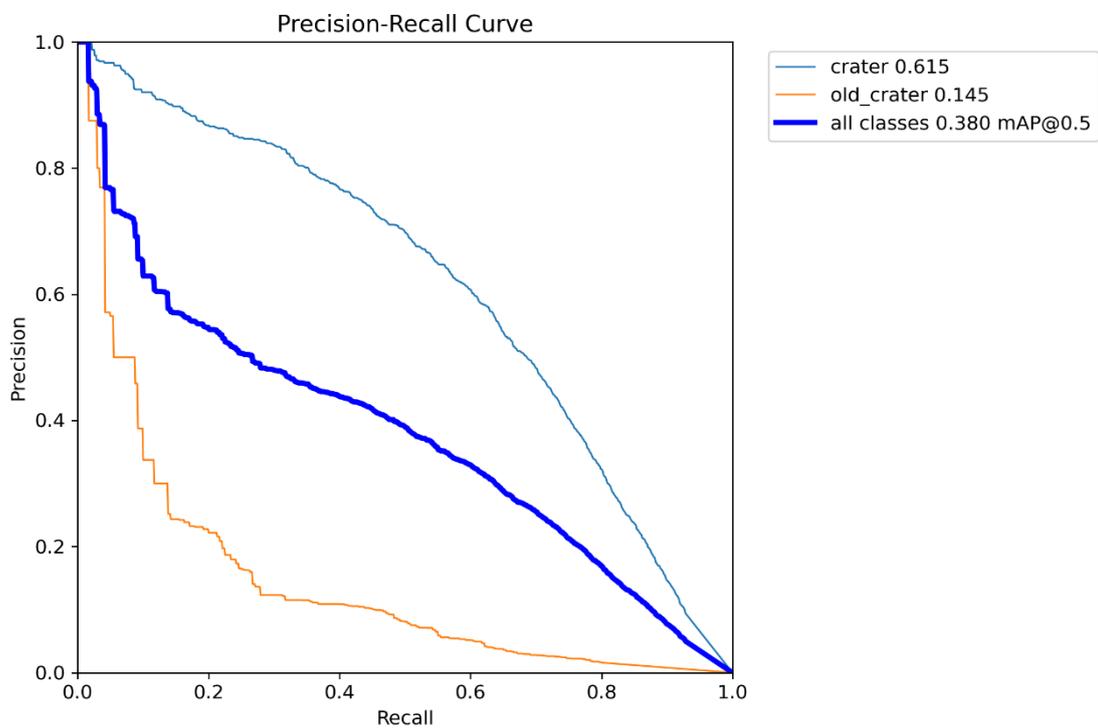


図 3.3 エンセラドスモデルの PR 曲線

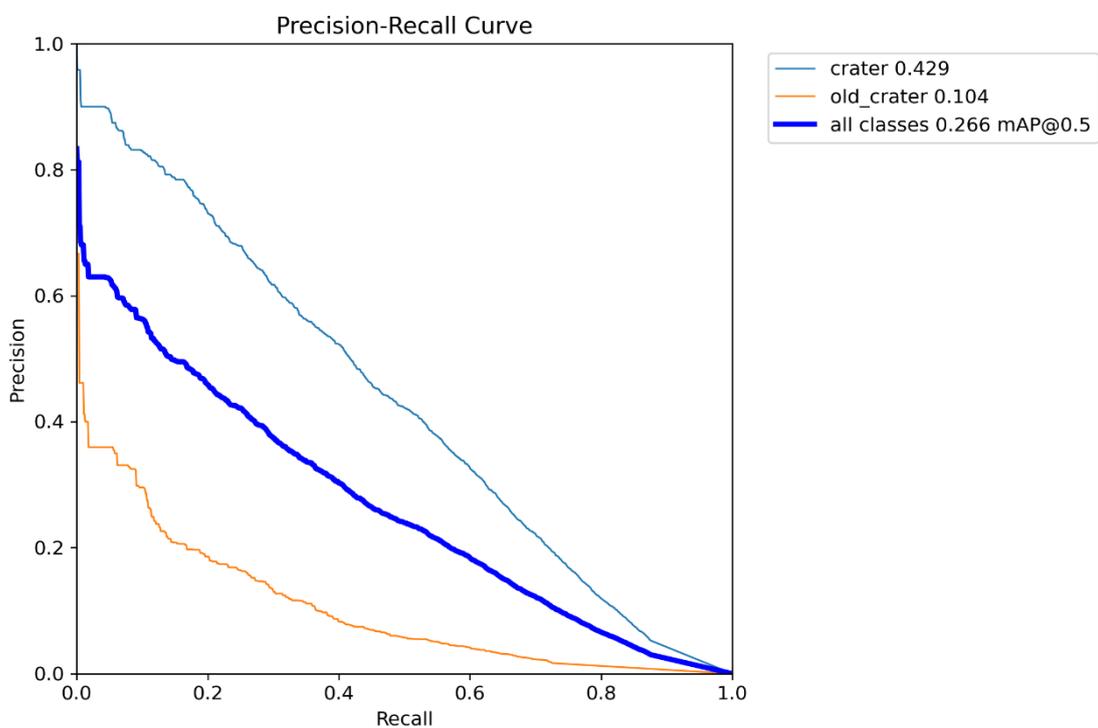


図 3.4 月モデルの PR 曲線

3つの結果を比べてみると、エンセラドスモデルの場合に各平均適合率、全クラス平均適合率（mAP）が最も大きい値となった。どの訓練データでも old_crater のクラスは値が小さく mAP を小さくしている原因となった。

また、物体検出後の画像をランダムで 10 枚抽出して参照し、クラスごとに適合率と再現率の計算を行った結果が表 3.2 および表 3.3 である。

表 3.2 ラベル「crater」の適合率・再現率

crater	エンセラドス・月	エンセラドス	月
Precision	0.6855	0.8736	0.5806
Recall	0.5166	0.3602	0.4265

表 3.3 ラベル「old_crater」の適合率・再現率

old_crater	エンセラドス・月	エンセラドス	月
Precision	0.4483	0.8824	0.6154
Recall	0.4643	0.5357	0.2857

3.3 mAP の比較

各訓練データについて、IoU が 0.5 および 0.5~0.95 の場合の mAP 値の推移を図 3.5 および図 3.6 に示す。

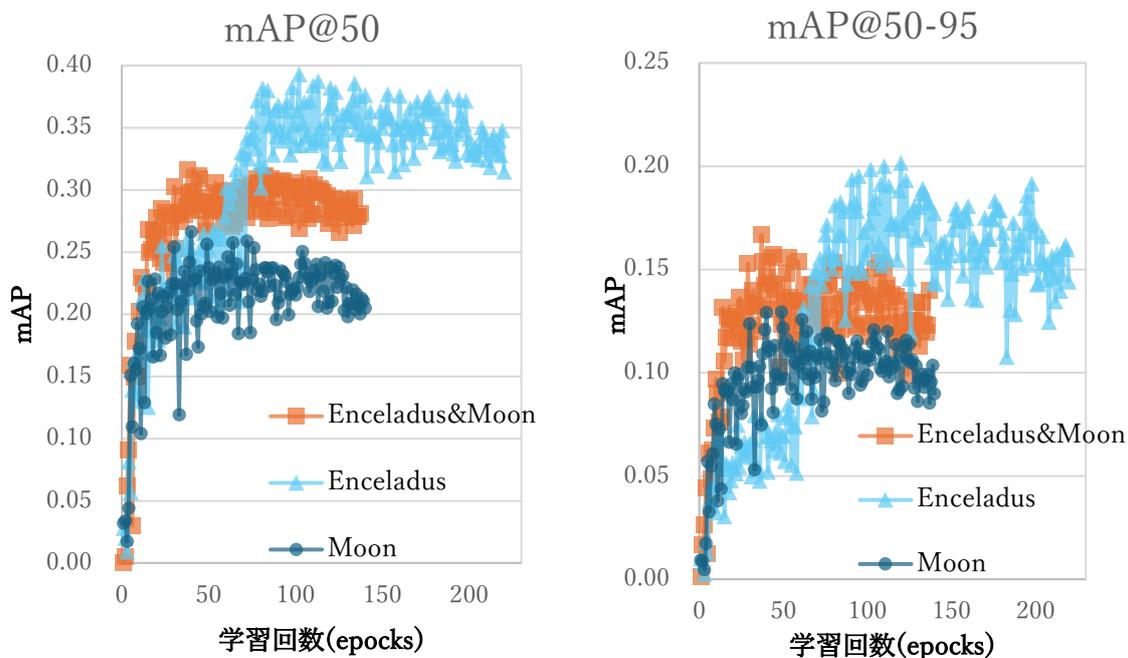


図 3.5 IoU=0.5 (左) と IoU=0.5~0.95 (右) のときの各訓練データの mAP。IoU の変化は検出の難易度に関わってくる。IoU が大きくなるほど検出の難易度は高くなるため、必然的に IoU=0.5~0.95 の mAP は低下する。

エンセラドスモデルの mAP がどちらの IoU 条件においても高い値をとることが分かる。エンセラドスモデルの mAP の最高値は IoU=0.5 のとき 0.39284、IoU=0.5~0.95 のとき 0.20159 となった。一般的に性能の良いとされるモデルは mAP@50 (IoU=0.5) で 0.8 程度であるので、今回学習したモデルはどれも性能が高いとは言えない。また、月モデルと月・エンセラドスモデルでは mAP のピーク値に達した後は緩やかに低下していく傾向が見られたが、エンセラドスモデルでは学習回数 50 付近でもう一度大きく mAP 値が上昇している。これは、前者二つの訓練データが 50 回以内の学習回数で早期終了に達したのに対し、エンセラドスモデルは学習の改善がみられる状態であったからだと考えられる。

3.4 モデルの決定と物体検出の結果

3.3 節までの結果から、最も性能の良いと考えられるエンセラドスモデルを用いて物体検出を行った。物体検出を行った元画像を図 3.6 および図 3.7 に示す。これらの画像は米国 NASA の PHOTOJOURNAL Image Galleries より取得した[6]。

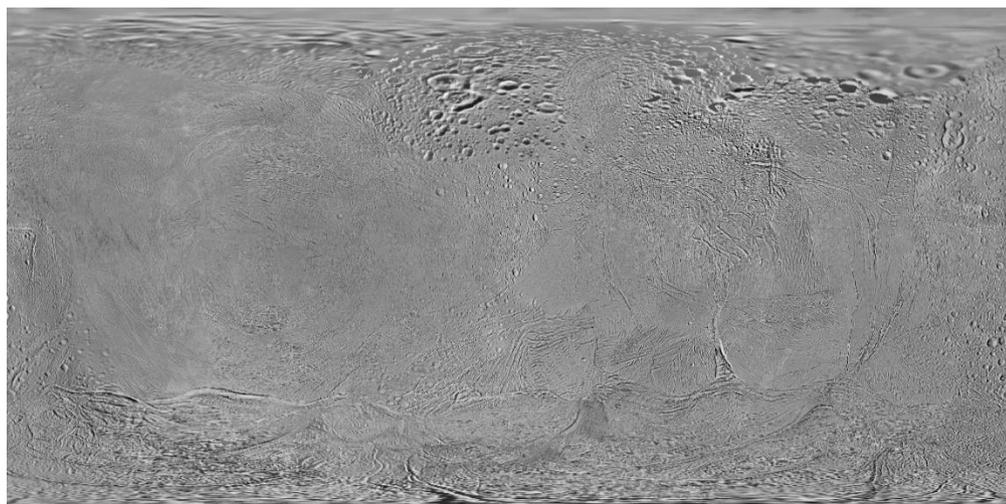


図 3.6 物体検出に使用したエンセラドスの全球画像 (PIA14937)。NASA の HP から取得したエンセラドスの正射投影図。実際の画像はさらに外枠があり、そこには緯度経度の座標情報が書かれており検出の障害となるため削除した。画像サイズが非常に大きい (7193×3586 pixel) ので横方向に 10 分割、縦方向に 20 分割した 360×359 pixel サイズの計 200 枚の画像に対して検出を行った。中心経度は 0° (土星方向) である。

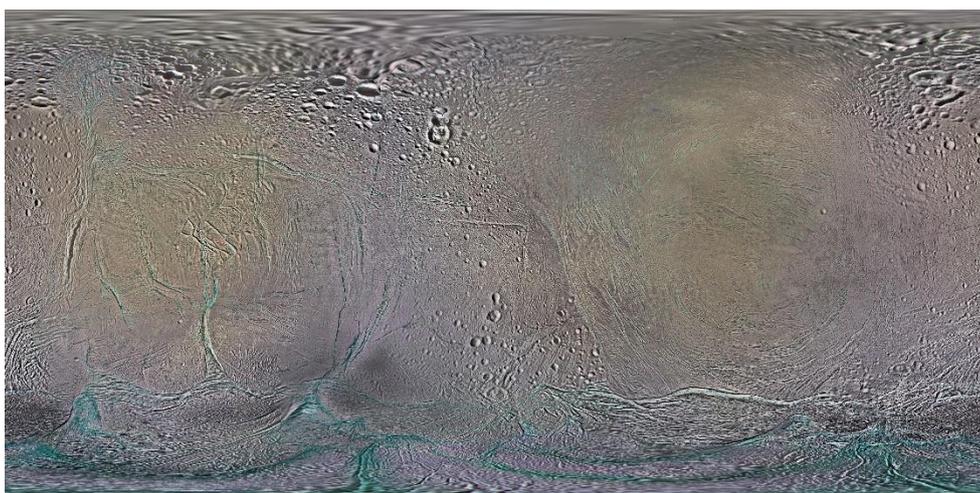


図 3.7 物体検出に使用したエンセラドスの全球画像 (PIA18435)。図 3.6 よりも解像度が高く (15960×7980 pixel)、中心経度が 180° E (反土星方向) となっている。

正射投影図の場合、緯度によってクレーターの形が大きく変わるので、赤道付近と高緯度地域に分けて検出結果を見ると図 3.8 に示すようになった。

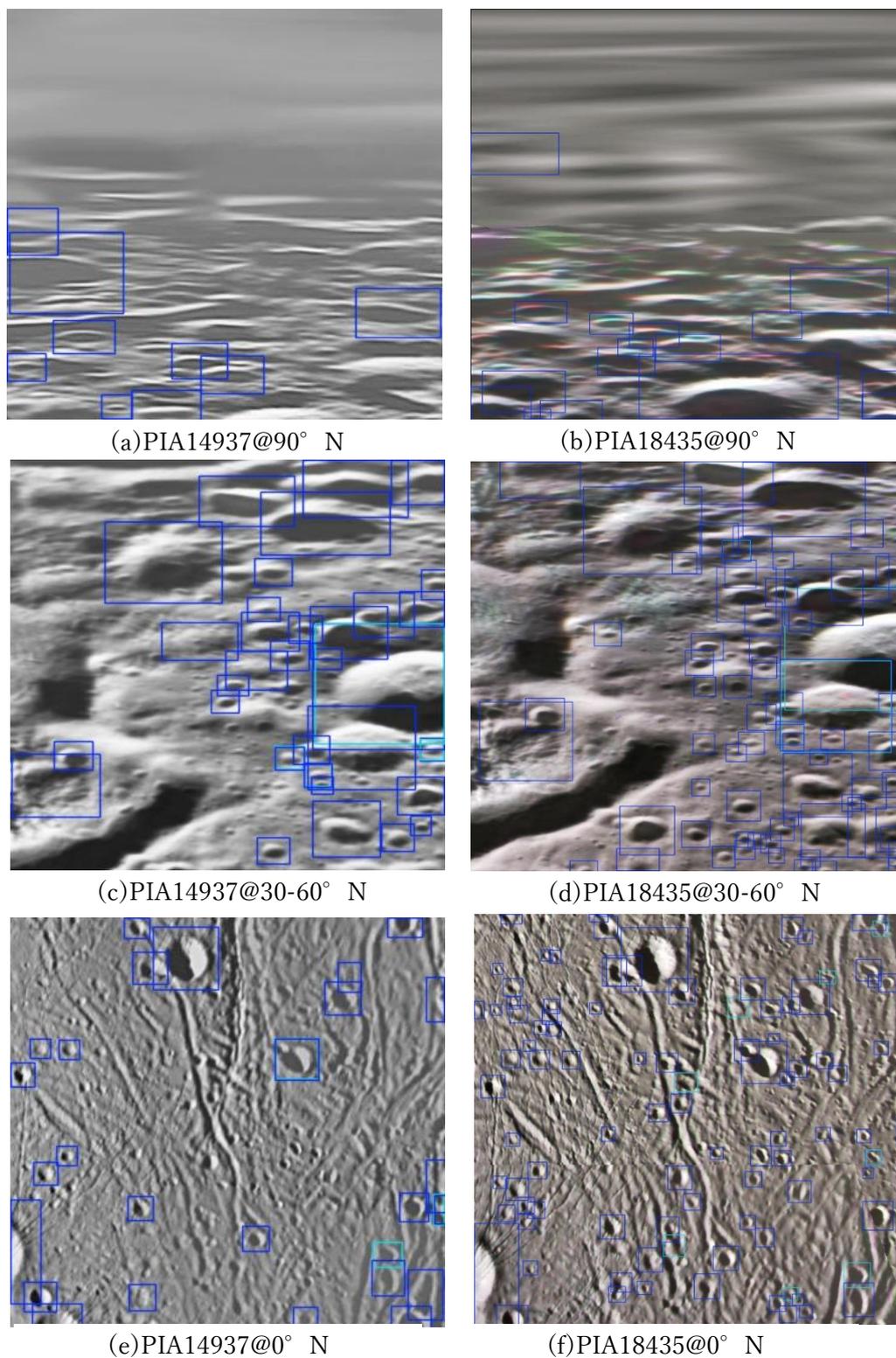


図 3.8 低解像度写真（左）と高解像度写真（右）の各緯度における検出結果の例

3.5 氷の亀裂に対するモデル精度

解像度の異なる2つの元画像（PIA14937、PIA18435）において、物体検出後の画像から、氷の亀裂による表面模様をクレーターと誤ってカウントした数（clacks）を手動で数え、全物体検出数と信頼度に応じて比較するヒストグラムを作成した。

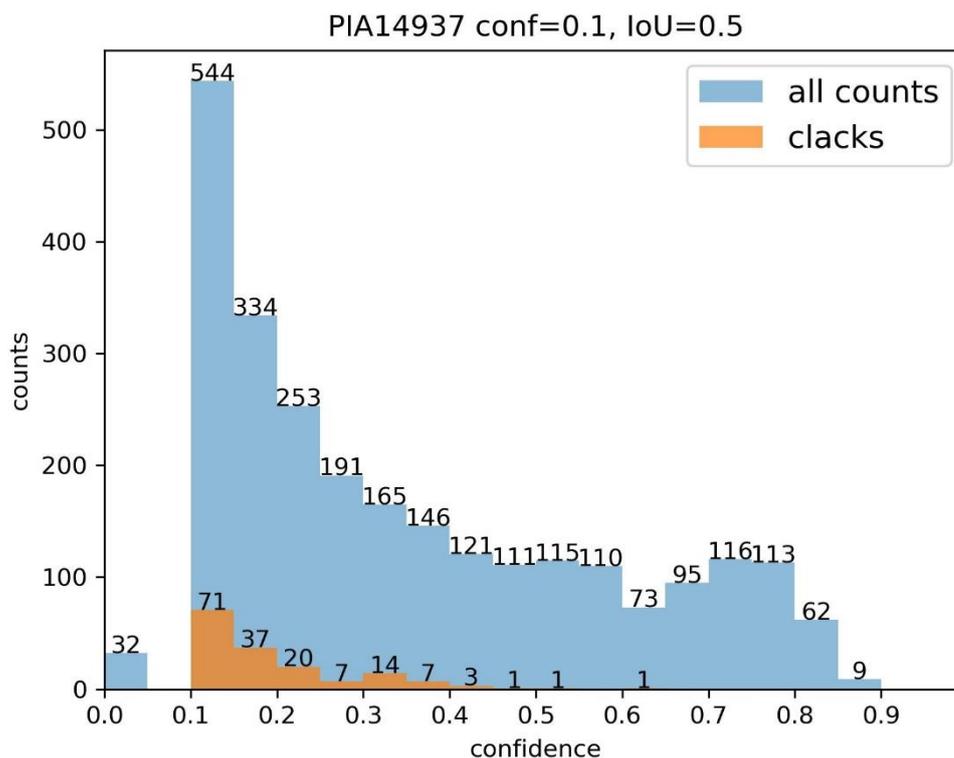


図 3.9 低解像度画像（PIA14937）を用いた場合の全物体検出数と氷の模様（氷の模様・亀裂の数）を信頼度 0.05 ごとに階級分けしたヒストグラム。各棒の上の数字はその階級において検出された物体の数を表している。物体検出の閾値を信頼度 0.1 以上、IoU 値 0.5 以上としているため信頼度 0.1 以下の検出は行われないが、例外として無検出であった画像には 0 の信頼度を入力して表示した。

低解像度画像（PIA14937）に対する結果を図 3.9 に示す。図 3.9 を見ると全物体検出数（all counts）に対して亀裂の数（clacks）の割合は小さいことが分かった。信頼度 0.0~1.0 における全物体検出数は 2558 であり、その内 162 個が氷の模様・亀裂に起因することが分かった。その割合は 6.33% であり、エンセラドス地球上の全クレーターのカウント数は 2396 となった。

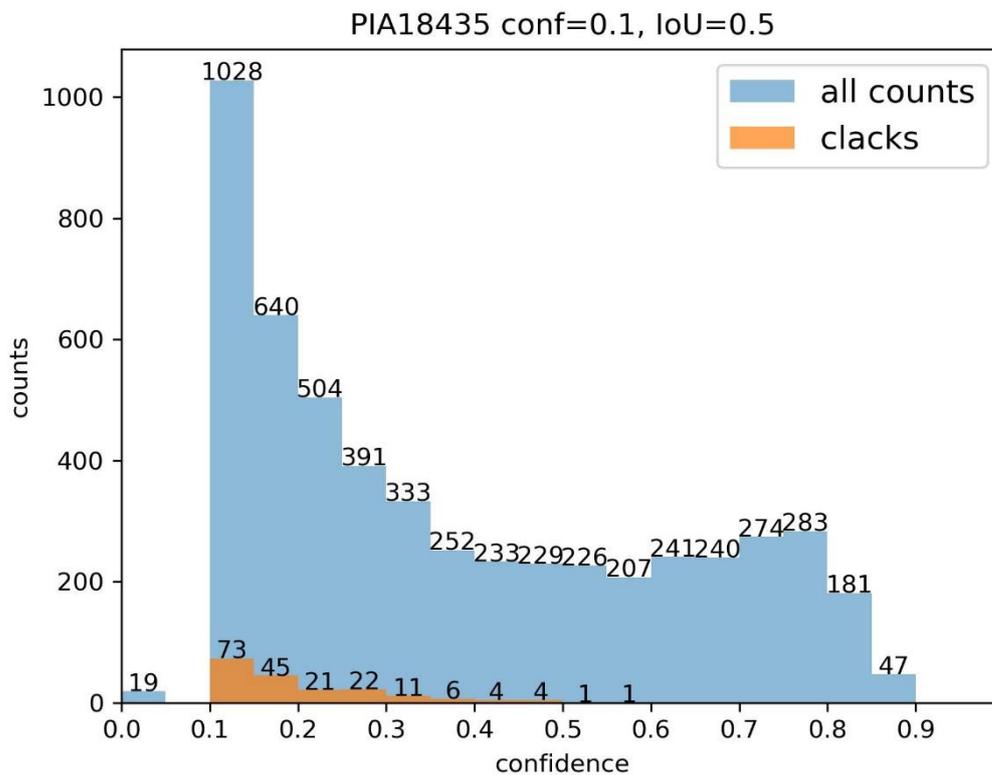
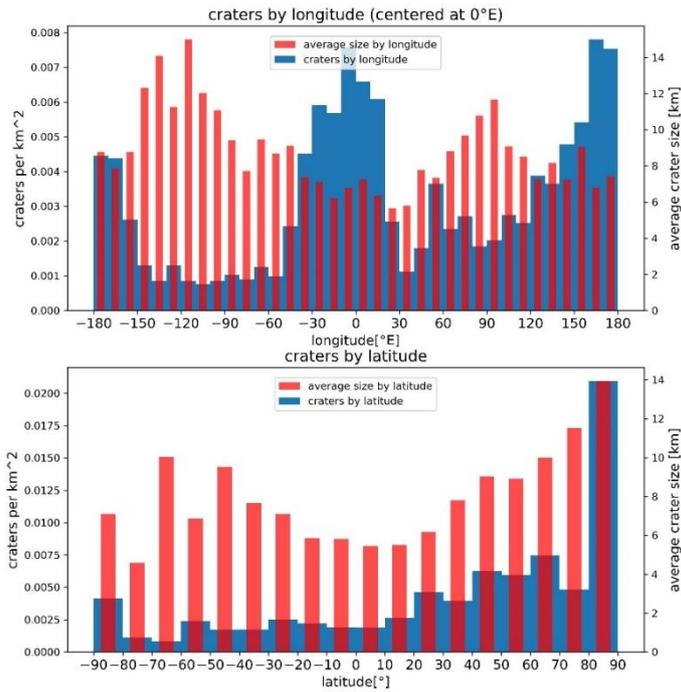


図 3.10 高解像度写真 (PIA18435) を用いた場合の全物体検出数と氷の模様の数との比較

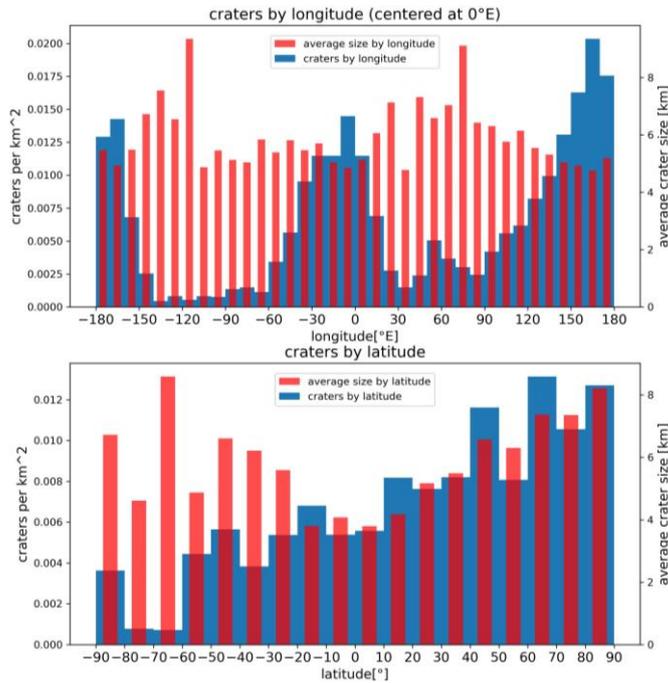
一方、高解像度写真 (PIA18435) を用いた場合の結果を図 3.10 に示す。図 3.10 より、全物体検出数 (all counts) は 5309、氷の亀裂の数 (clacks) は 188 であったのでエンセラドス全球上の全クレーターのカウント数は 5121 である。全物体検出数のうち氷の亀裂が占める割合は 3.54%であった。

3.6 クレーターの密度とサイズ

検出されたクレーターの密度とサイズのエンセラドス全球上での分布 (位置関係) は図 3.11 に示すようになった。



(a)低解像度 (PIA14937)



(b)高解像度 (PIA18435)

図 3.11 経度・緯度ごとのクレーターの平均直径と個数密度。(a)は地図の経度 0° 、緯度 0° を中心としてカウントした時の個数密度(青)とクレーターの直径(赤)を緯度・経度binごとの平均値として表している。(b)はより高画質な画像を用いて物体検出し、(a)と同様に表した図。

4. 考察・評価

4.1 モデルの性能

3章の結果をもとにモデルの性能を評価する。

4.1.1 損失における性能差

3.1節の結果では、3種類のモデルの各損失の差はあまり見られなかった。僅かではあるが月モデルはすべての損失が他のモデルよりも高かった。これは岩石で構成されている月の表面と氷で構成されているエンセラドスの表面とでクレーターを特徴づける要素が変わっているためと考えられる。

3.2節の結果では曲線が0から遠い（曲線で囲まれた積分値が大きい）ほど適合率Pと再現率Rが高いと言えるが、エンセラドスモデルが性能が一番高いと言える。再現率Rが0.1以下での適合率Pが大きいため予測した物体の正解率は非常に高いが、その内正解を検出した割合は低い（間違えた予測をしたがたまたま正解であった割合が高い）ので正しい予測ができているとは言い難い。逆に再現率Rが0.8弱までは適合率Pが0.2以上を保っているため、予測した物体が正解であった割合が比較的高く、その内正解を検出した割合も高いまま検出を行っていると考えられる。表3.2、表3.3を見ても適合率Pのほうが高く再現率Rはやや低い傾向にあると見られる。以上のことから、エンセラドスモデルは誤検出は少ないが見逃している正解Boxの存在も多くなる可能性のあるモデルであると考えられる。

エンセラドス・月モデルは表3.2、表3.3を見ると適合率P、再現率Rともにバランスよく高い傾向にあるとみられる。誤検出が少なく、かといって正解のBoxを見逃している数が多いわけでもないモデルであると考えられる。

月モデルに関してはold_craterの再現率が特に低いことが気になった。これは月の表面における風化具合がエンセラドスのそれと変わっているからだと考えられる。岩石である月の表面は大気もないため風化が遅く、古いクレーターの縁でもはっきり残っているものが多い。一方でエンセラドスの表面は氷でできており、熱変性を受けて岩石よりも劣化が早くなる傾向にある。そのためはっきりとしないクレーターが多く、月モデルのold_craterの特徴量と合わなかったのではないかと考えられる。

これらのことを踏まえると機械学習によるクレーターの検出では氷天体では氷天体の学習データ、岩石天体では岩石天体の学習データを使用する方がより物体検出の精度を上げられるのではないかと考えられる。逆説的に言えば、今回作成したエンセラドスモデルは月や他の岩石天体でのクレーター検出をすると精度が落ち、汎用的に使用できるモデルではない可能性がある。しかし、一方で氷天体特有の氷の模様がない分、誤検出の確率が低くなり、結果的に精度が良くなる可能性もある。これは他天体への適用など、今後の検証によってはっきりわかるものであると推測する。

いずれにしても今回のエンセラドスモデルでは、エンセラドス自身の大方のクレーター密度分布を作成できることが分かったので精度としては十分であると判断した。

4.1.2 氷の亀裂に対する信頼度の閾値

3.5 節の結果から全物体検出数に占める氷の模様の数割合は 6.33%、高解像度写真では 3.54%であることが分かった。階級ごとにみると信頼度 0.1~0.15 のとき、氷の模様・亀裂が最も多く存在しているがそれでも物体検出数は 544 であり、氷の模様が占める割合は 13.1%となった。閾値を 0.15 にすると全体の約 2 割のクレーターが失われるので得策とは言えない。氷の模様が多すぎてクレーターの検出が困難になるのであれば閾値が必要であると考えたが、全体に占める割合が少ないのと失われるクレーターの数を考慮すると閾値を設定しなくてもよいと考えられる。

4.1.3 解像度が及ぼす物体検出への影響

物体検出の際に使用した画像は 2 枚であり、そのうち一つはもう一つの画像よりも高画質であった。検出画像 1 枚当たりの大きさは PIA14937 が 360 × 359 pixel、PIA18435 が 798 × 798 pixel であり、1 辺当たりの長さは 2.2 倍、ピクセル数では約 5 倍の解像度である。物体検出の結果として、全クレーター数は PIA14937 が 2396、PIA18435 は 5121 となった。クレーター数としては 2 倍以上の差が出る結果となった。

解像度の高い画像のほうが細かいところまで見えることは当然であるが、これほど検出に差が出るとなると物体検出における画像の解像度は非常に重要であると考えられる。PIA14937 で検出した最小サイズのクレーター直径は 10.0 pixel であり、PIA18435 は 16.1 pixel となった。PIA14937 の赤道

上での単位ピクセルあたりの長さは 220 m/px、PIA18435 は 99.2 m/px なので、最小クレーターサイズはそれぞれ 2.20 km、1.60 km と算出され、PIA18435 の方がより小さいクレーターを検出できることが分かった。このことから、物体検出モデルが検出できる最小サイズが 10~16 pixel 程度であり、解像度に応じて検出精度が大きく変化することが考えられる。

図 3.11 のクレーター密度を見ると、高解像度の経度方向の最高クレーター密度は 0.0200 個/km²、低解像度は 0.008 個/km² 程度となった。さらに、経度方向でのクレーターの最大直径を見ると、高解像度では 9 km、低解像度では 15 km 程度となった。これらの結果は画像の解像度が最小クレーターを見つけ出す精度と誤検出の確率に大きく関与していることが考えられる。

4.1.4 先行研究との比較

先行研究では主に北緯 70° から南緯 40° の間で 5240 個のクレーターを特定した。その範囲外でのクレーター数はごく少数であった。本研究結果では誤検出も含めて全体で合計 5121 個のクレーターを特定した。

先行研究よりも正確さや学習データ数は劣っているものの、図 3.11 と図 4.1、図 4.2 のグラフを比べるとクレーターの分布からわかる傾向は経度方向では似たようなものとなったが、緯度方向では主に高緯度での違いが大きくなった。これは、本研究では高緯度での誤検出が多く、それを排除しない上での検出結果としてまとめたためこのような違いが生まれたと考えられる。研究結果としてこのような違いが大きくなるのであれば、明らかにミスと言える検出結果は手動で省いたうえで理学的議論に用いるべきであると考えられる。ただ、本研究では先行研究では行われていなかった北緯 70° 以上と南緯 40° 以上でのクレーターの数のカウントを行ったため、そこを含めたエンセラドスのクレーターに関する機械学習の議論を行えるという点では新しい発見となった。

一方で、40-50° N でのクレーター数が最も多いことや高緯度地域でのクレーター数が少ないことが分かる。経度に関しては本研究の結果を東経 0-360° で考えると先行研究の結果のグラフを 180° 回転させたような形になった。これはエンセラドスの地図の経度と先行研究の内容を見ると、先行研究の経度グラフが東経ではなく西経によるものである可能性が高いと考えられる。

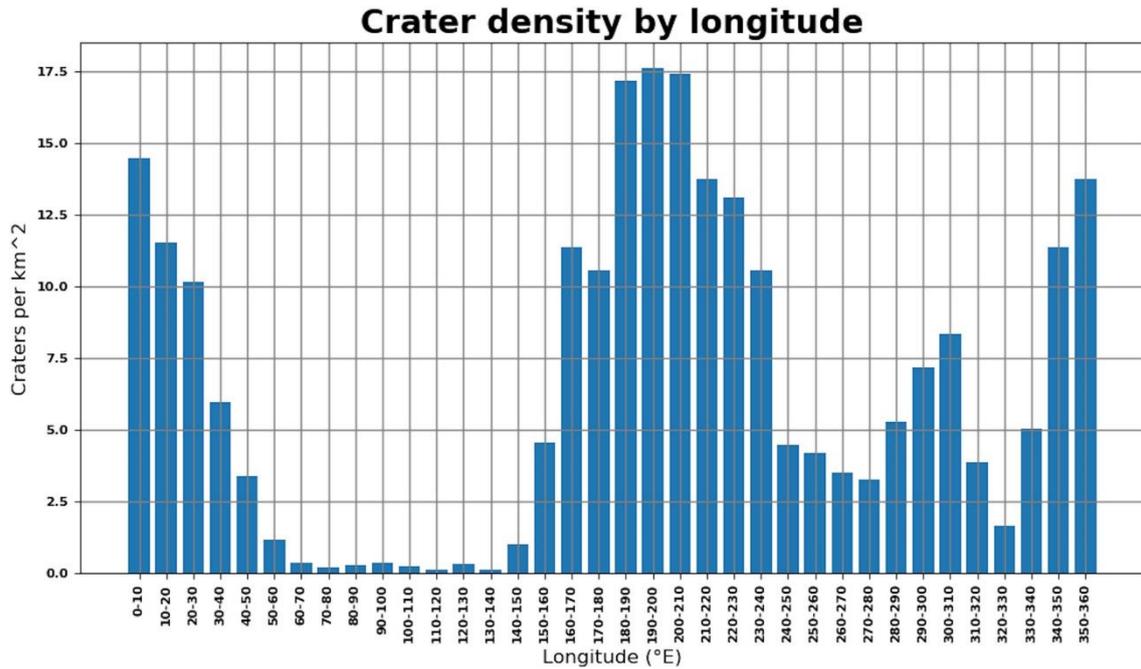


図 4.1 経度ごとの平均クレーター数 (Blanco-Rojas et al., 2024) [9]

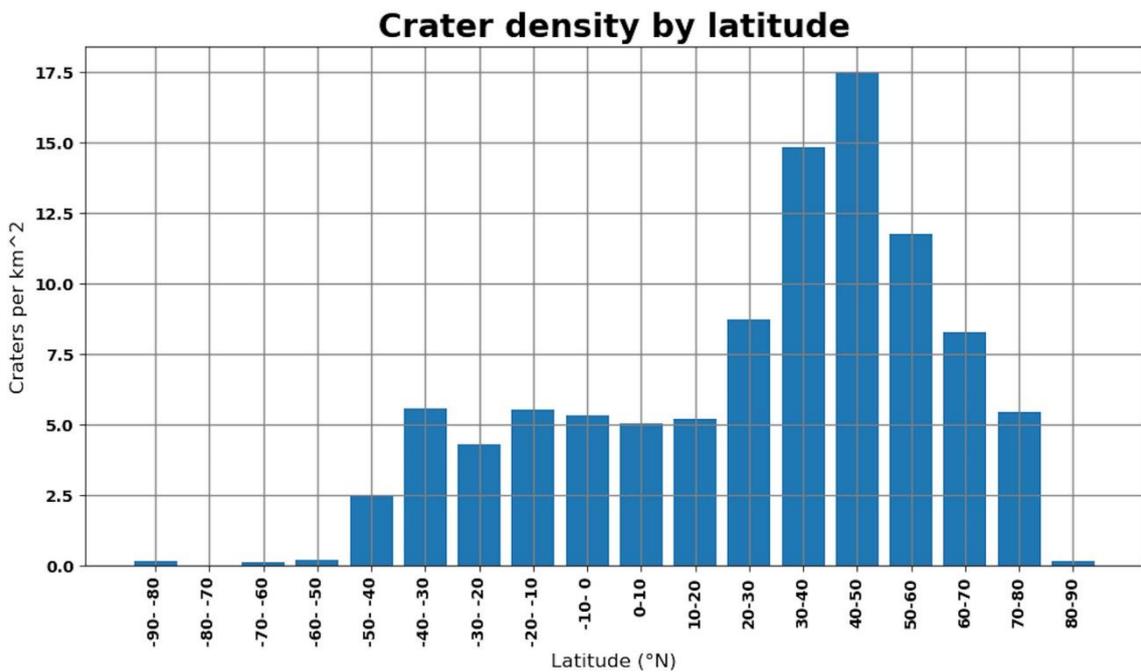


図 4.2 緯度ごとの平均クレーター数 (Blanco-Rojas et al., 2024) [9]

図 4.3 は経度 0° E (上) と 180° E (下) を中心としてエンセラドスのクレーターを検出し密度分布をモルワイデ図法で表したものである。検出させる画像の経度の中心軸を 180° 変えたと同じ元画像でもクレーターの密度

分布が変わった。上の図は下の図と比べると密度が小さくなっている。クレーターのカウント数はほぼ同じであるためこのように密度が変わるのは本研究の検出モデルが画像の端での検出を苦手としている可能性がある。先行研究では東経 0° 、北緯 30° でのクレーター密度は $0.025\sim 0.030$ craters/km² であり[9]、本研究では 0.0150 craters/km² 程度となった。これは先行研究で使用した学習モデル (Keras) がエンセラドスの経度 0° 付近でのクレーター検出が得意であった、もしくは YOLOv8 が苦手であったかのどちらかが考えられる。

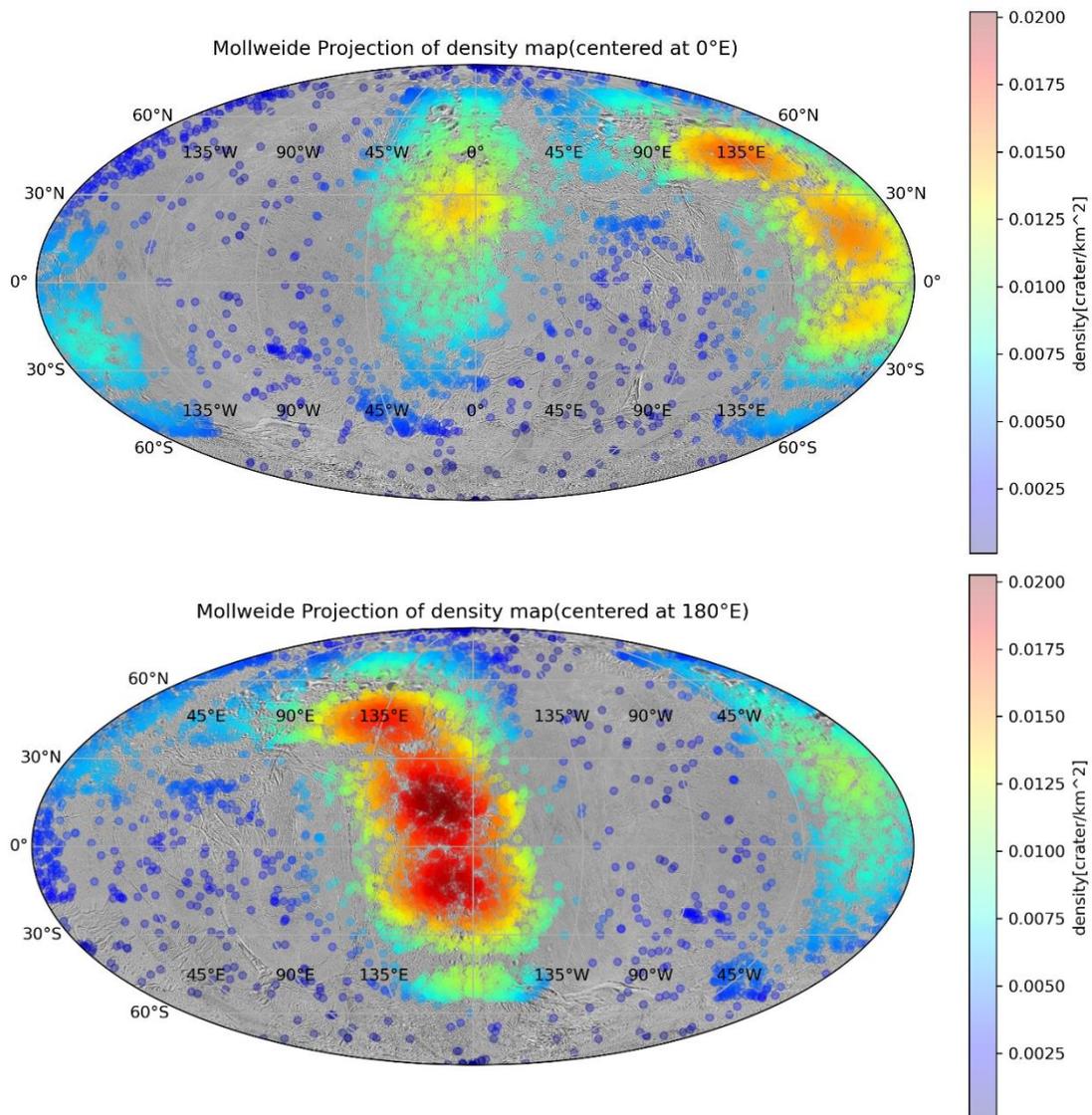


図 4.3 エンセラドスのクレーター密度分布

4.2 クレーターマップとエンセラドスの地形

エンセラドスモデルでエンセラドスの物体検出を行った結果を用いてエンセラドスのクレーターマップを作成した。赤道から中緯度付近までのクレーターはある程度検出することはできたが高緯度付近ではひしゃげたクレーターになってしまうので著しく精度が落ちた。よって、高緯度付近の分布は実際の衛星写真とよく見比べる必要がある。

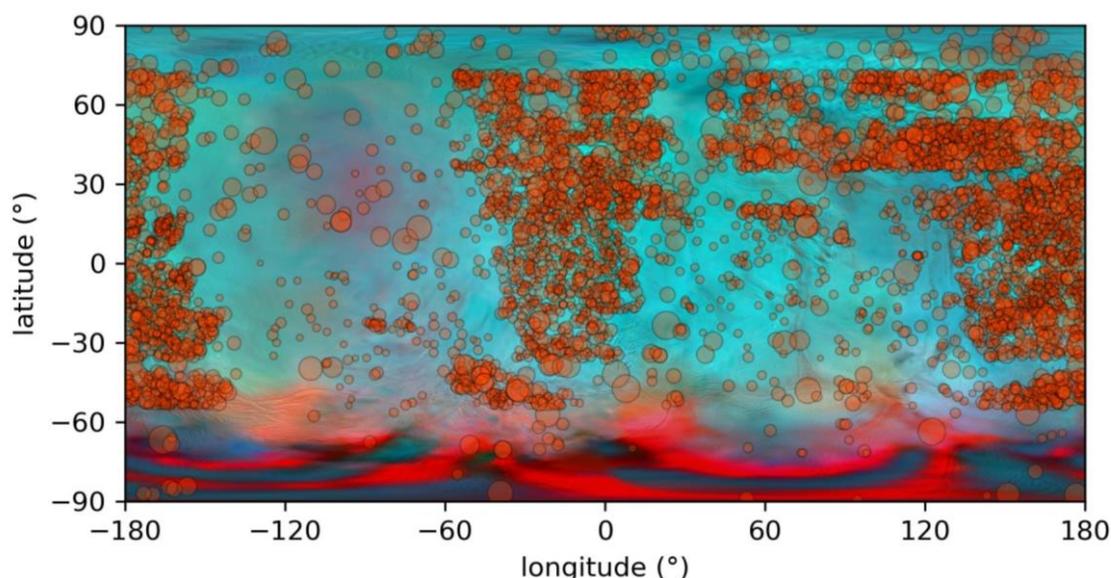


図 4.4 エンセラドスのクレーター分布とヒートマップ。赤外線センサーで計測されたエンセラドスの表面のヒートマップ (PIA24027) [6]の上に本研究で検出したクレーターの大きさと分布を表した図。正射投影図となっている。

エンセラドスは主星となる土星との間に潮汐ロックがかかっていることから、地球における月と同様に常に同じ面を土星に向けて公転している。その面は経度 0° の面であり、クレーターマップを見るとその面はクレーターが集中的に分布していることがわかる。ヒートマップではその地域は熱が少なく氷の流動も比較的穏やかであると考えられるため、この地域の氷は南半球よりも古いものであると考えられる。反対側の 180° の経度線に沿った帯上で最もクレーターの密度が高くなっており、クレーターの分布としてはやや東側に多くみられるのに対し、これらの間 ($30\text{--}120^\circ$ E) の赤道地域ではクレーターが極端に少なく、滑らかな氷の模様が広がっている。 $60\text{--}120^\circ$ W の緯度 60° N 地域はヒートマップではやや赤みを帯びており、熱を受けている地域であることがわかる。この辺りの氷は土星と垂直な面であることに加え、熱を多く受け氷が薄く、クレーターが残りやすいため他地域よりもクレー

ーターが少ないのではないかと考えられる。

南極域には「タイガーストライプ」(虎の模様のような構造)と呼ばれる巨大な亀裂があり(図 4.5)、そこからプルーム噴出が起きていることが確認されている[2]。巨大な亀裂が確認されていることとヒートマップの関係を見ると南極域の氷層が薄いことが伺える。南極域の熱量が非常に大きいことと極域に対するインパクトの起こりにくさを考えると、南極域にほとんどクレーターが存在しない地理的な理由になっていると考えられる。

以上のことを踏まえると潮汐ロックを受けている衛星の特徴とクレーターマップから見て取れることは一致している。赤道以南の地域にも少なからずクレーターが存在しているのを見るとやはり経度 110° E や 70° E のエリアのクレーターの少なさは不思議である。氷が流動して主にその地域を移動して北極まで運ばれているのなら些か早すぎるのではないかと感じるし、そうであるのならば最近ではクレーターがほとんど発生していないと考えられる。すべて北極とは言わず経度 0° と 180° の帯に拡散するように流されている可能性もある。今日の研究ではこの地域での大規模な再結成が行われたのではないかとされている (Mallory et al, 2024) [3]。東面と西面に位置するクレーターでは東経側のほうが多く高緯度に偏って分布している。また赤道付近にも大きな氷の亀裂が存在している。これはタイガーストライプにも似た特徴をもっているため過去のタイガーストライプである可能性が考えられる。

クレーターの大きさに注目すると、クレーターの密度が非常に高い地域ではクレーターの大きさが小さい傾向にあるのが分かる。北極域では密度は落ちるものの平均的なクレーターの大きさは大きいものが多いと言える。しかし、これは高緯度では精度が極端に落ちるクレーターマップから読み取るのは正確とは言えず実際の衛星写真や図 3.11 と比較しながら慎重に考察する必要がある。

図 3.11 (a) の経度方向のグラフではクレーター密度の高い地域では比較的クレーターのサイズが小さくなっていることが分かる。逆に、クレーター密度の低い地域では大きなクレーターが存在している。図 4.5 の北半球の写真を見ると確かに高緯度地域でのクレーターは目に見えて大きいものがあることがわかる。しかし、南極域でのクレーターは $80-90^{\circ}$ S に大きなクレーターは存在していないため、図 3.11 (a) の $80-90^{\circ}$ S のクレーター密度と直径は氷の模様・亀裂の誤検出である可能性

が高い。

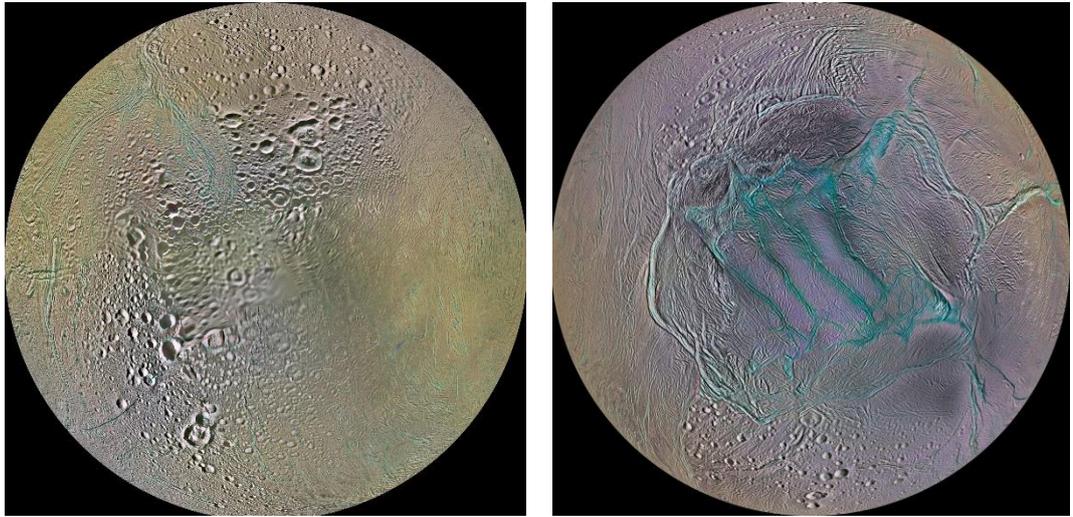


図 4.5 エンセラドスの北半球（左）と南半球の画像（右） [6]

5. 結論

土星の氷衛星であるエンセラドスのクレーターをカウントするため、月とエンセラドスの表面画像を用いて YOLOv8 に機械学習を行わせた結果、エンセラドスのみの学習データを用いて学習したモデルが最も性能の良いモデルとなった。クレーターマップを作成する際の精度としては十分ではあるが、厳密なクレーターの数やサイズを特定するに至るまでの精度は得られなかった。また、高緯度域での検出は不安定であるため、誤検出を防ぐために予め高緯度域を省くか別の事前処理を行うか、手動での正射投影図により大きく引き延ばされている誤検出データの排除が必要であることが分かった。検出できた最小のクレーターサイズは 1.60km であったため、これ以下のクレーターは検出できずカウントされておらず、実際のクレーター数は本研究の結果よりも遥かに多いと考えられる。また、検出に用いる元画像の解像度による影響が大きく、縦、横それぞれ 2 倍程度の元画像に適用した比較ではクレーターの検出数には倍以上の差が出ることから検出の際に用いる画像は高画質である方が良いと考えられる。

謝辞

本研究のテーマがこれまでの宇宙地球探査システム研究室における研究テーマになかった新規テーマの選定となったにも関わらず、たくさんの助言とご指導をしてくださった指導教員である高知工科大学システム工学群 山本真行 教授に心から感謝申し上げます。

本研究の資料作成、惑星地質学的な観点からのご指導を賜りました高知工科大学システム工学群 西川泰弘 特任助教に深く感謝申し上げます。

そして、同じ研究室に所属する皆様方には日々様々なサポートをしていただいたこと深く感謝いたします。

最後に、筆者の日々の生活の支えになってくれた家族や友人、親戚一同に感謝いたします。

参考文献

- [1]NASA/JPL-Caltech/Space Science Institute, “PIA17182: A Snowball in Space,” <https://photojournal.jpl.nasa.gov/catalog/PIA17182>, 2013/12/23 (投稿日)
- [2]NASA/JPL-Caltech/Space Science Institute, ” PIA17184: Encroaching Shadow”, <https://photojournal.jpl.nasa.gov/catalog/PIA17184>, 2014/07/28 (投稿日)
- [3]Frank Postberg et al., “Detection of phosphates originating from Enceladus’s ocean,” Nature, Vol 618, pp.489-493, 14 June 2023, DOI:10.1038/s41586-023-05987-9
- [4]Blanco-Rojas et al., “A Novel Approach to Impact Crater Mapping and Analysis on Enceladus, Using Machine Learning,” JGR Planets, Vol 129, 13 February 2024, Issue 2, DOI: 10.1029/2023JE008010
- [5]Glenn Jocher, “Ultralytics YOLOv8,” https://docs.ultralytics.com/ja/models/yolov8/#__tabbed_1_1, (2025/02/06 閲覧)
- [6]NASA, “PHOTOJOURNAL Image Galleries,” <https://photojournal.jpl.nasa.gov/target/enceladus>, (2025/02/07 閲覧)
- [7]Koga Takahiro, “【YOLO】 labelImg でアノテーション済みデータを増殖(data augmentation),” <https://qiita.com/garcoo/items/2ab8972ce304ffc8d225#%E3%82%A2%E3%83%8E%E3%83%86%E3%83%BC%E3%82%B7%E3%83%A7%E3%83%B3%E5%BE%8C%E3%83%87%E3%83%BC%E3%82%BF%E3%81%AE%E5%A2%97%E6%AE%96>, (2025/02/06 閲覧 data augmentation)
- [8]g-k, “ROC 曲線と PR 曲線-分類性能の評価方法を理解する②-,” <https://qiita.com/g-k/items/b47b9b0ee2015a3b0b94>, 2025/02/06 閲覧)
- [9]Blanco-Rojas et al., “A Novel Approach to Impact Crater Mapping and Analysis on Enceladus, Using Machine Learning,” JGR Planets, Vol 129, Issue 2, pp.6, 13 February 2024, DOI: 10.1029/2023JE008010