教師なしドメイン適応を用いた粒子形状評価

顧　　席銘† 窪内　将隆†† 山本　岳洋††† 大島　裕明†,†††

† 兵庫県立大学 大学院応用情報科学研究科　〒 650–0047 兵庫県神戸市中央区港島南町 7–1–28
†† 堺化学工業株式会社　〒 590–8502 堺市堺区戎島町 5–2
††† 兵庫県立大学 社会情報科学部　〒 651–2197 神戸市西区学園西町 8–2–1
E-mail: †{aa19z502,ohshima}@ai.u-hyogo.ac.jp, ††kubouchi-m@sakai-chem.co.jp,
†††t.yamamoto@sis.u-hyogo.ac.jp

あらまし　本研究では，粒子の顕微鏡写真画像を入力として，粒子形状の評価を自動的に行う手法を提案する．粉末状の化学製品などにおいて，製品の品質として，粒子の形状がどのようになっているかや，規定外の大きさの粒子が存在していないかといったことが問題となる．実際の現場では人手による評価が行われていることもあり，自動的に粒子形状の評価を行うことが求められている．本研究では，インスタンスセグメンテーション技術を用いて，この問題に取り組む．粒子の顕微鏡写真画像は容易に大量に作成することができる．一方，粒子のセグメンテーションのラベル付けには非常に大きなコストがかかる．そこで，本研究では，ドメイン適応を行うことで，比較的少ないラベル付けデータを用いて粒子のセグメンテーションを精度良く行い，その結果を分析することで粒子形状の評価を行う．

キーワード　インスタンスセグメンテーション，ディープラーニング，領域適応

1 はじめに

産業界における画像処理技術の応用範囲は極めて広い．品質管理が厳しい製造企業では, 機械化による省力化や検査基準の安定化のために画像処理技術が使われている．化学製品の検査, すなわち，粒子の検査では，目視によるの検査には一定の休息期間が必要である．専門家と初心者の単純検査能力には有意差は見られないが，専門知識を必要とするような検査項目についての比較ではその検出率に差が見られる．

一方で，機械学習手法の発達によって，ラベル付きデータが大量に存在すれば予測器を構築することができる．予測器の正解率を向上させるためには，大量のデータを用意する必要がある．データにラベルを付けてもらうのはコストが高いことである．教師なしドメイン適応とは，ラベル付きデータとラベルなしデータの両方を利用して，教師あり学習や教師なし学習よりもよい結果を出すために用いられる．教師なしドメイン適応が注目される動機として，適用可能なデータが多く得られ，かつ精度の向上が期待できることがあげられる．実際，現実世界で得られるデータでは，ラベルつきデータが少ないが，一方でラベル無しデータは簡単に，大量に得ることができる．

粒子の画像は特徴として，密集や形の違い，互いに重なるという三つ特徴がある．図 1 に示すように，画像セグメンテーションの処理は難しいものである．化学会社は新製品を開発する場合，新製品には新しい形がある．新製品が開発されるたびに対応するラベル付きデータが作成されるので，非常にコストが高い作業がかかる．本研究は将来的に，新しい形状の粒子の画像をラベルなしデータとしてモデルを直接訓練して，モデルを新しい形状に適応させることが可能であると考えられるので，長期的な実用性と使いやすさを備えていると考えられている．

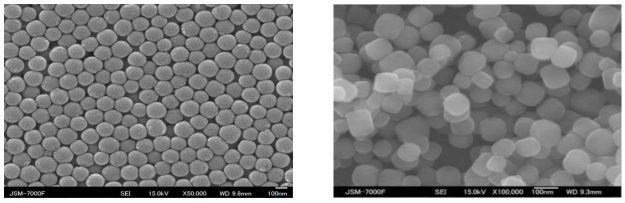


図 1　粒子画像の例

そこで本論文ではラベルなしデータを活用して，粒子のセグメンテーションモデルをより多くの形の粒子に適応させるために，ADDA [2] というドメイン適応手法を使用する．

本研究の構成は以下の通りである．2 節ではニューラルネットワークや Data Augmentation などに関する関連研究を述べる．3 節ではデータセットの作成と問題構造ついて説明する．4 節では粒子のセグメンテーションついて説明する．5 節では実験結果を報告し，最後にまとめと今後の課題について述べる予定である．

2 関連研究

2.1 ニューラルネットワークに関する研究

近年，画像認識においては畳みこみニューラルネットワークを使った手法の実績が顕著である．ResNet と呼ばれる手法では 2 層を学習するごとに出力と 2 層前の入力を足し合わせる手法を取り入れることによって多層化を行っても良い結果を得ることができるようになっている [5]．ResNet では 2 層を学習するごとに出力と 2 層前の入力を足し合わせる手法を取り入れることによって多層化を行っても勾配消失問題が生じなくなり良い結果を得ることができるようになっている．精度が劇的に良くなるわけではないが，ResNet を使えば 1000 層以上の多層化

を行ったうえで計算を行うことも可能である．このような素晴らしい層が作られているが，一方でディープラーニングには学習に必要なデータ数が膨大に必要であるという弱点が存在している．ResNet はラベル付きの画像を学習することで高い精度を得ることできるが，画像データをスカラー値として入力するため，同じ画像でも少しゆがめただけで別のラベルの画像として判別してしまうという弱点が存在する．

2.2 Data Augmentation に関する研究

同じ画像でも少しゆがめただけで別のラベルの画像として判別してしまうという弱点に対応するために Data Augmentation という手法が用いられている [8]．Data Augmentation は画像に回転やズームなどを加えることによって画像の枚数を増やす手法を示している．画像の枚数を増やすことで少しゆがめただけの画像にも対応できるようになる．また，mixup と呼ばれる Data Augmentation の手法を用いている [17]．この手法は二つの画像を混ぜることによって一つの画像を生成し，学習用画像を生成する手法である．混ぜた画像のラベルは二つの画像をどの割合で混ぜたかによって決まる．たとえば，A の画像と B の画像を 0.2 と 0.8 の割合で混ぜた場合，ラベルも A の画像のラベルも 0.2 の割合とし，B の画像のラベルも 0.8 とすることで学習させる．このように盛んにディープラーニング手法の研究や環境の整備が行われたことにより様々な分野からディープラーニングの応用が注目されている [6], [7], [15].

2.3 半教師あり学習に関する研究

半教師あり学習はデータの取得が容易で，モデルのパフォーマンスを効果的に向上させることができるため，近年，多くの半教師あり学習方法が研究されている．例えば，疑似ラベル [10] はディープニューラルネットワークの半教師あり学習のシンプルで効率的な方法である．この手法では，まず，ラベルがつけられている画像を用いて分類モデルを構築する．その後，ラベルがつけられていない画像の分類を行い，そのうち，確信度高く分類できた画像を選択する．もともとラベルがつけられている画像と選択された画像を用いて，新たに分類モデルを構築するということを行う．Fixmatch [16] は疑似ラベルと Consistency Regularization [9], [14] を合わせた手法である．Consistency Regularization とはラベルなしデータの画像にノイズを加えても，モデルの出力値が変わらないようにする方法である．FixMatch では，まず，左右反転等の弱い変換を与えたラベルなし画像を学習中のモデルで予測値を計算する．そして，一番確信度が高いラベルの予測値がある閾値を超えたもののみを，ラベルを 1，それ以外を 0 にする one-hot 形式のラベルにして，それを擬似ラベルとして扱う．次に，同じ画像に対して複数種類 Data Augmentation 組み合わせる等の強い変換をかけた画像を，同じモデルで予測値を出す．最後にその予測値と，先ほどの擬似ラベルを正解として Cross Entropy をかけることで，Consistency Regularization の制約をかける．

2.4 データ不足に関する研究

画像認識に関連して，様々な大規模データセットが存在している．例えば，ImageNet [1] というデータセットがある．ImageNet における約 120 万枚の画像に対して，各画像にどのようなオブジェクトが写っているかラベル付けされたデータセットが存在する．そこで対象とされているオブジェクトの種類は 1,000 種類である．それらの 1,000 種類のオブジェクトに対しては，最新の深層学習技術を用いることで，高い認識精度を実現することが可能である．これは，ひとえにデータ量が十分に存在するためであるということができる．しかし，社会の実問題においては，それら 1,000 種類のオブジェクト以外においても認識したいということがあるだろう．

そのような場合に用いられる技術は，いくつか存在する．一つは，転移学習 [12] という手法である．転移学習では，まず，ImageNet のデータセットを用いて 1,000 種類のオブジェクトを認識するようにモデルの学習を行う．そのようにして用意されたモデルを，事前学習モデルと呼ぶ．事前学習モデルは，大きく，画像から特徴抽出を行う部分と，得られた特徴をもとにして分類を行う部分に分けられる．転移学習では，事前学習モデルから分類を行う部分を取り除き，分類を行うニューラルネットワークを新規で追加し，自分が分類したいデータを用いて，その分類を行うニューラルネットワークの部分だけを新たに学習するということを行う．この場合，特徴抽出を行う部分については，一般的な画像ですでに学習が行われているものをそのまま利用するということになる．

ドメイン適応は転移学習の一種であり，ドメインは違うものの上でタスクは同じのケースを指す．つまり，ソースドメインとターゲットドメインのデータは同じであり，特徴分布が違う場合に有効な手法である．Ganin らは DANN [3] というニューラルネットワークを提案したした．この提案はドメイン適応を方針で開発された．主に対抗的なニューラルネットワークとドメイン適応を合わせる．対抗的な方法を使用して，ネットワークを新しいドメインに適合させる．訓練に連れって，抽出された特徴は同じ分布になる．

3 問題定義

本節では，実験で用いるデータである粒子の説明をおこない，データ作成について述べる．また，その際の粒子のセグメンテーション問題や粒子形状の評価問題についても説明する．

3.1 粒子の写真

本研究で使用する粒子の写真は顕微鏡で拡大された化学製品の写真である．製品の違いによって粒子の形も異なる．写真は粒子であるかどうか判別する際に用いるため，識別するオブジェクトの種類は 1 種類のみである．図 1 に示すように，ラベル付きの粒子の写真は様々な形がある．

3.2 データと問題構造

本研究で用いる画像データは，堺化学に提供されたデータである．その中で，本研究で利用したラベル付き画像の枚数を表 1 に掲載する．粒子の種類は形における名前を付ける．「不規則な円形」粒子や「四角」粒子や「円形」粒子は図 2 で示す．

表 1　粒子の種類とデータにおける写真枚数

種類	ラベル付き	ラベルなし
「不規則な円形」	4	103
「四角」	74	84
「円形」	22	45
人工データ	35	0
新形粒子	0	158

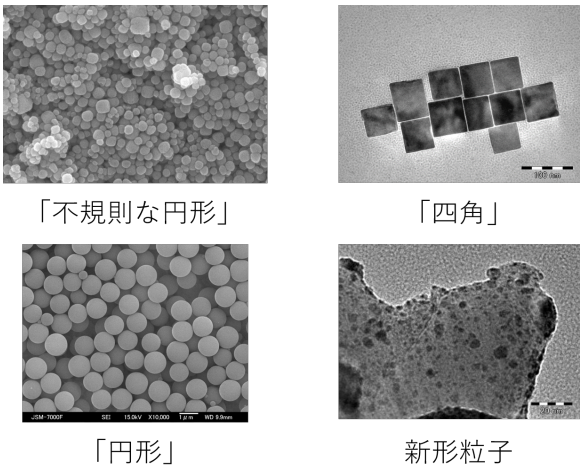


図 2　画 像 の 例

本研究では，ラベル付きデータを，訓練データとテストデータに分割した．比率は，概ね 9:1 とした．訓練データは，モデルの構築に使用する．テストデータは，学習されたモデルを評価するために使用する．

まず，粒子形状の評価の全体流れ（図 3）を説明する．入力は 1 枚の画像である．入力された画像は，はじめに粒子のセグメンテーションを行う．この時，出力は写っている粒子にマスクをつけられた写真である．次に，マスクを付けた粒子の形状の評価を行う．評価指標は円相当径，円形度，絶対最大長，対角幅，面積，球体積，正方形相当径，四角度の 8 種類で構成される．この 8 種類の指標によって，粒子の形状が良いかどうかを判断する．

ここでは，セグメンテーション問題と，粒子形状に評価問題という 2 つ問題に取り込むをこととなる．

3.3　粒子のセグメンテーション問題

粒子のセグメンテーションは写真に写っている粒子を一個ずつ識別する必要がある．そのために，本研究はインスタンスセグメンテーション技術を使用する．インスタンスセグメンテーションは物体の領域をオブジェクトごとにピクセルレベルで検出および分類する手法である．本研究は粒子を一個ずつ識別出来る後マスクを付けられる．分類する時粒子という 1 つ種類のみがある．

3.4　モデル性能の評価指標

本研究では複数モデルを構築した．お互いにセグメンテーション性能を向上するかどうかを評価するために，Faster R-CNN [13]，Mask R-CNN，COCO などのオブジェクト検出モデルを評価するために使用される最も一般的な方法を使用する．評価指標は 6 つがある（表 2）．AP，$AP_{IOU=0.5}$，$AP_{IOU=0.75}$，AP_{small}，AP_{medium}，AP_{large} である．この AP は IoU がベースとなっている．IoU とは下の式（1）のように，正解の領域と予測の領域の面積を 2 つの領域の共通部分の面積で割った値である．

表 2　評価指標リスト

評価指標	意味
AP	平均精度
$AP_{IOU=0.5}$	IoU>0.5 時の AP
$AP_{IOU=0.75}$	IoU>0.75 時の AP
AP_{small}	面積<32 × 32 ピクセル時の AP
AP_{medium}	32 × 32 ピクセル<面積<96 × 96 ピクセル時の AP
AP_{large}	面積>96 × 96 ピクセル時の AP

$$IoU = \frac{\text{予測の領域} \cap \text{正解の領域}}{\text{予測の領域} \cup \text{正解の領域}} \tag{1}$$

3.5　粒子形状の評価問題

粒子形状の評価はセグメンテーションモデルから出力の写真によって，粒子の 8 種類のパラメータを検出する．次に，平均値と標準偏差を計算して，1 粒子ごとの形状を評価する．

4　粒子のセグメンテーション

本節はまず提案アプローチを説明する．次に，インスタンスセグメンテーションの手法を説明する．そして，画像の事前処理する時，どのようなの Data Augmentation 手法を使用するかを説明する．最後，モデルの汎化性能を向上するため，訓練データに人工データを追加することを説明する．

4.1　提案アプローチ

本節は提案アプローチを紹介する．全体構造は図 5 のようになっている．本研究は Mask R-CNN を 2 つの部分に分割する．一部分は特徴抽出器である．もう一部分は予測器である．特徴抽出器は入力された画像から粒子の特徴を抽出する．この時 1 枚画像に対して，写真の中に粒子が存在する可能性のある領域の特徴を出力するので，出力は複数特徴マップである．予測器は出力されたいろいろな特徴によって，該当粒子の種類の分類，boundingbox の推定，マスクの予測を行う．

ドメイン適応はドメインとターゲットがある．本研究のドメインはラベル付きデータである．ターゲットはラベルなしデータである．提案手法では，まず，事前学習を行う．事前学習はラベルがつけられている粒子画像を用いてモデルを構築する．この時 Mask R-CNN 特徴抽出器はソース特徴抽出器と呼ぶこととする．

次に，敵対適応を行う．この部分は 2 つの特徴抽出器を使用する．1 つはソース特徴抽出器で，もう 1 つはターゲット特徴抽出器である．このソース特徴抽出器は（図 9）Step1 を訓練したのソース特徴抽出器である．ターゲット特徴抽出器はソー

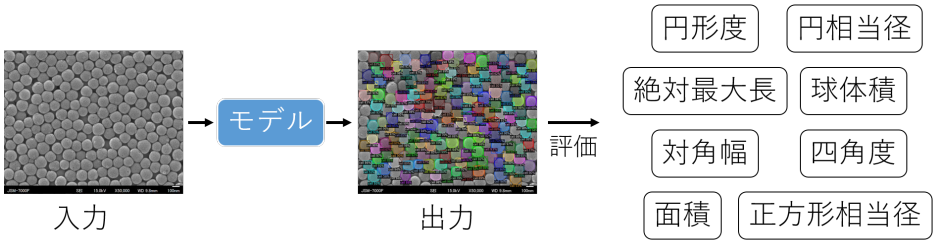


図 3　形状評価の全体流れ

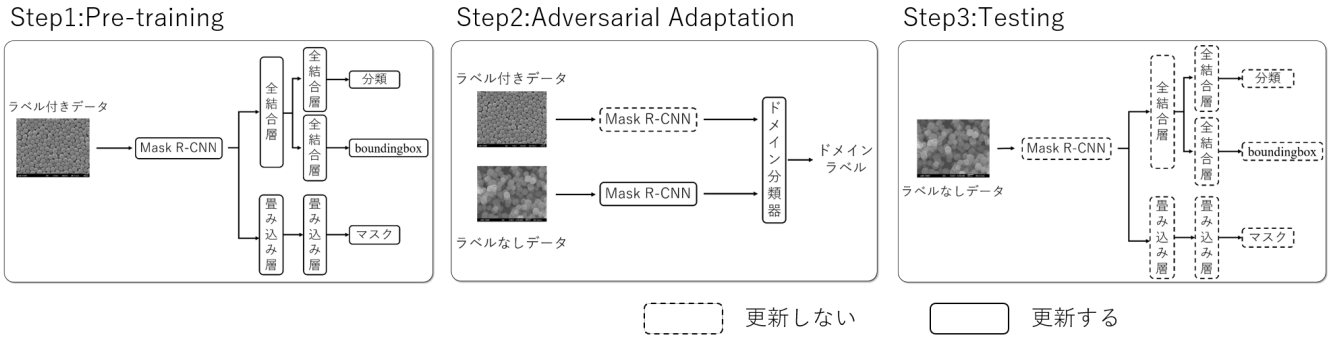


図 4　提案アプローチの構造

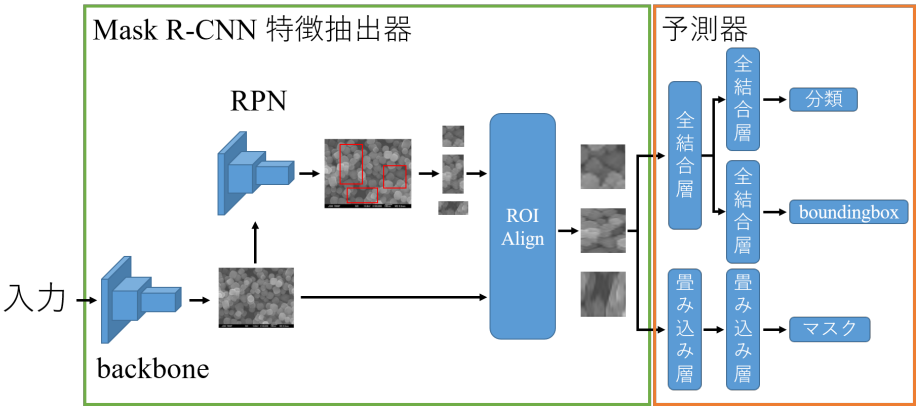


図 5　Mask R-CNN の構造

ス特徴抽出器と同じ構造を使用する．ターゲット特徴抽出器はソース特徴抽出器の最終パラメータを初期値として使用する．訓練する時，ソース特徴抽出器はフリーズする．ターゲット特徴抽出器は訓練するにつれてパラメータを更新する．この部分がドメイン分類器を構築する．ドメイン分類器は入力された画像がドメインかターゲットかを判断する．すなわち，入力された画像はラベル付き画像かどうかを判断する．この部分の損失関数の設定は通常の機械の損失関数と異なる場合がある．この部分は敵対的な訓練を行うので，損失関数は逆に設定する．すなわち，もしドメイン分類器はラベル付きかどうかを分けることができれば，損失値が大きくなる．分類できなければ，損失値は小さくなる．分類できる場合，2 つの特徴抽出器は違う特徴を抽出する．この場合，2 つのドメインの共通の特徴が効果的に抽出されていないことを意味する．

最後に，(図 9) Step2 を訓練したターゲット特徴抽出器を使用して，テストデータの予測を行う．特徴抽出された後，Step1 を訓練した予測器に入力する．こと時，予測を行うので，全体のパラメータは更新しない．

本研究の提案手法はまずラベル付きデータによって，ソース特徴抽出器を訓練する．次に，ソース特徴抽出器とターゲット特徴抽出器を抵抗的な訓練を行う．これにより，ターゲットドメイン特徴抽出器は，新しい形状の粒子画像のどの特徴が予測器の結果に効果的であるかを知ることができる．最後に，ラベル付きデータで訓練された予測器を使用して，ターゲット特徴抽出器によって出力される特徴を予測する．予測器はラベル付きデータで訓練されているため，ラベルなしデータから抽出された特徴がラベル付きデータと同じ分布になる場合，効果的に予測できる．

4.2　インスタンスセグメンテーション

本節は与えられた粒子画像のセグメンテーションモデルの構築について説明する．本研究では 3 つのモデルを構築する予定である．1 つは粒子画像で訓練するモデル．このモデルはベースラインモデルと呼ぶこととする．もう 1 つは粒子の画像と人

工で生成した粒子見たい画像を使用して，モデルを訓練する．このモデルは汎化モデルと呼ぶこととする．もう 1 つ今後の実験で構築する予定である．

本研究はインスタンスセグメンテーション技術で粒子のセグメンテーションを行う．Facebook の He らは開発した Mask R-CNN [4] というインスタンスセグメンテーションモデルを用いる（図 5）．本研究で使用するモデルは，事前に COCO データセット [11] で訓練したモデルである．COCO データセットは Microsoft で作成されたデータセットである．主に画像の認識，セグメンテーション問題に対し，訓練データとして用いるデータセットである．モデルは以下の 6 つのステップからなる．

- 画像の読み込み
- Data Augmentation
- CNN による画像の特徴化
- 特徴マップから ROI を抽出
- ROI Align で ROI の形を同一化
- 分類，boundbox の推定，マスクの予測

まず，このモデルは，入力として画像を 1 枚読み込む．入力された画像は様々な方法で処理を行う．このような操作は，Data Augmentation と呼ばれる手法である． Data Augmentation は，深層学習において訓練時に用いるデータを変形させてデータ量を増やす手法である．過学習を避けることができ，モデルの汎用性が向上するとされ，広く一般に利用されている．本研究では，Detectron2 に付属している Data Augmentation 手法を使用した．以下のような処理を行った．

- 画像の輝度を 0.5〜1.5 倍に調整
- 画像のコントラストを 0.5〜1.5 倍に調整
- 水平方向で反転

次に，様々な手法で変形された写真は backbone ニューラルネットワークに入力する．本研究では，ResNet50 は backbone ニューラルネットワークとして使用する．入力された写真の特徴マップを出力する．

さらに，RPN という物体候補領域（ROI）を抽出するネットワークを使用して，ROI を抽出する．RPN は，入力画像中から物体候補領域（ROI）を抽出するためのネットワークである．

そして，予測器のネットワークに入力するために，特徴の形は同じになることが必要である．ROI Align は抽出された違う形の ROI の形を同一化するための方法である．

最後，同一化された ROI の分類，boundbox の推定，マスクの予測を行う．粒子は複数種類があるが，本研究は主に粒子のセグメンテーションによる化学製品の品質を検査するために，種類は粒子という 1 種類を設定する．

このようにして構成されたモデルによって，入力された画像から粒子はオブジェクトごとで識別することが可能となる．

4.3 Data Augmentation 手法の選択

本節は新しい Data Augmentation 手法の選択を説明する．

Detectron2 に付属している Data Augmentation 手法はいくつかある．粒子のセグメンテーションにおいて，どのような Data Augmentation 手法が良いかを実験した．輝度やコントラストの調整と水平方向で反転をする場合，粒子の形に大きな影響がないので，この 3 つ手法は全部使用するのかよいと分かった．次に，粒子の形を少し変換する場合，モデルの汎化性能が良くなると考えられるので，以下の 2 つ手法を試すこととした．



図 6　人工画像の例

- 高さを 20%〜100%でクロップ，幅を 20%〜100%でクロップ
- 画像を 0.5〜1 倍にリサイズ

画像のクロップとリサイズによって粒子の形が変更されるため，具体的にどのような影響があるかを確認する必要がある．そこで以下の 4 つ比較モデルを構築した．比較実験の結果は節 5.1 で説明する．

- クロップとリサイズしないモデル
- クロップするモデル
- リサイズするモデル
- クロップとリサイズするモデル

4.4 汎化モデルの学習

本節はモデルを訓練する時，人工データを追加することを説明する．本節のモデルは節 4.2 を説明したように構築する．

画像セグメンテーションモデルには 2 種類のモデルを構築する．ベースラインモデルを使うデータは顕微鏡で撮影した化学製品の画像である．汎化モデルを使うデータはベースラインモデルが使用するデータかつ人工画像を使用する．人工画像は図 6 の示すような画像である．人工画像は自分で生成した輪郭が明確な重なっている画像（a）と複数同じオブジェクト等で写っている画像（b）で構成された画像である．人工画像は合計 35 枚である．

モデルの訓練のために，それぞれのデータを訓練データとテストデータに分割する．比率は 9:1 とした．その結果，ベースラインモデルでは訓練データが 91 枚であり，汎化モデルでは訓練データが 121 枚である．2 種類のモデルを比較するために，テストデータは同じデータセットを使用する．テストデータは 14 枚である．それらの画像を用いてベースラインモデルと汎化モデルを構築する．学習率は 1.5×10^{-5} と設定した．

ベースラインモデルと汎化モデルの結果は節 5.2 で説明する．

5 実　　験

本節は本研究を構築したモデルの実験を説明する．節 5.1 は違う Data Augmentation 手法を使用したモデルの比較実験を説明する．節 5.2 はベースラインモデルと汎化モデルの比較実

表 3　比較モデル構築に用いる画像数

	訓練データ	テストデータ
「不規則な円形」	3	1
「四角」	68	6
「円形」	19	3

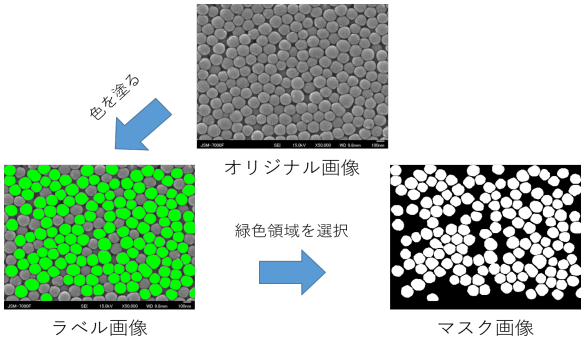


図 7　マスク画像の作成方法

験を説明する．

5.1　Data Augmentation 比較モデルの結果

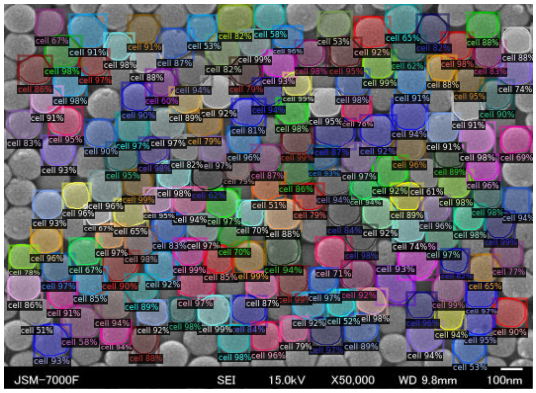
これら 4 つのモデルを使用するデータはラベル付きデータである．人工データは使用しない．使用するデータは表 3 で示す．

節 4.2 の説明を通してモデルの構築を行う．モデルは表 3 のデータを使用する．全部ラベル付きデータである．

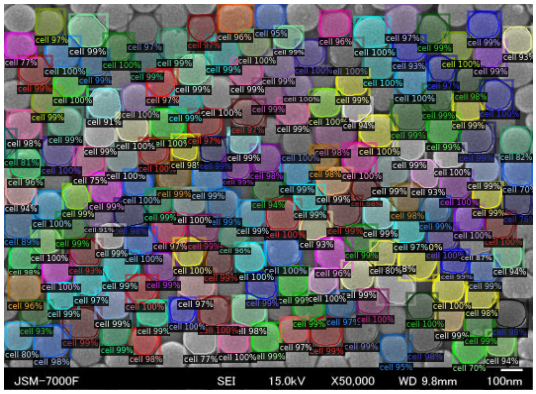
モデルの学習を始める前に入力データについて説明する．画像のセグメンテーション問題の入力は 2 枚の写真である．1 枚は粒子のオリジナル画像（図 7 の上の画像），1 枚は対応するマスク画像（図 7 の左側の画像）である．マスク画像はラベル画像（図 7 の右側の画像）により生成される．まず，ラベル画像中緑色を選択する．次に，選択されたポリコンの座標を取得する．そして，様々なな座標で囲まれた範囲を白色を塗る．その以外の範囲は黒色を塗る．ラベル画像中から色を選択する際，小さい領域を選択する可能性がある．例えば，画像の縁は非常に細い緑色の線であり，このようなの線は人間の目では見えにくい．本研究は面積 15×15 ピクセル未満の領域を除いた．

構築されたモデルは表 3 のデータを使用する．結果は表 4 を示す．この結果から見ると，クロップする後，モデルの正解率は向上した．リサイズする，クロップしない場合，一番良い結果が得られた．クロップするモデルの正解率はリサイズするモデルの正解率より低い原因はある粒子の画像の問題である．ある粒子の画像の中央に小さな部分にある，残りは空白である．クロップする場合，空白区域を選択する可能性がある．この原因で，無駄な訓練を行った．リサイズする場合，AP_{small} の正解率は大幅に向上した．すなわち，画像が小さくなった場合，面積は小さい粒子の識別の正解率は向上した．リサイズとクロップ両方する場合，大きな向上は見えない．そして，このモデルの AP_{small} と AP_{large} はリサイズのみするモデルより低いである．全体的に考えて，リサイズのみする十分と思う．本研究以降は下の 4 つ Data Augmentation 手法を使用する．

- 画像を 0.5〜1 倍にリサイズ
- 画像の輝度を 0.5〜1.5 倍に調整
- 画像のコントラストを 0.5〜1.5 倍に調整
- 水平方向で反転



(a) ベースラインモデル



(b) 汎化モデル

図 8　予 測 結 果

5.2　ベースラインモデルと汎化モデルの比較結果

構築された 2 つモデルは表 5 のデータを使用する．汎化モデルは 35 枚人工画像を追加した．「不規則な円形」，「四角」，「円形」のデータはベースラインモデルと同じデータを使用する．Mask R-CNN を用いて 2 つのモデルを訓練する．2 つのモデルは 10,000 回ずつ訓練する．

テストデータを使用して 2 つモデルの評価を行った．テストデータは同じデータセットを使用する．テストデータは人工データを含めて，ラベル付きデータ中から 14 枚をランダムに選択する．各モデルを評価結果を表 6 を示す．結果からみると，平均精度は向上した．でも，AP_{small} の平均精度は減少した．AP_{large} の平均精度は大幅に向上した．

図 9 は同じ画像で 2 つモデルの予測結果である．図に示すように，人工データを追加する前，粒子の輪郭の周辺の予測結果は良い精度とは言えない．輪郭は明確な粒子の様なの人工データ画像（図 6）を追加した後，モデルは粒子の輪郭は重要な情報と認識した．面積は小さいの粒子の写真に対する精度は低い．本研究では，異常な粒子を検出するためには，大きな粒子のみを検出することが重要であると考えられる．従って，大きな粒子の写真に対する精度が向上していることから，ベースラインよりも良い手法であると考えられる．

表 4　比較モデルの結果

	クロップとリサイズしないモデル		クロップするモデル		リサイズするモデル		クロップとリサイズするモデル	
	bbox	seg	bbx	seg	bbox	seg	bbx	seg
AP	25.37	24.60	28.93	28.97	28.77	29.04	29.40	29.56
$AP_{IOU=0.5}$	37.60	37.60	33.48	33.51	43.32	34.33	34.38	34.39
$AP_{IOU=0.75}$	32.59	30.24	32.50	33.44	33.22	33.31	33.38	34.30
AP_{small}	0.49	0.49	13.14	12.56	14.67	15.14	13.07	13.33
AP_{medium}	20.91	20.10	32.71	32.93	31.66	32.06	33.19	33.60
AP_{large}	80.91	82.36	92.00	90.31	89.84	87.87	88.75	86.64

表 5　モデル構築に用いる画像数

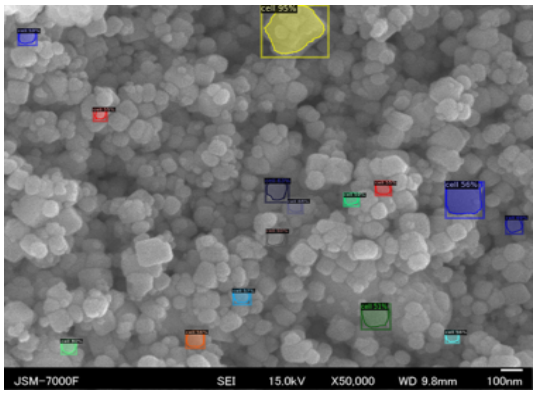
	ベースラインモデル	汎化モデル
「不規則な円形」	4	4
「四角」	67	67
「円形」	20	20
人工データ	0	30

表 6　ベースラインモデルと汎化モデルの結果

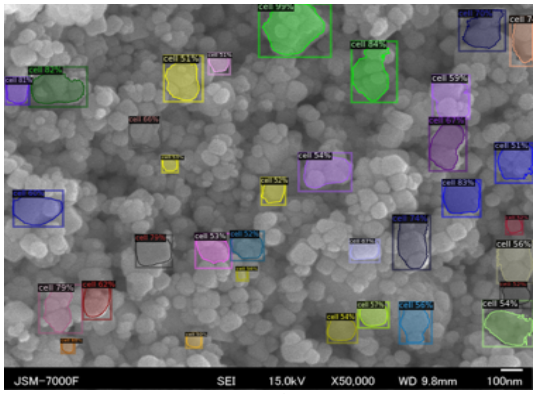
	ベースラインモデル		汎化モデル	
	bbox	seg	bbx	seg
AP	27.99	28.05	32.26	32.33
$AP_{IOU=0.5}$	32.62	32.63	37.57	37.58
$AP_{IOU=0.75}$	36.61	31.62	36.56	36.57
AP_{small}	4.93	4.76	4.13	3.99
AP_{medium}	31.90	32.98	36.68	36.83
AP_{large}	65.88	65.92	80.32	80.25

表 7　ベースラインモデルと汎化モデルの結果

	ベースラインモデル		汎化モデル		ドメイン適応モデル	
	bbox	seg	bbx	seg	bbx	seg
AP	27.09	28.05	32.26	32.33	28.11	28.50
$AP_{IOU=0.5}$	32.62	32.63	37.57	37.58	35.61	35.62
$AP_{IOU=0.75}$	36.61	31.62	36.56	36.57	35.47	35.33
AP_{small}	4.93	4.76	4.13	3.99	5.39	4.81
AP_{medium}	31.90	32.98	36.68	36.83	30.17	31.22
AP_{large}	65.88	65.92	80.32	80.25	78.17	75.67



(a) 汎化モデル



(b) ドメイン適応モデル

図 9　予 測 結 果

5.3　ドメイン適応したモデルの結果

ドメイン適応モデルを使用するデータはラベル付きデータとラベルなしデータである．ラベル付きデータとラベルなしデータは表 1 が示したデータを使用する．テストデータはラベル付きデータから選択する．具体的には表 3 の様なになる．節 4.1 の説明を通してモデルの構築を行う．step1 を訓練する時使用のデータは人工データも含む．step2 を行う時，ラベルなしデータを使用する．step3 は表 3 のテストデータで評価する．

この 3 つモデルは全部 Data Augmentation を使用した．ベースラインモデルは粒子の写真のみ使用する．汎化モデルは訓練データに人工データを追加した．ドメイン適応モデルは人工データ追加モデルの基礎上でドメイン適応を行う．モデルを評価結果を表 7 を示す．この結果から見ると，汎化モデルの正解率は一番高い．ドメイン適応モデルの正解率は少し低いである．

新しい形状の粒子データは全部ラベルなしデータのため，正解率を計算することができない．直接予測を見る．図 9(a) は汎化モデルの予測結果．図 9(b) はドメイン適応モデルの予測結果．この予測結果から見ると，ドメイン適応した後，新しい形状の粒子の識別率が向上した．

6　まとめと今後の課題

本研究では，ドメイン適応を用いて，粒子のセグメンテーション手法を提案した．ラベルなしデータを効果的に使用してモデルを構築することで，より高い精度を得ることを目指す．

今後の課題としては，より多くの損失関数を試すことで，より良いの損失関数を模索する予定である．そして，粒子の測量は損失関数を用いるために，全体的な訓練を行う手法を検討する予定である．

謝　　辞

本研究の一部は JSPS 科学研究費助成事業 JP18H03243, JP17H00762, JP18H03244 による助成を受けたものです．ここに記して謝意を表します．

文　　献

[1] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Proceedings of the IEEE Conference on Computer Vision and Ppattern Recognition*, pp. 248–255, 2009.

[2] Tzeng Eric, Hoffman Judy, Saenko Kate, and Darrell Trevor. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2962–2971, 2017.

[3] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *The Journal of Machine Learning Research*, Vol. 17, No. 1, pp. 2096–2030, 2016.

[4] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2961–2969, 2017.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

[6] Dong Huk Park, Lisa Anne Hendricks, Zeynep Akata, Anna Rohrbach, Bernt Schiele, Trevor Darrell, and Marcus Rohrbach. Multimodal explanations: Justifying decisions and pointing to the evidence. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8779–8788, 2018.

[7] Angjoo Kanazawa, Shubham Tulsiani, Alexei A. Efros, and Jitendra Malik. Learning category-specific mesh reconstruction from image collections. In *Proceedings of the European Conference on Computer Vision*, pp. 386–402, 2018.

[8] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Proceedings of the International Conference on Neural Information Processing Systems*, pp. 1097–1105, 2012.

[9] Samuli Laine and Timo Aila. Temporal ensembling for semi-supervised learning. In *Proceedings of the International Conference on Learning Representations*, 2017.

[10] Dong-Hyun Lee. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In *Proceedings of the Workshop on challenges in representation learning, International Conference on Machine Learning*, Vol. 3, 2013.

[11] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *Proceedings of the European Conference on Computer Vision*, pp. 740–755, 2014.

[12] Lorien Y Pratt. Discriminability-based transfer between neural networks. *Advances in neural information processing systems*, Vol. 5, pp. 204–211, 1992.

[13] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 39, No. 6, pp. 1137–1149, 2016.

[14] Mehdi Sajjadi, Mehran Javanmardi, and Tolga Tasdizen. Regularization with stochastic transformations and perturbations for deep semi-supervised learning. In *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1163–1171, 2016.

[15] Guohua Shen, Tomoyasu Horikawa, Kei Majima, and Yukiyasu Kamitani. Deep image reconstruction from human brain activity. In *Proceedings of the International Joint Conference on Neural Networks*, pp. 1049–1056, 2017.

[16] Kihyuk Sohn, David Berthelot, Chun-Liang Li, Zizhao Zhang, Nicholas Carlini, Ekin D Cubuk, Alex Kurakin, Han Zhang, and Colin Raffel. Fixmatch: Simplifying semi-supervised learning with consistency and confidence. *arXiv preprint arXiv:2001.07685*, 2020.

[17] Hongyi Zhang, Moustapha Cisse, Yann N. Dauphin, and David Lopez-Paz. mixup: Beyond empirical risk minimization. In *Proceedings of the International Conference on Learning Representations*, 2018.