



ISPRS 写真測量とリモートセンシングのジャーナル

195巻、2023年1月、ページ192～203

リモートセンシングのクロスドメインシーン分類のための半教師あり双方向アライメント

魏黄^a、Yilei Shi^b、シヨン・ジトン^c、チー・ワン^c、シャオ・シャン・ジュー^a

もっと見る

概要 | 共有 | 引用

<https://doi.org/10.1016/j.isprsjprs.2022.11.013>

権利とコンテンツを取得する

クリエイティブ・コモンズの下でライセンス

オープンアクセス

抽象的な

リモートセンシング (RS) 画像シーン分類は、その広範な応用の可能性からますます注目を集めています。従来の完全教師ありアプローチでは、通常、手動でラベル付けされた大量のデータが必要です。利用可能な RS 画像が増えるにつれ、これらのラベルなしデータをどのように活用するかが緊急の課題となっています。半教師あり学習は、少数のラベル付きデータを使用して多数のラベルなしデータの自己学習をガイドする直感的な戦略です。ただし、異なるデータセット間ではドメインが大幅にシフトするため、これをデータセット間 (つまり、ドメイン間) のシーン分類に適用するのは困難です。この目的を達成するために、ドメインシフトを軽減し、完全にラベル付けされた RS シーン データセット (ソース ドメイン) から制限付きラベル付き RS シーン データセット (ターゲット ドメイン) に知識をさらに転送できる半教師ありドメインアダプテーション (SSDA) は、次のようになります。実現可能な解決策。この論文では、RS クロスドメイン シーン分類のための双方向サンプルクラスアライメント (BSCA) と呼ばれる SSDA 方法を提案します。BSCA は、教師なしアラインメント (UA) と教師ありアラインメント (SA) という 2 つのアライメント戦略で構成されており、どちらもドメインシフトの減少に貢献します。UA は、クラス ラベルを必要とせず、ドメイン間の最大平均不一致の距離を短縮することに集中します。対照的に、SA は、クラスを意識して、ソース サンプルから関連ターゲットクラス センターへ、およびターゲット サンプルから関連ソース クラス センターへの両方の分布調整を達成することを目的としています。提案された方法の有効性を検証するために、広く使用されている 4 つの RS シーン分類データセットに基づいて構築された RS-SSDA ベンチマークで広範なアブレーション、比較、および視覚化実験が行われます。実験結果は、いくつかの最先端の方法と比較して、当社の BSCA がコンパクトな特徴表現と低エントロピー分類境界により優れたクロスドメイン分類パフォーマンスを達成していることを示しています。私たちのコードは<https://github.com/hw2hwei/BSCA>で入手できます。

前の

次

キーワード

リモートセンシング; 半教師ありドメイン適応; クロスドメイン分類; サンプルクラスの双方向アラインメント

1. 導入

リモートセンシング (RS) 画像シーン分類 (Wang et al., 2020, Cheng et al., 2017, Bai et al., 2022, Ma et al., 2021a, Xiong et al., 2022) は、その特徴によりますます注目を集めています。土地利用 (Ma et al., 2017) や都市計画 (Rottensteiner et al., 2012, Qiu et al., 2019) などの広範な現実世界の応用。ただし、RS 画像シーン分類のほとんどの最先端の方法は、完全に教師ありの深層学習モデルに基づいており、手動でラベル付けされた多数のデータに大きく依存しています。対照的に、ラベルのない RS 画像がますます蓄積され、研究者がアクセスできるようになりました。このような状況において、ラベルなし RS 画像をいかに有効に活用するかが喫緊の課題となっている。

RS 画像シーン分類に関するいくつかの半教師あり学習 (SSL) 作品 (Han et al., 2018, Miao et al., 2022) は、ラベルなしデータの自己学習によってアノテーションへの大きな依存を軽減するために試みられています。ただし、クロスデータセット (クロスドメイン) シフトのために、他の既存のデータセットの大規模なラベル付きデータを直接利用することはできません。具体的には、異なる時間、さまざまなスケール、さまざまな撮影機器、固定されていない撮影角度など、RS シーン分類データセット全体でドメインのシフトを引き起こす多くの要因があります。実際には、これらの要因が混在することでドメインの移行がさらに複雑になります。一方、いくつかの教師なしドメイン適応 (UDA) 方

法がRSクロスドメインシーン分類のために研究されています (Ma et al., 2021b, Zheng et al., 2022b, Yu et al., 2022, Ahmed et al., 2021, Wurm et al., 2019)、既存のラベル付き RS 画像データセット (ソース ドメイン) を利用して、共有クラス空間内でラベルなし RS 画像データセット (ターゲット ドメイン) の分類を支援することを目的としています。ただし、UDA の一般化は、ソース ドメインとターゲット ドメイン間のブリッジとしてドメイン シフトを軽減するために不可欠なラベル付きターゲット データが欠如しているため制限されます。この観点から、SSL と UDA を組み合わせた半教師ありドメイン アダプテーション (SSDA) は、ガイダンスに従ってアノテーションが豊富なソース ドメインからマージナル ラベルが付けられたターゲット ドメインに知識を転送できるため、実現可能なソリューションとなります。いくつかのラベル付きターゲット データの一部。それでも、UDA と同様に、SSDA もドメイン シフトに悩まされています。

特徴の位置合わせは、ドメインのシフトを減らすための経験的に効果的な方法であるため、ターゲット ドメインの分類パフォーマンスを向上させることができます。機能レベルの観点から見ると、SSDA の現在の機能アライメントは、サンプルアライメント (Motian et al., 2017) とドメインアライメント (Kim and Kim, 2020) の 2 種類に大別できます。サンプルの調整は、クロスドメインのクラス内インスタンスを調整することを目的としていますが、ドメインの調整は、2 つのドメインのグローバルな機能分布を調整することを目的としています。中間戦略として、クラスアライメントはノイズの多いサンプルに対する感度が低いサンプルアライメントよりも堅牢であり、焦点化操作によりドメインアライメントよりも粒度が細くなります。正確で安定したクラスアライメントを行うには、トレーニング段階で複数のサンプルからクラス中心を計算する必要があります。残念ながら、同じクラス内のサンプルの数は各ミニバッチでかなり制限されているため、クラスごとの中心の正確かつ堅牢な計算やさらなるクラスの調整が妨げられます。

高い計算負荷をかけずに柔軟かつ堅牢なクラスレベルの特徴アライメントを実現するために、RS クロスドメインシーン分類用の双方向サンプルクラスアライメント (BSCA) と呼ばれる新しい全体的な SSDA フレームワークを考案しました。これは、2 つの特徴アライメントモジュールで構成されます。

(1) 教師なしアライメント (UA) モジュールが導入され、ラベル付きサンプル (ソース サンプルとラベル付きターゲット サンプルを含む) から抽出された特徴と、ラベルなしターゲット サンプルから抽出された特徴の間のグローバルな分布調整が実現されます。クラス情報は要求されません。これは、ソースドメインとラベルなしターゲットドメインの間だけでなく、ラベル付きターゲットドメインとラベルなしターゲットの間の最大平均不一致 (MMD) (Gretton et al., 2012a, Pan et al., 2010) の距離を短縮することによって達成されます。ドメイン;

(2) 独自の教師付きアライメント (SA) モジュールが提案され、サンプルから対応するクロスドメインクラスセンターまでの特徴アライメントを実現します。ここで、クラスセンターは、ソースサンプルとターゲットサンプルから抽出された特徴を個別に保存できるメモリバンクメカニズムによって計算されます。SA には、2 種類のクラスセンターに対応する、ソースメモリバンクとターゲットメモリバンクの 2 つの分離されたメモリバンクがあります。ソースサンプルから抽出された特徴を例にとると、トレーニングの反復ごとに 2 つの役割を同時に果たします。(a) 同じクラスのソースメモリバンクを更新するために使用されます。(b) ターゲットメモリバンクから計算される、同じクラスのターゲットクラスの中心に位置合わせされます。同様に、ターゲットサンプルから抽出された特徴は、ターゲットメモリバンクを更新するために使用され、対応するソースクラスセンターに合わせられます。したがって、我々の SA は、ソースサンプルからターゲット中心へ、およびターゲットサンプルからソース中心への両方の特徴アライメント、つまり双方向アライメントを実現できます。さらに、メモリバンクの恩恵を受けて、過去のサンプルの豊富なクラス認識機能を保存し、少量のコンピューティングリソースでリアルタイムのサンプルクラス機能の調整に使用できます。つまり、BSCAのSAは、ミニバッチの限界を超えた豊富なクラス認識特徴量から堅牢なクラスセンターを計算し、オンラインでクラスごとの特徴量調整をサポートすることができます。

提案された BSCA の有効性を客観的に評価するために、広く使用されている 4 つの RS 画像シーン分類データセットの共通クラスを使用して RS-SSDA ベンチマークを構築し、そのメソッドをいくつかの最先端の SSL、UDA、このベンチマークの SSDA メソッド。実験結果は、我々の BSCA が比較手法を上回っており、ベンチマークの合計 12 の適応シナリオの最高の平均分類パフォーマンスを達成していることを示しています。全体として、私たちの貢献は次のように要約できます。

- 我々は、教師ありアライメント (SA) と教師なしアライメント (UA) の 2 つのモジュールで構成される、RS-SSDA用の双方向サンプルクラスアライメント (BSCA) 方法を提案します。SA は、ソース サンプルからターゲット クラス センターへ、およびラベル付き/擬似ラベル付きターゲット サンプルからソース クラス センターへの双方向の特徴アライメントを実現することを目的としています。対照的に、UA は、クラス情報を要求せずに、ラベル付きサンプルとラベルなしサンプルから抽出された特徴間のグローバルな位置合わせに焦点を当てます。
- 提案手法の有効性を検証するために、広く使用されている 4 つの RS 画像シーン分類データセットの 7 つの共通クラスを含む RS-SSDA ベンチマーク データセットを収集し、それに基づいて広範な実験を実行します。
- いくつかの最先端の SSL/UDA/SSDA 方法と比較して、私たちが提案する BSCA は、上記の RS-SSDA ベンチマークで最良の平均分類結果を取得し、私たちの方法の有効性を実証しています。

2. 関連作業

2.1. 半教師ありドメイン適応

SSDA は、SSL と UDA を組み合わせたものとして最近注目されている研究分野です。私たちの知る限り、SSDA 手法には大まかに 3 種類あります。敵対的トレーニングベースの SSDA (Tzeng et al., 2017, Jiang et al., 2020)、エントロピー最適化ベースの SSDA (Grandvalet et al., 2005, Saito et al., 2020)、および特徴アライメントベースの SSDA (Motian et al., 2017, Kim and Kim, 2020, Li et al., 2021)。具体的には、Tzeng et al. (2017) SSDA

問題とディープモデルの一般化に対処するためのシャムアーキテクチャを提案しました。このアーキテクチャは、ラベル付けされたソースサンプルとターゲットサンプルの特徴を処理するために、分類と対照的セマンティックアライメント (CCSA) 損失を導入することによって、識別的な埋め込み部分空間を学習します。ただし、ラベルのないターゲットサンプルの使用は無視されます。齊藤ら。(2019) SSDA モデルを敵対的に最適化する新しいミニマックスエントロピー (MME) アプローチによってこの分野を強調しました。MME は、特徴分類器に対してはラベルなしターゲットデータの条件付きエントロピーを最大化し、特徴エンコーダーに対しては最小化することを交互に行うことができます。ジャンら。(2020) ドメインギャップを越えて敵対的な例を導くための一般的な双方向敵対的トレーニング (BiAT) を考案しました。Li と Hospedales (2020) は、マルチソースの教師なしドメイン適応 (MSDA) と SSDA の両方に対して、計算的に扱いやすく実際に効果的なオンライン最短パスメタ学習フレームワークを提案しました。Kim and Kim (2020) は、ドメイン内の不一致を減らすことで機能を調整することを目的として、主にアトラクション、摂動、探索の3つのスキームで構成される SSDA フレームワークを提示しました。リーら。(2021) 敵対的適応クラスタリング損失を導入してラベルのないターゲットデータの特徴をクラスにグループ化し、次にクラスごとの機能を実装することで、ドメイン間適応とドメイン内適応の両方を達成するクロスドメイン適応クラスタリング (CDAC) と呼ばれるアプローチを考案しました。ドメイン間の調整。

2.2. RS シーン分類のドメイン適応

近年、RS画像シーン分類のドメイン適応 (DA) に焦点を当てた研究が増えています (Song et al., 2019, Lu et al., 2020, Zheng et al., 2021, Zhang et al., 2020, Zhu et al., 2021, Zheng et al., 2022a, Lasloun et al., 2021)。ソングらでは、(2019) では、特徴部分空間内のソースドメインとターゲットドメインを位置合わせするために、RS画像シーン分類のDAに対して、CNNモデルに追加された新しい部分空間位置合わせレイヤーが提案されました。その結果、対象ドメインの分類に適応するようにCNNモデルを最適化できます。Zhang et al. (2020) では、ソースドメインデータを無関係な部分空間にマッピングすることを回避しようとする部分空間相関最大化 (SCM) と、無関係な部分空間へのマッピングを回避しようとする動的統計分布調整 (DSDA) で構成される、RS画像シーン分類用の相関部分空間動的分布調整方法が提案されました。クロスドメイン配信の不一致を軽減します。アテンションベースのマルチスケール残差適応ネットワーク (AMRAN) が、クロスシーン分類タスク用に提案されました (Zhu et al., 2021)。AMRAN では、周辺分布と条件付き分布の両方が考慮され、マルチスケールアテンションメカニズムを使用して堅牢な特徴と完全な情報が抽出されました。Zhengらでは、(2022a) では、単一ソース複数ターゲットのドメイン適応タスクがRSアプリケーション向けに検討され、2段階適応ネットワーク (TSAN) と呼ばれる新しいアルゴリズムが提示されました。ソースドメインと混合マルチターゲットドメイン全体、および (2) 複数のターゲットドメインの固有の特徴を学習するために、混合マルチターゲットドメインをその擬似ドメインラベルで分割する自己教師あり学習を採用します。Lasloun et al. (2021) では、セクション2.1で述べたMMEアルゴリズム (Saito et al., 2019) をRS画像シーン分類の目的でマルチソースSSDAに適用しました。

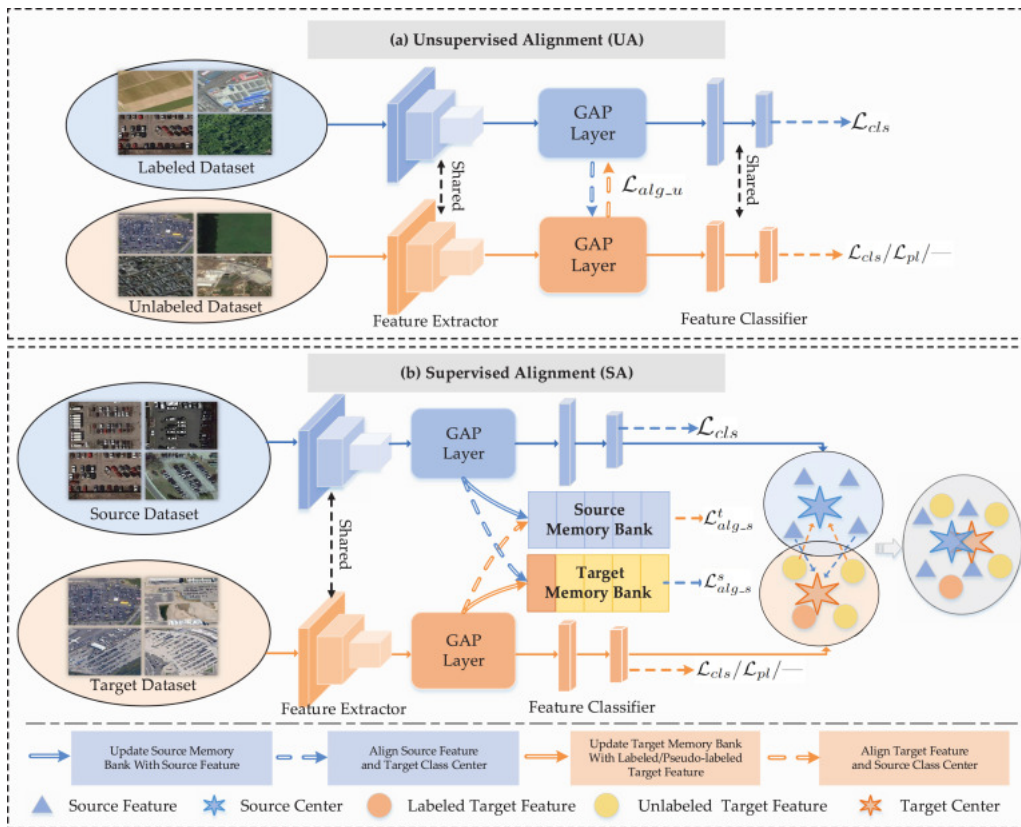
私たちの研究は、ドメイン不変の特徴を取得するために、単一ソースドメインから単一ターゲットドメインに転送可能な知識を学習することを目的として、単一ソース、単一ターゲットSSDA設定におけるRS画像シーン分類に主に焦点を当てています。この表現は、マルチソース/マルチターゲットの研究にも役立ちます。

2.3. メモリバンク

メモリバンクモジュールは、ニューラルネットワークの外部に追加情報を保存でき、参照用の辞書として使用できるため、広く適用されています。たとえば、最も明白な情報を外部メモリモジュールに書き込むことを学習できるメモリ拡張型時間双方向学習ネットワークは、人間の行動認識のためにYuan et al.で提案されました。(2019)。Lv et al.では、リアルタイムフレーム構築のプロトタイプとして通常のダイナミクスをエンコードする双方向プロトタイプユニットが提示されました(2021)。対比学習の観点から、教師なし視覚表現学習のために、運動量コントラスト (MoCo) の移動平均キューエンコーダを備えたメモリバンクに基づいて双方向辞書が構築されました (He et al., 2020)。アロンソらでは(2021)では、メモリバンクメカニズムを対照学習と組み合わせて、セグメンテーションネットワークを強制して、半教師ありセマンティックセグメンテーションのクロスドメインのクラス内サンプルと同様のピクセルレベルの特徴表現を取得しました。ただし、ラベル付きデータの特徴ベクトルでのみ更新され、ラベルなしデータの利用は無視されます。

3. 方法論

このセクションでは、ユニバーサルSSDAのいくつかの表記法を紹介することから始まり、次に共有ネットワークアーキテクチャを示し、最後に提案されたBSCAをそのトレーニング手順の概要とともに詳細に説明します。BSCAベースのSSDAアーキテクチャのワークフローを図1に示します。



ダウンロード:高解像度画像のダウンロード(1MB)

ダウンロード:フルサイズの画像をダウンロード

図1. RS-SSDA用に提案されたBSCAのワークフローは、次の2つの部分で構成されます。(a) クラス情報を持たずにクロスドメインのクラスに依存しない機能を調整するUA。(b) クロスドメインのクラス認識機能をクラスラベル/疑似ラベルの要求に合わせて調整するSA。

3.1. 表記法

SSDAの設定には、3つのサブドメインがあります: 完全にラベル付けされたソースドメイン \mathcal{S} 、限定されたラベルが付けられたターゲットドメイン \mathcal{T}_l 、ラベルのないターゲットドメイン \mathcal{T}_u 。彼らのサンプルセットは次のように表されます。 $\mathcal{D}_s = \{(\mathbf{x}^s, \mathbf{y}^s)\}_{i=1}^{N_s}$ 、 $\mathcal{D}_l = \{(\mathbf{x}^l, \mathbf{y}^l)\}_{i=1}^{N_l}$ 、そして $\mathcal{D}_u = \{(\mathbf{x}^u)\}_{i=1}^{N_u}$ 、それぞれ、ここで \mathbf{x} 、 \mathbf{y} 、そして N は、それぞれ画像、それに関連付けられたラベル、およびそのドメインのサンプル番号です。ここ N_l よりもはるかに小さいです N_u 。ソースラベルとターゲットラベル、 \mathbf{y}^s そして \mathbf{y}^l 、同じクラス空間 $\{1, \dots, K\}$ を持ちます。ここで、 K はクラスの数です。

3.2. 共有ネットワークアーキテクチャ

いくつかの既存のSSDA文献(Chen et al., 2019, Saito et al., 2019)に従って、この論文ではCNNモデルをRSクロスドメインシーン分類に使用します。これは3つのコンポーネントで構成されています: CNNベースの特徴抽出器 \mathcal{E} 、グローバル平均プリーング(GAP)レイヤー \mathcal{G} 、および特徴分類器 \mathcal{C} 。これらはすべて、ソースドメインとターゲットドメインの間で共有されます。

まずは特徴マップ $\mathbf{m} \in \mathbb{R}^{H \times W \times C}$ ここで、 H 、 W 、および C は、それぞれ空間の高さ、空間の幅、および特徴の寸法であり、画像から抽出されます。 \mathbf{x} 特徴抽出器による \mathcal{E} と表されます。

$$\mathbf{m} = \mathcal{E}(\mathbf{x}). \quad (1)$$

次に、グローバル特徴ベクトル $\mathbf{f} \in \mathbb{R}^C$ から計算されます \mathbf{m} GAP層経由 \mathcal{G} (Lin et al., 2013)として

$$\mathbf{f} = \mathcal{G}(\mathbf{m}) = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W \mathbf{m}_{ij}. \quad (2)$$

ついに \mathbf{f} 特徴分類器に入力されます \mathcal{C} クラスごとの確率的予測ベクトルを取得するには $\mathbf{p} \in \mathbb{R}^K$ と表されます。

$$\begin{cases} \mathbf{f} = \frac{\mathbf{f}}{\|\mathbf{f}\|}, \\ \mathbf{p} = \sigma(\mathcal{C}(\mathbf{f})/T), \end{cases} \quad (3)$$

ここで、最初の方程式は l_2 正規化(Ranjan et al., 2017)。用語 σ はSoftmax関数で、予測ベクトルのすべての要素の合計を1に正規化できます。これは、要素の値を意味します。 \mathbf{p} は、対応するクラスの確率として見ることができます。 T 温度のハイパーパラメータです。

SSDA文献の一般的な慣行(Saito et al., 2019, Li et al., 2021)に従って、両方のソースドメインからのすべてのラベル付きターゲットドメイン \mathcal{D}_l 分類損失(標準のクロスエントロピー損失)を介してネットワークをトレーニングするために使用されます。

$$\mathcal{L}_{ds} = -\sum_{\{x,y\} \in \mathcal{D}_s \cup \mathcal{D}_t} y \log p_y. \quad (4)$$

3.3. サンプルクラスの双方向アライメント

BSCAはUAとSAで構成されており、図1に示すような教師なしおよび教師ありの特徴位置合わせを実現することを目的としています。

3.3.1. 教師なしアライメント

図1の「教師なしアライメント」部分に示すように、UAは主に、ラベル付きサンプルから抽出された特徴とラベルなしサンプルから抽出された特徴との間の教師なしドメイン間アライメントに集中する。カーネルヒルベルト空間(RKBS) (Borgwardt et al., 2006) の再現におけるMMD (Gretton et al., 2012a, Pan et al., 2010) の距離は、これら2つの分布の平均値の類似性を次のように評価するために使用されます。

$$\mathcal{L}_{alg-u} = d_{MMD}^2 = \left\| \frac{1}{N_s + N_t} \sum_{i=1}^{N_s + N_t} \phi(\mathbf{f}_i^{sl}) - \frac{1}{N_u} \sum_{i=1}^{N_u} \phi(\mathbf{f}_i^u) \right\|_{\mathcal{H}}^2, \quad (5)$$

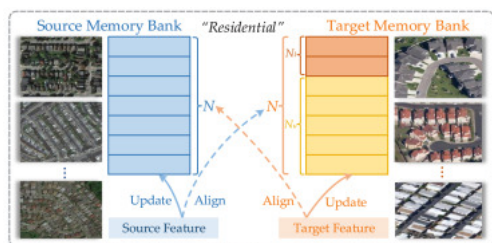
どこ \mathbf{f}^{sl} は、式(1)と2)によって抽出されたグローバル特徴ベクトルを表します。(1)~(2)ソースドメインとラベル付きターゲットドメインの画像から、および \mathbf{f}^u は、ラベルのないターゲットドメインの画像から抽出されたグローバル特徴ベクトルを表します。関数 ϕ 投影できるガウス動径基底関数(RBF)カーネルです。 \mathbf{f}^{sl} そして \mathbf{f}^u RKBSへ。特徴表現の能力を向上させるために、MMDのマルチカーネルバリエーション、つまりMK-MMD (Gretton et al., 2012b, Long et al., 2015) が距離測定に使用されます。

ここで、UAは、ソースデータとターゲットデータの間の唯一の調整ではなく、すべてのラベル付きドメイン(ソースドメインとラベル付きターゲットドメインを含む)とラベルなしターゲットドメイン間の特徴分布の調整に焦点を当てていることに注意してください。その理由は、ラベル付きターゲットデータは教師ありトレーニングによって堅牢で識別的な特徴表現を取得でき、完全にラベル付けされたソースデータと連携して、ターゲットデータから抽出された特徴に対するより堅牢なアンカーとして、ドメインレベルでより包括的な特徴分布を提供できるためです。ラベルのないデータ。

3.3.2. 監視付きアライメント

図1(b)に示すように、SAはメモリバンク機構を介してサンプルからクロスドメインクラスセンターまでの特徴アライメントを実現することを目的としています。

SAには、ソースメモリバンクとターゲットメモリバンクがあり、次のように表されます。 \mathcal{M}^s そして $\mathcal{M}^t \in \mathbb{R}^{K \times N \times C}$ 、それぞれソースサンプルとターゲットサンプルから抽出された特徴を保存します。ここ K はデータセットのクラス番号です。 C はフィーチャのチャネル寸法であり、 N は各クラスの項目番号です。 \mathcal{M}^s そして \mathcal{M}^t 。ターゲットメモリバンク \mathcal{M}^t 2つのサブコンポーネントで構成されます: ラベル付きターゲットメモリバンク $\mathcal{M}^l \in \mathbb{R}^{K \times N_s \times C}$ ラベルのないターゲットメモリバンク $\mathcal{M}^u \in \mathbb{R}^{K \times N_u \times C}$ 、どこ $N_l + N_u = N$ 。これらのメモリバンクはゼロに初期化されます。



ダウンロード:高解像度画像のダウンロード (400KB)

ダウンロード:フルサイズの画像をダウンロード

図2. RS クロスドメインシーン分類のSAの図。例としてクラス「住宅」を使用します。(この図の凡例における色の参照の解釈については、この記事のWebバージョンを参照してください。)

メモリバンクベースのSAのパイプラインを図2に示す。図では、青い長方形はソース画像から抽出された特徴を表し、オレンジ色の長方形はラベル付きターゲット画像から抽出された特徴を表し、黄色の長方形はラベルのない信頼性の高いターゲット画像から抽出された特徴を表します。SAはメモリバンクに基づいて、サンプルアライメントとクラスアライメントの両方の利点を組み合わせており、サンプルは柔軟な最適化を可能にし、クラスセンターは安定した堅牢なアンカーを提供します。具体的には、計算量の多い特徴抽出プロセスを必要とせずに、メモリバンクに保存された複数の特徴から堅牢なクラスセンターを計算できます。一方、ミニバッチのサンプルから抽出された特徴は、関連するクロスドメインクラスセンターにリアルタイムで位置合わせできます。その結果、このようなサンプルクラスのアライメントは余分なコンピューティングリソースをほとんど消費せず、トレーニングフェーズ中に柔軟かつ堅牢な機能アライメントの両方を実現します。

SAの詳細なプロセスは、次の3つのステップに分けることができます。

(a)特徴とラベルのペアを生成します。まず、特徴とラベルのペアを生成する必要があります $[\mathbf{f}, \mathbf{y}]$ 各反復で。方程式の使用(1)~(2)の画像 \mathbf{x}^s 、 \mathbf{x}^l 、そして \mathbf{x}^u の組み合わせで入力されます。 \mathcal{E} そして \mathcal{G} 、次に対応するグローバル特徴ベクトル \mathbf{f}^s 、 \mathbf{f}^l 、そして \mathbf{f}^u が出力されます。これら3つのベクトルのうち、 \mathbf{f}^s そして \mathbf{f}^l 関連するラベルと直接組み合わせることができます \mathbf{y}^s そして \mathbf{y}^l 特徴とラベルのペアを形成する $[\mathbf{f}^s, \mathbf{y}^s]$

そして $[\mathbf{f}^l, \mathbf{y}^l]$ 。のために \mathbf{f}^u には関連付けられたラベルがないため、最大アクティブ化クラスが擬似ラベルとして使用されます。 $\hat{\mathbf{y}}^u$ 、として

$$\hat{\mathbf{y}}^u = \underset{i=\{1,\dots,K\}}{\operatorname{argmax}} \mathbf{p}_i^u, \quad (6)$$

どこ \mathbf{p}^u から計算されます \mathbf{f}^u 式を介して (3)。規約 \mathbf{f}^u 、 \mathbf{p}^u 、そして $\hat{\mathbf{y}}^u$ 疑似特徴とラベルのペアを形成します $[\mathbf{f}^u, \mathbf{p}^u, \hat{\mathbf{y}}^u]$ 。

(b)機能とラベルのペアを使用してメモリ バンクを更新します。次に、メモリ バンクを更新するための合理的な更新戦略を設計することが重要です。メモリ バンクは、サンプル クラスのアライメントのアンカーを提供するため、SSDA のパフォーマンスを決定します。 \mathcal{M}^s として \mathcal{M}^l 対応する機能によって更新されます \mathbf{f}^s として \mathbf{f}^l ; ノイズラベルの干渉を可能な限り減らすため、 \mathcal{M}^u 信頼度の高いものによって更新されます \mathbf{f}^u 、その疑似クラス確率がしきい値よりも高い τ 。更新戦略はアルゴリズム 1 に示されています。その中で、 $\mathcal{M}[y, 1:N-1]$ からのアイテムを表します $\mathbf{1}$ に $N-1$ クラスの y 対応するメモリバンク内 \mathcal{M} 。BSCA のオンライン学習フレームワークに従って、メモリ バンクは反復ごとにそれに応じて更新されます。(1)遅延クラス センターではなく、リアルタイムで画像から抽出された特徴と一致するリアルタイム クラス センターを生成します。(2)柔軟な項目を生成します。 \mathcal{M}^u これにより、誤って擬似ラベルが付けられたターゲット サンプルが繰り返し置き換えられるため、そのサンプルの占有を回避できます。メモリ バンクのもう 1 つの利点は、BSCA をクラスのサンプル数に依存しないようにできることです。技術的には、BSCA は、アソシエイト クラス センターの計算のために、クラスごとに少なくとも 1 つのラベル付きソース サンプルと 1 つのラベル付きターゲット サンプルのみを要求します。

Algorithm 1 Updating Memory Banks

Input: Feature-label pairs: $\{[\mathbf{f}^s, \mathbf{y}^s]\}$, $\{[\mathbf{f}^l, \mathbf{y}^l]\}$, and $\{[\mathbf{f}^u, \mathbf{p}^u, \hat{\mathbf{y}}^u]\}$, the size of a mini-batch M , the entropy threshold τ , memory banks: \mathcal{M}^s , \mathcal{M}^l , and \mathcal{M}^u

Output: updated memory banks: \mathcal{M}^s , \mathcal{M}^l , and \mathcal{M}^u

```

1: while  $i < M$  do
2:   Update:  $\mathcal{M}^s[y_i^s, 1:N-1] \leftarrow \mathcal{M}^s[y_i^s, 2:N]$ 
3:   Update:  $\mathcal{M}^s[y_i^s, N] \leftarrow \mathbf{f}_i^s$ 
4: end while
5: while  $i < M$  do
6:   Update:  $\mathcal{M}^l[y_i^l, 1:N_i-1] \leftarrow \mathcal{M}^l[y_i^l, 2:N_i]$ 
7:   Update:  $\mathcal{M}^l[y_i^l, N_i] \leftarrow \mathbf{f}_i^l$ 
8: end while
9: while  $i < M$  do
10:  if  $E_p < \tau$  then
11:    Update:  $\mathcal{M}^u[y_i^u, 1:N_u-1] \leftarrow \mathcal{M}^u[y_i^u, 2:N_u]$ 
12:    Update:  $\mathcal{M}^u[y_i^u, N_u] \leftarrow \mathbf{f}_i^u$ 
13:  end if
14: end while
15: return  $\mathcal{M}^s$ ,  $\mathcal{M}^l$ , and  $\mathcal{M}^u$ 

```

[ダウンロード:高解像度画像のダウンロード \(399KB\)](#)

[ダウンロード:フルサイズの画像をダウンロード](#)

一方、信頼性の高いラベルのないサンプルの擬似ラベルを使用して、ラベルのないターゲット ドメインの重みを改善するようにネットワークをトレーニングすることもできます。

$$\begin{cases} E_p = -\sum_{i=1}^K \mathbf{p}_i^u \log \mathbf{p}_i^u, \\ \mathcal{L}_{pl} = -\sum_{\mathbf{x}^u \in \mathcal{Q}_u} \mathbf{1}\{E_p < \tau\} \cdot \hat{\mathbf{y}}^u \log \mathbf{p}_{\hat{\mathbf{y}}^u}^u, \end{cases} \quad (7)$$

どこ $\mathbf{1}\{E_p < \tau\}$ は、エントロピー合計値が次の場合に値が 1 になる指標関数です。 E_p エントロピー閾値より小さい τ 、それ以外の場合は 0 です。

(c)クロスドメインのサンプルクラスの調整を行います。最後に、ソース ドメインからターゲット ドメインへ、およびターゲット ドメインからソース ドメインへの両方で、機能センターとクラス センター間のクロスドメイン調整を行います。アライメントの前に、次のように示されるソースとターゲットのクラス中心を計算する必要があります。 $\mathbf{c}^s, \mathbf{c}^t \in \mathcal{R}^{K \times C}$ 。の k クラスの中心は、関連するメモリバンク内のそのクラスのすべての項目の平均であり、次のように定式化されます。

$$\begin{cases} \mathbf{c}_k^s = \frac{1}{N} \sum_{i=1}^N \mathcal{M}_{ki}^s, \\ \mathbf{c}_k^t = \frac{1}{N} \left(\sum_{i=1}^{N_l} \mathcal{M}_{ki}^l + \sum_{i=1}^{N_u} \mathcal{M}_{ki}^u \right). \end{cases} \quad (8)$$

ここで、トレーニングの開始時には、メモリ バンクを更新するための十分な機能がないこと、つまり一部のアイテムが初期のゼロ状態のままであることに言及する価値があります。干渉を防ぐため、この段階ではこれらのゼロ項目はクラス センターでは計算されません。

ソースとターゲットのクラス中心は、フィーチャの位置合わせのアンカーとして使用されます。 \mathbf{f}^s 、 \mathbf{f}^l 、そして \mathbf{f}^u 。全体の教師付きアライメント損失 $\mathcal{L}_{alg,s}$ ソース フィーチャのアライメント損失の 2 つの方向で構成されます $\mathcal{L}_{alg,s}^s$ およびターゲット フィーチャのアライメント損失

$\mathcal{L}_{alg,s}^t$ 、次のように計算されます。

$$\begin{cases} \mathcal{L}_{alg,s}^s = \frac{1}{M} \sum_{i=1}^M (\mathbf{f}_i^s - \mathbf{c}_{y_i^s}^t), \\ \mathcal{L}_{alg,s}^t = \frac{1}{M} \left[\sum_{i=1}^{M^l} (\mathbf{f}_i^l - \mathbf{c}_{y_i^l}^t) + \sum_{i=1}^{M^u} (\mathbf{f}_i^u - \mathbf{c}_{y_i^u}^s) \right], & \hat{y}_i^u > \tau, \\ \mathcal{L}_{alg,s} = \mathcal{L}_{alg,s}^s + \mathcal{L}_{alg,s}^t. \end{cases} \quad (9)$$

ここで、機能とラベルのペア $[\mathbf{f}^s, \mathbf{y}^s]$ 、 $[\mathbf{f}^l, \mathbf{y}^l]$ 、そして $[\mathbf{f}^u, \hat{\mathbf{y}}^u]$ はステップ (1) により提供される。Liらに続いて、(2021)のとき \mathbf{f}^u が位置合わせに使用される場合、RandAugment 技術によって同じターゲット画像の拡張部分から抽出された対応物に置き換えられます (Cubuk et al., 2020)。ミニバッチ内では、 M^u ラベルのないフィーチャの番号です \mathbf{f}^u 誰の疑似ラベル $\hat{\mathbf{y}}^u$ より大きい τ 、そして \mathbf{f}^u に最も近い $\mathbf{c}_{y_i^u}^s$ 測定値として平均二乗誤差 (MSE) を使用します。BSCA では、クラス センターは特徴のアンカーとしてのみ機能し、対応する生の画像への勾配逆伝播は行われません。その結果、このようなフィーチャクラスの調整では、トレーニング フェーズ中にほとんどコンピューティングリソースが消費されません。

3.3.3. BSCA の全体的な損失とトレーニング手順

BSCA ベースの SSDA モデルの全体的な損失は、教師付き分類損失の組み合わせです。 \mathcal{L}_{cls} 、疑似ラベル分類損失 \mathcal{L}_{pl} 、教師なし特徴アライメント損失 $\mathcal{L}_{alg,u}$ 、および監視されたクロスドメインフィーチャクラスのアライメント損失 $\mathcal{L}_{alg,s}$ 、次のように定式化されます

$$\mathcal{L} = \mathcal{L}_{cls} + \mathcal{L}_{pl} + \alpha \mathcal{L}_{alg,u} + \beta \mathcal{L}_{alg,s}, \quad (10)$$

どこ α として β の重みです $\mathcal{L}_{alg,u}$ として $\mathcal{L}_{alg,s}$ 。

私たちのメソッドのワークフローを明確に理解するために、BSCA のトレーニング手順の繰り返しを以下にまとめます。各反復で、画像の 3 つのミニバッチ、 \mathbf{x}^s 、 \mathbf{x}^l 、そして \mathbf{x}^u からランダムに抽出されます。 \mathcal{D}_s 、 \mathcal{D}_l 、そして \mathcal{D}_u 、それぞれ、共有特徴抽出器と特徴分類器に送信されます。したがって、グローバル機能の 3 つのミニバッチ (\mathbf{f}^s 、 \mathbf{f}^l 、そして \mathbf{f}^u) とクラスごとの予測 (\mathbf{p}^s 、 \mathbf{p}^l 、そして \mathbf{p}^u) が得られます。その中で、 \mathbf{p}^s として \mathbf{p}^l クロスエントロピー損失の計算に使用されます \mathcal{L}_{cls} ；信頼性の高い部分 \mathbf{p}^u 疑似ラベル分類損失の計算に使用されます \mathcal{L}_{pl} ； \mathbf{f}^u との連結 \mathbf{f}^s として \mathbf{f}^l 教師なしアライメント損失の計算に使用されます $\mathcal{L}_{alg,u}$ ； \mathbf{f}^s との連結 \mathbf{f}^l として \mathbf{f}^u 教師ありアライメント損失の計算に使用されます $\mathcal{L}_{alg,s}$ 。次に、上記の損失から式 (1) により全体の損失が得られます。(10) であり、勾配逆伝播とモデルの最適化に使用されます。ついに、 \mathbf{f}^s ソースメモリバンクを更新するために使用されます。 \mathbf{f}^l ラベル付きのターゲットメモリバンクと、その信頼性の高い部分を更新するために使用されます。 \mathbf{f}^u ラベルのないターゲットメモリバンクを更新するために使用されます。



[ダウンロード:高解像度画像のダウンロード\(2MB\)](#)

[ダウンロード:フルサイズの画像をダウンロード](#)

図3. 4 つの RS シーン データセットの 7 つの共通クラスの例。さまざまなスケールや角度など、異なるデータセット間でドメインのシフトがあることが観察できます。

4. 実験

このセクションでは、ベンチマーク データセットや実装の詳細など、いくつかの実験設定を紹介します。次に、提案された BSCA の境界を調査するために 2 種類のハイパーパラメータ調整実験が実行されます。続いて、提案された方法といくつかの最先端の SSL/UDA/SSDA 方法を比

較する実験が行われます。次に、BSCA の各コンポーネントの個別の効果を調査するためにアブレーション研究が実行されます。最後に、直感的な比較のために、BSCA と他の方法の特徴分布の視覚化が提供されます。

4.1 。実験設定

ベンチマーク データセット。提案された方法を包括的に評価するために、広く使用されている 4 つの RS 画像シーン分類データセット (RSSCN7 (Zou et al., 2015)、NWPU-RESISC45、AID、および WHU) から 7 つの共有クラスを含む RS-SSDA ベンチマークを収集して構築します。RS19。各データセットはソース ドメインとターゲット ドメインの両方として機能するため、合計 12 の適応シナリオがあります。データセットごとに、ソース ドメインとして機能する場合、すべての画像にクラス ラベルが付けられますが、ランダムのみが使用されます。 N_l 各クラスの画像にはラベルが付いており (ラベル付きターゲット ドメイン)、残りの画像はすべてラベルなし (ラベルなしターゲット ドメイン) です。ラベルのないターゲット データはすべて、教師なしの方法でモデルのトレーニングに使用されます。このうち、クラスあたり 6 つのサンプルが検証セットとして使用され、残りのすべてのデータがテスト セットとして使用されます。発売年、撮影解像度、元画像サイズ、クラス名、サンプル番号を表1に、その例を図3に示します。以下のセクションでは、NWPU-RESISC45、AID、RSSCN7、および WHU-RS19 をそれぞれ N、A、R、および W と省略します。

実装の詳細。私たちは、広く使用されている 3 つの高度な CNN バックボーン、VGG16 (Simonyan and Zisserman, 2014)、ResNet34 (He et al., 2016)、および EfficientNet_B0 (Tan and Le, 2019)に基づく実験を採用しています。これらはすべて ImageNet で事前トレーニングされています。(Deng et al., 2009)。それらの特徴マップは、式 1 のようにGAP層によって対応するグローバル特徴ベクトルに埋め込まれます。(2)フォローアップ SSDA 操作の場合、最後の完全接続層は特徴分類子によって置き換えられます。 \mathcal{L} 式で説明されています。(3)温度パラメータを使用した場合 T これは、既存の SSDA の研究 (Saito et al., 2019、Kim and Kim, 2020、Li et al., 2021) と一致する 0.05 に設定されています。ソースおよびターゲット メモリ バンクの項目数は 32 です。ラベル付きターゲット メモリ バンクのメモリ バンクは $2 \times N_l$ したがって、ラベルのないターゲット メモリ バンクのメモリ バンクが残ります ($32 - 2 \times N_l$)。方程式では、(7)、(9)、エントロピーしきい値 τ アブレーション、比較、視覚化実験のターゲット ドメインとして 4 つのデータセットすべてに対して 0.5 に設定されます。式では、(10)、教師なしアライメント損失の重み α および監視されたアライメント損失 β はそれぞれ 10 と 0.025 に設定されます。BSCA の上記の特定のハイパーパラメータに加えて、公平な比較のために、すべての一般的なトレーニング設定はこのホワイトペーパーで使用されている他の SSDA メソッドと同じに保たれています。具体的には、運動量 0.9 の確率的勾配降下法(SGD) がモデルをトレーニングするためのオプティマイザーとして使用されます。学習率は 0.01 で初期化され、ミニバッチを使用すると 0.005 の重み減衰で減少します。 M サイズは 24 に設定されます。すべてのメソッドのモデルは、12 の適応シナリオすべてで 2000 回の反復でトレーニングされ、50 回の反復ごとに検証セットによって検証されます。トレーニング段階では、検証パフォーマンスが最も優れたものが保存され、トレーニング後にテスト セットによってテストされます。

表 1。RS 画像シーン分類セットの特徴。

	NWPU-RESISC45	援助	RSSCN7	WHU-RS19
年	2017年	2016年	2015年	2010年
解像度(m)	0.2~30	0.5~8	-	0.5
画像サイズ	256×256	600×600	400×400	600×600
クラス 1:農地	700	370	400	50
クラス 2:森林	700	250	400	53
クラス 3:産業用	700	390	400	53
クラス 4:牧草地	700	280	400	61
クラス5:駐車場	700	390	400	50
クラス 6:住宅	700	410	400	54
クラス 7:川	700	410	400	56

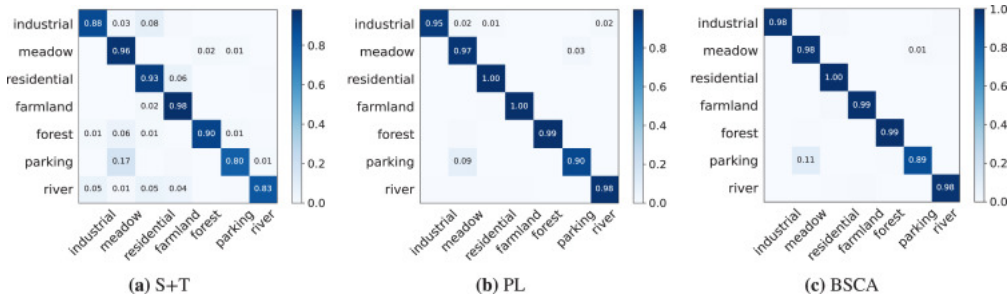
実験は、2 つの GeForce RTX 2080Ti GPU 上の PyTorch 1.9.11 (Paszke et al., 2017)に基づいて実装されています。

4.2 。アブレーション研究

提案された BSCA の目的関数全体には、監視分類損失の 4 つの要素があります。 \mathcal{L}_{cls} ラベル付きデータの場合、擬似ラベル分類損失 \mathcal{L}_{pl} 信頼性の高いラベルなしデータ、教師なしアライメント損失の場合 \mathcal{L}_{alg_u} 教師なしクロスドメイン特徴と教師ありサンプルクラスのアライメント損失の間 \mathcal{L}_{alg_s} クラス認識/疑似クラス認識機能の中にあります。知識伝達に対するそれぞれの効果を検証するために、このサブセクションでは、ResNet34 を CNN バックボーンとして使用した 12 の適応シナリオすべてに対して BSCA のアブレーション研究を実装します。実験結果を表 2 に示す。

表 2。CNN バックボーンとしてResNet34を使用した 3 ショット設定でのアブレーション研究の分類精度 (%)。

\mathcal{L}_{cls}	\mathcal{L}_{pl}	\mathcal{L}_{alg_s}	\mathcal{L}_{alg_u}	あ→N	R→N	W→N	N→あ	R→あ	W→あ	N→R	あ→R	W→R	N→W	あ→W	R→W	平均
✓				90.7	77.4	84.3	93.3	86.0	94.2	75.9	72.5	72.9	89.9	98.3	92.5	85.7
✓	✓			94.6	86.5	94.2	95.9	94.7	97.2	78.9	79.0	81.6	97.6	97.9	93.4	91.0 (+5.3)
✓	✓	✓		91.9	89.8	93.2	96.7	96.0	96.7	80.5	79.9	82.3	97.9	99.0	96.9	91.7 (+6.0)
✓	✓		✓	96.6	89.0	94.5	96.8	96.8	97.6	81.6	81.9	84.1	97.9	99.0	96.2	92.7 (+7.0)
✓	✓	✓	✓	96.8	89.7	94.7	97.2	97.2	97.5	84.1	82.6	84.0	97.9	99.7	98.3	93.3 (+7.6)



ダウンロード:高解像度画像をダウンロード (418KB)

ダウンロード:フルサイズの画像をダウンロード

図4. S+T, PL, および N で提案されている BSCA の手法の混同行列→CNN バックボーンとして ResNet34 を使用した適応シナリオ。

のみ使用する場合 \mathcal{L}_{cls} 、12 の適応シナリオすべてにわたって 85.7% の平均分類精度が得られます。埋め込み後 \mathcal{L}_{pl} 、ラベルのないターゲットデータの組み込みにより、平均精度は 91.7% まで劇的に増加します。これにより、分類境界のエントロピーが減少し、ソースドメインの過剰な影響が軽減されます。さらに、両方とも \mathcal{L}_{alg_s} として \mathcal{L}_{alg_u} をベースにさらなるパフォーマンス向上に貢献します。 $\mathcal{L}_{cls} + \mathcal{L}_{pl}$ 程度が異なります。追加するだけであることがわかります \mathcal{L}_{alg_s} の 91.0% に対して 0.7% のゲインを得ることができます。 $\mathcal{L}_{cls} + \mathcal{L}_{pl}$ おそらく、誤って疑似ラベルが付けられたターゲットデータによって制限されます。幸いなことに、と組み合わせた後、 \mathcal{L}_{alg_u} 、個人の改善 \mathcal{L}_{alg_s} さらに 0.6% 上昇するため、最終的な BSCA では 93.3% という最良の結果が得られ、これはベースラインより 7.6% の利点があります。 \mathcal{L}_{cls} 同様に 2.3% 以上 $\mathcal{L}_{cls} + \mathcal{L}_{pl}$ 追加のモデルパラメータを追加することなく、アブレーションの結果は、提案された BSCA の各コンポーネントの有効性を示しています。

疑似ラベルトレーニングでは、必然的にノイズの多いラベルが生成されます。 \mathcal{D}_u 分類パフォーマンスを向上させながら。ただし、アブレーションの結果は、特徴アライメント段階では、ターゲットドメインのクラス中心が BSCA のこれらのノイズ疑似ラベルによって悪影響を受けないことを明らかにしています。幸いなことに、この問題を軽減できる 5 つの要素があります。(1) ソースドメインとターゲットドメインは同じクラス空間を共有するため、ソースドメインでトレーニングされたモデルはターゲットドメインでもある程度使用できます。(2) 信頼性の高いラベルのないターゲットサンプルのみを疑似ラベルトレーニングに使用できます。これにより、誤って疑似ラベルが付けられたターゲットサンプルの影響を大幅に軽減できます。(3) ラベル付けされたターゲットサンプルから抽出された特徴は、ターゲットメモリバンク内の一定数のアイテムを占め、特に初期段階では純粋なサンプルが提供されます。(4) ターゲットメモリバンク内の多数の項目は、平均的な演算によって少数のノイズフィーチャの干渉を弱めることができます。したがって、ターゲットメモリバンクの更新プロセスは、フォローアップの調整に十分堅牢です。(5) BSCA の UA モジュールは、ソースドメインとラベルのないターゲットドメインの間でクラスに関係のないドメインレベルのアライメントを実現するため、後者の分布をより安定させ、ノイズ疑似ラベルの影響を少なくすることができます。

S+T の混同行列 (\mathcal{L}_{cls})、PL ($\mathcal{L}_{cls} + \mathcal{L}_{pl}$)、完全な BSCA を図 4 に示します。特に「住宅」のクラスでは、コンポーネントを追加するほど、より明確な混同行列が得られることがわかり、アルゴリズム設計の有効性がさらに検証されます。

4.3. ハイパーパラメータ調整実験

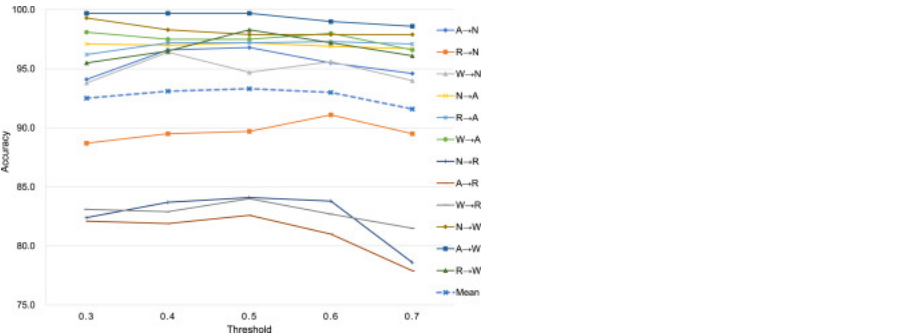
BSCA のパフォーマンス境界をさらに調査するために、このサブセクションでは、エントロピーしきい値調整とショット数 (クラスごとのラベル付きサンプルの数) 調整という 2 種類のハイパーパラメータ調整実験が実行されます。

エントロピー閾値調整の実験結果を表 3 に報告し、図 5 にプロットする。0.3 は低いいため、トレーニングに使用できる疑似ラベル付きターゲットサンプルの数が比較的少なくなり、モデルの適合が不十分になります。対照的に、0.7 は非常に高いため、ラベルのないデータが多すぎると疑似ラベルが付けられ、自己学習され、誤って疑似ラベルが付けられたデータの過剰適合につながります。全体として、図 5 の傾向から、0.5 が最良の平均分類パフォーマンスを達成していることがわかります。次の実験では、エントロピーのしきい値はデフォルトで 0.5 に設定されています。

一方、ショット数チューニングの実験結果を表 4 および図 6 に示す。一般に、特に基本性能 (1 ショット) が 95% 未満の場合、ショット数の増加に応じて分類精度は向上します。ショット数が 3 から 10 に増加すると、平均精度は 0.7% 増加し、パフォーマンスの向上は限られています。実際の応用の観点から、我々は 3 ショット設定での実験に焦点を当てますが、これはいくつかの古典的な SSDA 作業の実践にも準拠しています (Saito et al., 2019, Li et al., 2021)。

表 3. CNN バックボーンとして**ResNet34**を使用した 3 ショット設定での閾値調整実験の分類精度 (%)。

Threshold	A→N	R→N	W→N	N→A	R→A	W→A	N→R	A→R	W→R	N→W	A→W	R→W	Mean
0.3	94.1	88.7	93.8	97.1	96.2	98.1	82.4	82.1	83.1	99.3	99.7	95.5	92.5
0.4	96.6	89.5	96.4	97.0	97.2	97.5	83.7	81.9	82.9	98.3	99.7	96.5	93.1
0.5	96.8	89.7	94.7	97.2	97.2	97.5	84.1	82.6	84.0	97.9	99.7	98.3	93.3
0.6	95.5	91.1	95.6	96.9	97.3	98.0	83.8	81.0	82.7	97.9	99.0	97.2	93.0
0.7	94.6	89.5	94.0	96.7	97.1	96.6	78.6	77.9	81.5	97.9	98.6	96.1	91.6

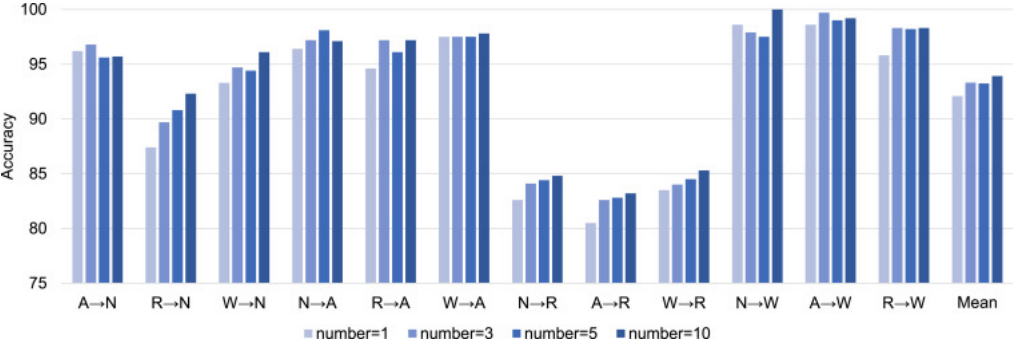


[ダウンロード:高解像度画像のダウンロード \(264KB\)](#)
[ダウンロード:フルサイズの画像をダウンロード](#)

図5. ResNet34 に基づいて提案された BSCA のさまざまなエントロピーしきい値による分類精度。

表 4. CNN バックボーンとして**ResNet34**を使用した、クラスごとのさまざまな数のラベル付きサンプルの分類精度 (%)。

Shot number	A→N	R→N	W→N	N→A	R→A	W→A	N→R	A→R	W→R	N→W	A→W	R→W	Mean
1	96.2	87.4	93.3	96.4	94.6	97.5	82.6	80.5	83.5	98.6	98.6	95.8	92.1
3	96.8	89.7	94.7	97.2	97.2	97.5	84.1	82.6	84.0	97.9	99.7	98.3	93.3
5	95.6	90.8	94.4	98.1	96.1	97.5	84.4	82.8	84.5	97.5	99	98.2	93.2
10	95.7	92.3	96.1	97.1	97.2	97.8	84.8	83.2	85.3	100.0	99.2	98.3	93.9



[ダウンロード:高解像度画像をダウンロード \(407KB\)](#)
[ダウンロード:フルサイズの画像をダウンロード](#)

図6. クラスごとにラベル付けされたターゲット サンプルの数が異なる場合の分類精度。

4.4。比較実験

提案されているRS-SSDAベンチマークの12種類の適応シナリオに基づいて、RSクロスドメインシーン分類に対するBSCAの有効性を公平かつ包括的に評価するために、私たちの手法といくつかの最先端のUDAとの比較実験を実施します。、SSL、および3つの一般的なCNNバックボーン上のSSDAアルゴリズム。

4.4.1。ベースライン

RSクロスドメインシーン分類におけるSSL/UDAに対するSSDAの利点を検証するために、BSCAを以下と比較します。(1) FixMatch (Xiong et al., 2021, Sohn et al., 2020)、1つの古典的なSSL手法。半教師ありRSシーン分類に適用され、この分野で最先端の結果の1つが達成されました。(2) ECB-FAM (Ma et al., 2021b)、RSクロスドメインシーン分類用の最新のUDAアルゴリズムの1つ。より包括的な評価のために、BSCAをこの研究で再現したいいくつかのSSDA手法とさらに比較します。(3) S+Tは、トレーニングにソースデータとラベル付きターゲットデータのみを使用する基本的な手法です。(4) ADDA (Wang et al., 2018, Tzeng et al., 2017)。これは、我々の知る限りではRS-SSDAに関する最初の研究です。2つの個別の抽出器と共有分類器に基づいて、ソースデータとターゲットデータを敵対的にトレーニングします。(5) RevGrad (Lasloun et al., 2021, Lu et al., 2019)。これはRSクロスドメインシーン分類に一般的に適用され、DANNの方法 (Ganin et al.,

2016、Saito et al., 2016)から開発されました。、2019年)。勾配反転レイヤーを介して、ソース フィーチャとターゲット フィーチャのフィーチャ分類器の識別を逆に低下させる可能性があります。(6)SSDAN (Lasloum et al., 2021)。これは、MME の方法 (Saito et al., 2019)から派生した RS-SSDA の現在の最先端のアプローチです。あるいは、特徴分類器のラベルなしターゲット データの条件付きエントロピーを最大化し、特徴抽出器の条件付きエントロピーを最小化します。(7)ENT (Grandvalet et al., 2005、Saito et al., 2019)。これは、ユニバーサル SSDA 分野で広く適用されている古典的な半教師あり学習方法 (Grandvalet et al., 2005) です。より信頼性の高い分類のために分類境界のエントロピーを減少させることを目的としています。(8)CDAC (Li et al., 2021)。これはコンピュータ ビジョンにおける最先端の SSDA 手法の 1 つであり、本稿では比較のために RS-SSDA に適用されています。ラベルのないターゲット データの特徴をドメイン間のクラスターにグループ化することにより、ドメイン間とドメイン内の両方の適応に焦点を当てます。

ここでは、公式コード (S+T & ENT & SSDAN、² ADDA、³ RevGrad、⁴および CDAC⁵) に基づいてすべての SSDA メソッドを実装します。CNN モデルに対するこれらの手法の堅牢性は、広く使用されている 3 つの CNN バックボーンで評価されます。(1) VGG16 (Simonyan および Zisserman, 2014)。これは、複雑な視覚表現のための古典的なディープ CNN アーキテクチャです。(2) ResNet34 (He et al., 2016)。勾配消失の問題を回避するためにスキップ残差接続を利用します。(3) EfficientNet_B0 (Tan and Le, 2019)。これはニューラル アーキテクチャ探索技術によって設計されており、モデルパラメータを効率的に削減します。

表5. CNN バックボーンとしてVGG16を使用した 3 ショット設定での RS-SSDA データセットからの 12 の適応シナリオに関する比較実験の分類精度 (%)。最良の結果は太字で示されています。

タイプ	方法	あ→N	R→N	W→N	N→あ	R→あ	W→あ	N→R	あ→R	W→R	N→W	あ→W	R→W	平均
SSL	FixMatch (Xiong et al., 2021)		77.9			96.3			67.5			93.4		83.8
宇田	ECB-FAM (Ma et al., 2021b)	90.5	74.6	80.6	93.9	79.5	92.3	69.7	71.0	68.2	94.1	97.7	87.6	83.3
SSDA	S+T	92.6	83.6	88.0	94.1	88.9	93.8	72.9	72.4	76.9	96.9	98.3	92.4	87.6
	ADDA (Wang et al., 2018)	81.8	69.2	73.9	90.5	82.4	77.9	61.8	64.8	66.7	86.1	89.6	80.9	77.1
	RevGrad (Lasloum et al., 2021)	88.2	81.7	90.0	93.9	88.9	89.8	73.8	71.9	77.5	96.9	98.3	93.4	87.0
	SSDAN (ラスルーム他, 2021)	94.6	84.7	89.5	96.8	94.2	95.5	79.6	80.6	81.8	98.3	99.3	96.9	91.0
	耳鼻咽喉科 (Grandvalet et al., 2005)	86.5	72.0	86.9	92.2	86.0	91.7	60.4	68.0	73.0	80.2	98.6	89.9	82.1
	CDAC (リーら, 2021)	85.6	74.6	71.7	89.9	81.8	79.5	65.0	70.2	69.8	81.2	89.2	73.6	77.7
	私たちのBSCA	96.5	94.2	95.9	96.5	94.5	96.5	78.0	82.8	84.7	98.6	99.0	99.3	93.0

表6. CNN バックボーンとしてResNet34を使用した 3 ショット設定での RS-SSDA データセットからの 12 の適応シナリオの比較実験の分類精度 (%)。最良の結果は太字で示されています。

タイプ	方法	あ→N	R→N	W→N	N→あ	R→あ	W→あ	N→R	あ→R	W→R	N→W	あ→W	R→W	平均
SSL	FixMatch (Xiong et al., 2021)		81.3			95.0			63.5			97.6		84.4
宇田	ECB-FAM (Ma et al., 2021b)	88.5	74.5	83.2	92.5	77.5	92.8	66.9	68.9	65.4	91.7	97.6	79.9	81.6
SSDA	S+T	89.9	76.4	84.4	91.7	84.5	91.6	71.4	68.4	69.4	92.0	97.6	91.0	84.0
	ADDA (Wang et al., 2018)	87.3	78.6	83.1	77.2	76.5	90.7	64.7	72.4	65.0	89.9	94.8	85.8	80.5
	RevGrad (Lasloum et al., 2021)	90.2	80.2	86.9	93.5	86.8	91.7	76.7	73.6	74.1	95.5	98.3	88.5	86.3
	SSDAN (ラスルーム他, 2021)	95.5	86.6	93.5	96.9	95.0	97.3	80.2	80.2	83.6	97.3	99.3	96.2	91.8
	耳鼻咽喉科 (Grandvalet et al., 2005)	79.5	67.7	77.2	91.2	91.2	87.5	66.6	60.5	62.6	89.9	91.0	83.7	85.4
	CDAC (リーら, 2021)	85.4	80.6	75.2	83.8	90.0	88.5	65.8	65.5	72.7	86.1	80.6	92.0	79.8
	私たちのBSCA	96.8	89.7	94.7	97.2	97.2	97.5	84.1	82.6	84.0	97.9	99.7	98.3	93.3

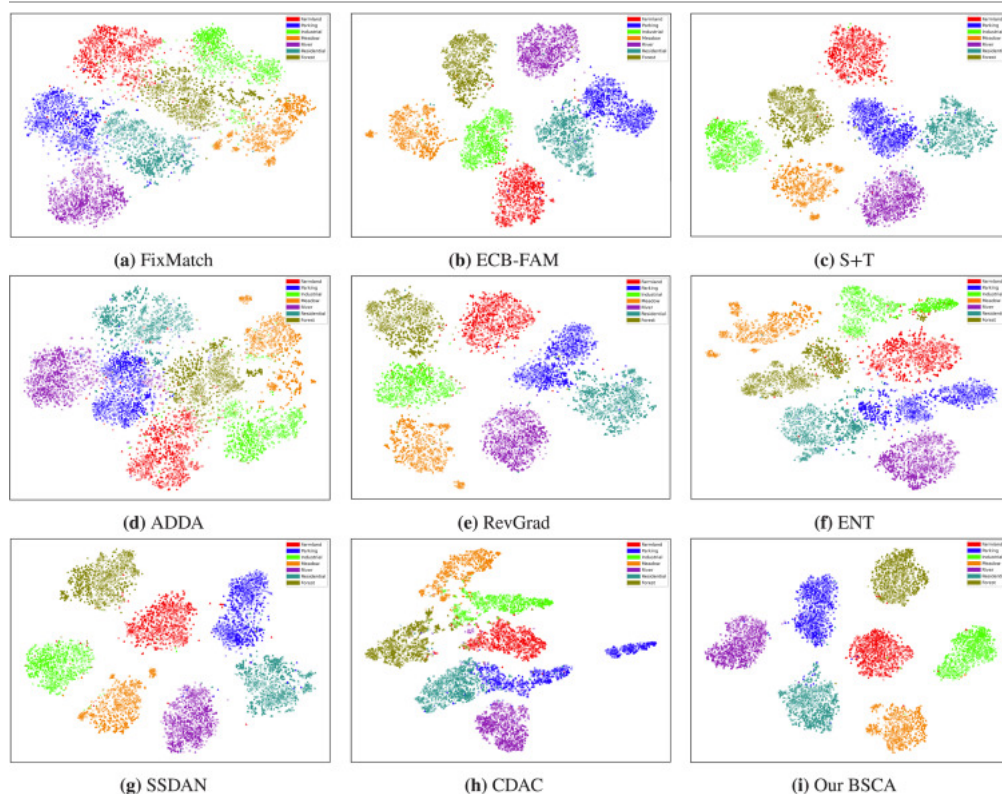
表7. CNN バックボーンとしてEfficientNet_B0を使用した 3 ショット設定での RS-SSDA データセットからの 12 の適応シナリオの比較実験の分類精度 (%)。最良の結果は太字で示されています。

タイプ	方法	あ、N	R、N	W、N	N、あ	R、あ	W、あ	N、R	あ、R	W、R	N、W	あ、W	R、W	平均
SSL	FixMatch (Xiong et al., 2021)		72.2			93.2			63.0			96.9		81.3
宇田	ECB-FAM (Ma et al., 2021b)	91.7	76.8	83.0	91.8	80.5	90.8	68.6	69.5	62.4	92.4	98.6	84.7	82.6
SSDA	S+T	91.1	81.5	86.0	92.6	84.2	91.4	71.4	71.9	66.2	93.1	97.6	90.4	84.8
	ADDA (Wang et al., 2018)	90.2	77.7	84.3	89.1	79.8	90.8	69.8	69.9	69.3	94.1	96.6	83.3	82.9

タイプ	方法	あ→N	R→N	W→N	N→あ	R→あ	W→あ	N→R	あ→R	W→R	N→W	あ→W	R→W	平均
	RevGrad (Lasloum et al., 2021)	90.9	78.4	84.5	92.7	84.4	91.5	72.8	69.6	72.0	94.1	99.0	89.2	84.9
	SSDAN (ラスルーム他, 2021)	90.6	79.6	83.7	90.8	86.0	89.6	73.1	73.8	72.0	96.5	99.0	87.2	85.2
	耳鼻咽喉科 (Grandvalet et al., 2005)	88.3	78.2	84.8	91.5	83.2	90.5	70.7	68.2	71.9	96.5	97.9	91.0	84.4
	CDAC (リーら, 2021)	87.0	84.7	86.0	88.5	78.6	84.8	68.5	66.3	68.5	88.9	92.7	81.2	81.3
	私たちのBSCA	96.1	88.3	93.8	96.5	95.8	96.4	77.0	78.9	83.0	98.3	98.6	98.6	91.8

4.4.2. 最先端の手法との比較

VGG16、ResNet34、および EfficientNet_B0 の比較結果をそれぞれ表 5、6、および 7 に報告します。全体として、提案された BSCA は、RS-SSDA ベンチマーク データセットの 12 の適応シナリオの平均分類精度に関して最高のパフォーマンスを達成します。BSCA は、VGG16、ResNet34、および EfficientNet_B0 でそれぞれ 93.0%、93.3%、および 91.8% の平均精度を獲得し、SSDAN の準最適な方法と比べて 2.0%、1.5%、および 6.6% のパフォーマンス上の利点があります。



ダウンロード:高解像度画像のダウンロード (3MB)

ダウンロード:フルサイズの画像をダウンロード

図 7. 提案された BSCA の手法と N の他の比較手法の特徴分布の t-SNE 視覚化→CNN バックボーンとして ResNet34 を使用した適応シナリオ。記号「x」そして「▲」は、それぞれソースサンプルとターゲットサンプルから抽出された特徴を表します。

CNN バックボーンの観点から見ると、ResNet34 のパフォーマンスが最高で、VGG16 がそれにわずかに遅れ、次に EfficientNet_B0 が続きます。ここで、EfficientNet_B0 が最悪のパフォーマンスを示していることに言及する価値があります。これは、トレーニングに十分なラベル付きデータがない場合、その縮小メカニズムによって制限される可能性があります。少数の焦点を当てたオブジェクトを含む自然画像とは異なり、RS シーン画像には地上の多数のオブジェクトが含まれます。したがって、アーキテクチャの縮小により、知識伝達プロセス中に RS 画像の一部の有用な情報が破棄される可能性があります。さらに、これらの CNN バックボーンをより包括的に評価するために、表 8 に示すように、パラメーター、計算複雑さ (積和演算、MAC によって評価)、および平均分類精度を含む 3 つの観点から BSCA を比較します。表から、EfficientNet_B0 のパラメーターと計算の複雑さが最小限であるため、パフォーマンスが比較的低いことがわかります。対照的に、ResNet34 は、効率的な残留接続と最大のモデルパラメーターの恩恵を受けて、中程度の計算複雑さで最高のパフォーマンスを達成します。

方法の種類の側面から見ると、FixMatch の SSL 方法は、UDA および SSDA よりも劇的なギャップを伴って最悪のパフォーマンスを示し、ドメイン適応戦略が半教師あり学習戦略よりも RS シーン画像のターゲットドメイン分類にとって価値があることを証明しています。ECB-FAM の UDA 方式の場合、そのパフォーマンスはソースドメインの影響を受けます。たとえば、VGG16 上の同じターゲットドメイン NWPU-RESISC45 の場合、精度は 90.5% (AID がソースドメイン) から 74.6% (RSSCN7 がソースドメイン) に低下し、その差は 15.9% です。対照的に、SSL と UDA を組み合わせた実験では、ほとんどの SSDA メソッドは、ソースドメインの影響を受けにくく、より優れた安定したパフォーマンスを示します。考えられる理由は、SSDA が UDA と SSL の利点を組み合わせているため、より堅牢になっているためです。

表8. さまざまな CNN モデルに基づく、パラメーター、計算の複雑さ、および 12 の適応シナリオの平均精度における BSCA の比較。

モデル	パラムス(男)↓	MAC (G) ↓	平均精度 (%)↑
VGG16	14.98	15.35	93.0
レスネット34	21.55	3.68	93.3
EfficientNet_B0	4.67	0.41	91.8

SSDA メソッドの観点から見ると、当社の BSCA は、テストされたすべての SSDA メソッドの中で最高かつ最も安定したパフォーマンスを実現します。S+T の基本的な方法と比較すると、SSDAN は ResNet34 および VGG16 に対して大幅な改善を示しています。ただし、SSDAN は EfficientNet_B0 でわずかなゲインしか得られず、異なる CNN バックボーンに直面した場合の不安定性が明らかになります。ADDA は、おそらくその独自の分離されたソース抽出器とターゲット抽出器のせいでパフォーマンスが低く、これはドメイン間でのモデル共有の重要性を明らかにしています。RevGrad はパフォーマンスに利点を示しておらず、ENT は一般にクロスドメイン分類にとって有害です。CDAC は RS-SSDA でパフォーマンスが悪くなります。これはおそらく、RS シーン クラス間でオブジェクトとフィーチャの重複が多いため、ラベルなしのフィーチャ類似性に基づく CDAC のクラスタリング戦略が無効になるためです。[この現象は、セクション4.2](#)で提供される特徴分布で直感的に観察できます。

特定の適応シナリオの観点から、これらの実験の結果は、WHU がターゲット ドメインとして機能する一部のシナリオを除いて、BSCA がほとんどのシナリオで最良の知識伝達を達成できることを示しています。その理由としては、サンプル数が少ないため、クロスドメイン アライメントの有効性とさらなる分類パフォーマンスが制限されることが考えられます。

一般に、実験結果は、SSL、UDA、およびその他の最先端の SSDA 方法と比較して、さまざまな RS 適応シナリオおよび CNN バックボーンに直面したときの、提案された BSCA の有効性、優位性、および安定性を示しています。

4.5。特徴分布の可視化

提案された BSCA の特徴調整効果を直感的に示すために、[図 7](#)では、t-SNE 手法 ([Van der Maaten および Hinton, 2008](#)) を使用して、N の適応シナリオにおける BSCA の特徴分布と他の比較方法を示しています。→ResNet34 に基づく。

FixMatch の SSL メソッドは、ターゲット機能を大まかに集約しますが、ソース機能の利用を無視するため、明らかなクロスドメインのセマンティック ギャップが生じます。ECB-FAM の UDA メソッドは、ドメイン間の不一致を減少させますが、「住宅」や「工業」などの一部のクラスではクラス境界が明確ではありません。SSDA 手法の場合、S+T のベースラインと比較して、RevGrad はドメイン シフトの削減にほとんど効果を示さず、ENT のエントロピー最小化操作はクロスドメインのクラス内機能を分離するというマイナスの役割を果たします。これは、その[分類結果](#)。SSDAN は、クロスドメインのクラス内分散に干渉することなく、クラスを相互に遠ざけることで適切に機能します。CDAC は RS-SSDA ではパフォーマンスが悪く、これはおそらく 2 つの画像に含まれるオブジェクトとフィーチャの類似性に基づく、ラベルのないターゲット フィーチャの適応[クラスタリング手法の影響](#)を受けています。ただし、建物や道路など、さまざまな RS シーン クラス間で重複するオブジェクトやフィーチャが多数あります。[図 7 \(h\)](#) に示すように、「工業」と「住宅」の特徴が大きく混在しており、教師なし[特徴クラスタリング](#)により特徴分布が不当に近くなっていることがわかる。対照的に、[図 7\(i\)](#)は、BSCA がクラス間の識別を増加させるだけでなく、クロスドメインのクラス内距離、つまりクラス内分散を減少させることができるため、最高の知識伝達パフォーマンスを達成できることを示しています。間違った擬似ラベルが付けられたターゲットサンプルがいくつかあるにもかかわらず。全体として、この視覚化は、BSCA がクロスドメイン機能調整を介してソース ドメインからターゲット ドメインへの知識の伝達を達成できることを示しています。

5. 結論

この論文では、リモート センシング (RS) クロスドメイン シーン分類のための半教師ありドメイン アダプテーション (SSDA) を目的として、ソース ドメインとターゲットの間の意味上の不一致を軽減する双方向サンプルクラス アライメント (BSCA) 方法を設計します。ドメイン。BSCA は、教師なしアライメント (UA) と教師ありアライメント (SA) の 2 種類のアライメント メカニズムで構成されます。まず、UA モジュールは、ソース ドメインとラベルなしターゲット ドメイン間だけでなく、ラベル付きターゲット ドメインとラベルなしターゲット ドメイン間でも、教師なしの方法でドメイン間の特徴アライメントに集中します。これは、それぞれの MMD を低減することによって実現されます。距離。次に、SA モジュールは、次のサブステップによってクラス レベルでのドメイン シフトをさらに削減することを目的としています。メモリ バンクから 2 つのドメインのクラスごとの中心を計算します。ソース サンプルからその関連するターゲット クラス センターへ、およびターゲット サンプルからその関連するソース クラス センターへの両方で、サンプルとクラスの特徴の調整を行います。サンプルから抽出された特徴を 2 つのドメインのクラスごとのメモリ バンクに保存します。いくつかの最先端の SSL/UDA/SSDA 手法と比較して、当社の BSCA は、コンパクトなクロスドメインのクラス内特徴表現と低エントロピー分類境界により、ラベルのないターゲット ドメインで優れた分類パフォーマンスを実現します。このような結果は、提案された方法の有効性と堅牢性を示しています。

今後の作業では、BSCA で誤った擬似ラベルが付けられたターゲット サンプルに起因するパフォーマンスのボトルネックを考慮して、カリキュラム学習などの動的なしきい値技術を介して擬似ラベルの品質を向上させることに焦点を当てます。

競合関係の宣言

著者らは、この論文で報告されている研究に影響を与えた可能性がある既知の競合する経済的利益や個人的関係を持っていないことを宣言します。

謝辞

この研究は、欧州連合の Horizon 2020研究革新プログラム(助成契約番号 [[ERC-2016-StG-714087](#)]、頭字語: *So2Sat*)、ヘルムホルツ優秀教授制度「地球観測におけるデータサイエンス - 都市研究のためのビッグデータ融合」の枠組みを通じてヘルムホルツ協会により作成(助成番号: [W2-W3-100](#))、国際未来 AI ラボ「AI4EO - 地球観測のための人工知能: 推論、不確実性、倫理およびその先」(認可番号: [01DD20001](#))の枠組みにおけるドイツ連邦教育研究省(BMBF)およびドイツ連邦省による「National Center of Excellence ML4Earth」の枠組みにおける経済問題と気候変動対策のための助成金(助成金番号: [50EE2201C](#))。

[おすすめ記事](#)

参考文献

[アーメッドら](#), 2021 Ahmed, N., Saha, S., Shahzad, M., Fraz, MM, Zhu, XX, 2021 年。衛星画像での森林マッピングのための漸進的教師なし深層転移学習。掲載: コンピューター ビジョンに関する IEEE/CVF 国際会議の議事録。752-761ページ。

[Google スカラー](#) [↗](#)

[アロンソほか](#), 2021 Alonso, I., Sabater, A., Ferstl, D., Montesano, L., Murillo, AC, 2021 年。クラスごとのメモリ バンクからのピクセルレベルの対比学習による半教師ありセマンティック セグメンテーション。掲載: コンピューター ビジョンに関する IEEE/CVF 国際会議の議事録。8219-8228ページ。

[Google スカラー](#) [↗](#)

[Bai 他](#), 2022 Bai L., Liu Q., Li C., Ye Z., Hui M., Jia X.

オクターブ畳み込みによるマルチスケール特徴融合共分散ネットワークを使用したリモート センシング画像シーン分類
IEEEトランス。地理学。遠隔監視員(2022)

[Google スカラー](#) [↗](#)

[ボルグヴァルトら](#), 2006 ボルグヴァルト K.M., グレットン A., ラッシュ M.J., クリーゲル H.P., シェルコフ B., スモーラ A.J.
カーネル最大平均不一致による構造化生物学的データの統合
バイオインフォマティクス, 22 (2006), pp. e49 - e57

[クロスリファレンス](#) [↗](#) [Scopus で表示](#) [↗](#) [Google スカラー](#) [↗](#)

[チェン他](#), 2019 チェン W.Y., リウ・Y.C., キラ Z., ワン Y.C.F., ファン J.B.

少数ショットの分類を詳しく見る

(2019年)

arXiv preprint [arXiv:1904.04232](#) [↗](#)

[Google Scholar](#) [↗](#)

[Cheng et al.](#), 2017 Cheng G., Han J., Lu X.

Remote sensing image scene classification: Benchmark and state of the art

Proc. IEEE, 105 (2017), pp. 1865-1883

[View in Scopus](#) [↗](#) [Google Scholar](#) [↗](#)

[Cubuk et al.](#), 2020 Cubuk, E.D., Zoph, B., Shlens, J., Le, Q.V., 2020. Randaugment: Practical automated data augmentation with a reduced search space. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. pp. 702-703.

[Google Scholar](#) [↗](#)

[Deng et al.](#), 2009 Deng J., Dong W., Socher R., Li L.J., Li K., Fei-Fei L.

Imagenet: A large-scale hierarchical image database

2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE (2009), pp. 248-255

[View in Scopus](#) [↗](#) [Google Scholar](#) [↗](#)

[Ganin et al.](#), 2016 Ganin Y., Ustinova E., Ajakan H., Germain P., Larochelle H., Laviolette F., Marchand M., Lempitsky V.

Domain-adversarial training of neural networks

J. Mach. Learn. Res., 17 (2016)

2096-2030

[Google Scholar](#) [↗](#)

[Grandvalet et al.](#), 2005 Grandvalet Y., Bengio Y., et al.

Semi-supervised learning by entropy minimization

CAP, 367 (2005), pp. 281-296

[Google Scholar](#) [↗](#)

[Gretton et al., 2012a](#) Gretton A., Borgwardt K.M., Rasch M.J., Schölkopf B., Smola A.

A kernel two-sample test

J. Mach. Learn. Res., 13 (2012), pp. 723-773

[View in Scopus](#) [Google Scholar](#)

[Gretton et al., 2012b](#) Gretton A., Sejdinovic D., Strathmann H., Balakrishnan S., Pontil M., Fukumizu K., Sriperumbudur B.K.

Optimal kernel choice for large-scale two-sample tests

Advances in Neural Information Processing Systems, Citeseer (2012), pp. 1205-1213

[View in Scopus](#) [Google Scholar](#)

[Han et al., 2018](#) Han W., Feng R., Wang L., Cheng Y.

A semi-supervised generative framework with deep learning features for high-resolution remote sensing image scene classification

ISPRS J. Photogramm. Remote Sens., 145 (2018), pp. 23-43

 [View PDF](#) [View article](#) [View in Scopus](#) [Google Scholar](#)

[He et al., 2020](#) He,K., Fan,H., Wu,Y., Xie,S., Girshick,R., 2020. Momentum contrast for unsupervised visual representation learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9729–9738.

[Google Scholar](#)

[He et al., 2016](#) He,K., Zhang,X., Ren,S., Sun,J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778.

[Google Scholar](#)

[Jiang et al., 2020](#) Jiang P., Wu A., Han Y., Shao Y., Qi M., Li B.

Bidirectional adversarial training for semi-supervised domain adaptation

IJCAI (2020), pp. 934-940

[CrossRef](#) [View in Scopus](#) [Google Scholar](#)

[Kim and Kim, 2020](#) Kim T., Kim C.

Attract, perturb, and explore: Learning a feature alignment network for semi-supervised domain adaptation

European Conference on Computer Vision, Springer (2020), pp. 591-607

[Google Scholar](#)

[Lasloun et al., 2021](#) Lasloun T., Alhichri H., Bazi Y., Alajlan N.

Ssdan: Multi-source semi-supervised domain adaptation network for remote sensing scene classification

Remote Sens., 13 (3861) (2021)

[Google Scholar](#)

[Li and Hospedales, 2020](#) Li D., Hospedales T.

Online meta-learning for multi-source and semi-supervised domain adaptation

European Conference on Computer Vision, Springer (2020), pp. 382-403

[CrossRef](#) [View in Scopus](#) [Google Scholar](#)

[Li et al., 2021](#) Li,J., Li,G., Shi,Y., Yu,Y., 2021. Cross-domain adaptive clustering for semi-supervised domain adaptation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2505–2514.

[Google Scholar](#)

[Lin et al., 2013](#) Lin M., Chen Q., Yan S.

Network in network

Comput. Sci. (2013)

[Google Scholar](#)

[Long et al., 2015](#) Long M., Cao Y., Wang J., Jordan M.

Learning transferable features with deep adaptation networks

International Conference on Machine Learning, PMLR (2015), pp. 97-105

[View in Scopus](#) [Google Scholar](#)

[Lu et al., 2019](#) Lu X., Gong T., Zheng X.

Multisource compensation network for remote sensing cross-domain scene classification

IEEE Trans. Geosci. Remote Sens., 58 (2019), pp. 2504-2515

[Google Scholar](#)

[Lu et al., 2020](#) Lu X., Gong T., Zheng X.

Multisource compensation network for remote sensing cross-domain scene classification

IEEE Trans. Geosci. Remote Sens., 58 (2020), pp. 2504-2515, [10.1109/TGRS.2019.2951779](#) [↗](#)

[View in Scopus](#) [↗](#) [Google Scholar](#) [↗](#)

[Lv et al., 2021](#) Lv,H., Chen,C., Cui,Z., Xu,C., Li,Y., Yang,J., 2021. Learning normal dynamics in videos with meta prototype network. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 15425–15434.

[Google Scholar](#) [↗](#)

[Ma et al., 2017](#) Ma L., Li M., Ma X., Cheng L., Du P., Liu Y.

A review of supervised object-based land-cover image classification

ISPRS J. Photogramm. Remote Sens., 130 (2017), pp. 277-293

 [View PDF](#) [View article](#) [View in Scopus](#) [↗](#) [Google Scholar](#) [↗](#)

[Ma et al., 2021b](#) Ma C., Sha D., Mu X.

Unsupervised adversarial domain adaptation with error-correcting boundaries and feature adaption metric for remote-sensing scene classification

Remote Sens., 13 (1270) (2021)

[Google Scholar](#) [↗](#)

[Ma et al., 2021a](#) Ma A., Wan Y., Zhong Y., Wang J., Zhang L.

Scenenet: Remote sensing scene classification deep learning network using multi-objective neural evolution architecture search

ISPRS J. Photogramm. Remote Sens., 172 (2021), pp. 171-188

 [View PDF](#) [View article](#) [View in Scopus](#) [↗](#) [Google Scholar](#) [↗](#)

[Miao et al., 2022](#) Miao W., Geng J., Jiang W.

Semi-supervised remote sensing image scene classification using representation consistency siamese network

IEEE Trans. Geosci. Remote Sens. (2022)

[Google Scholar](#) [↗](#)

[Motiian et al., 2017](#) Motiian,S., Piccirilli,M., Adjeroh,D.A., Doretto,G., 2017. Unified deep supervised domain adaptation and generalization. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 5715–5725.

[Google Scholar](#) [↗](#)

[Pan et al., 2010](#) Pan S.J., Tsang I.W., Kwok J.T., Yang Q.

Domain adaptation via transfer component analysis

IEEE Trans. Neural Netw., 22 (2010), pp. 199-210

[CrossRef](#) [↗](#) [Google Scholar](#) [↗](#)

[Paszke et al., 2017](#) Paszke A., Gross S., Chintala S., Chanan G., Yang E., DeVito Z., Lin Z., Desmaison A., Antiga L., Lerer A.

Automatic differentiation in pytorch

(2017)

[Google Scholar](#) [↗](#)

[Qiu et al., 2019](#) Qiu C., Mou L., Schmitt M., Zhu X.X.

Local climate zone-based urban land cover classification from multi-seasonal sentinel-2 images with a recurrent residual network

ISPRS J. Photogramm. Remote Sens., 154 (2019), pp. 151-162

 [View PDF](#) [View article](#) [View in Scopus](#) [↗](#) [Google Scholar](#) [↗](#)

[Ranjan et al., 2017](#) Ranjan R., Castillo C.D., Chellappa R.

L2-constrained softmax loss for discriminative face verification

(2017)

arXiv

[Google Scholar](#) [↗](#)

[Rottensteiner et al., 2012](#) Rottensteiner F., Sohn G., Jung J., Gerke M., Baillard C., Benitez S., Breitkopf U.

The isprs benchmark on urban object classification and 3d building reconstruction

ISPRS Ann. Photogram. Remote Sens. Spat. Inf. Sci., I-3 (2012) (2012), pp. 293-298

Nr. 1 1

[View in Scopus](#) [Google Scholar](#)

[Saito et al., 2019](#) Saito, K., Kim, D., Sclaroff, S., Darrell, T., Saenko, K., 2019. Semi-supervised domain adaptation via minimax entropy. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 8050–8058.

[Google Scholar](#)

[Simonyan and Zisserman, 2014](#) Simonyan K., Zisserman A.

Very deep convolutional networks for large-scale image recognition

(2014)

arXiv preprint [arXiv:1409.1556](#)

[Google Scholar](#)

[Sohn et al., 2020](#) Sohn K., Berthelot D., Carlini N., Zhang Z., Zhang H., Raffel C.A., Cubuk E.D., Kurakin A., Li C.L.

Fixmatch: Simplifying semi-supervised learning with consistency and confidence

Adv. Neural Inf. Process. Syst., 33 (2020), pp. 596–608

[Google Scholar](#)

[Song et al., 2019](#) Song S., Yu H., Miao Z., Zhang Q., Lin Y., Wang S.

Domain adaptation for convolutional neural networks-based remote sensing scene classification

IEEE Geosci. Remote Sens. Lett., 16 (2019), pp. 1324–1328

[CrossRef](#) [View in Scopus](#) [Google Scholar](#)

[Tan and Le, 2019](#) Tan M., Le Q.

Efficientnet: Rethinking model scaling for convolutional neural networks

International Conference on Machine Learning, PMLR (2019), pp. 6105–6114

[Google Scholar](#)

[Tzeng et al., 2017](#) Tzeng, E., Hoffman, J., Saenko, K., Darrell, T., 2017. Adversarial discriminative domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 7167–7176.

[Google Scholar](#)

[Van der Maaten and Hinton, 2008](#) Van der Maaten L., Hinton G.

Visualizing data using t-sne

J. Mach. Learn. Res. (2008), p. 9

[Google Scholar](#)

[Wang et al., 2018](#) Wang R., Collins L.M., Bradbury K., Malof J.M.

Semisupervised adversarial discriminative domain adaptation, with application to remote sensing data

IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE (2018), pp. 3611–3614

[CrossRef](#) [View in Scopus](#) [Google Scholar](#)

[Wang et al., 2020](#) Wang Q., Huang W., Xiong Z., Li X.

Looking closer at the scene: Multiscale representation learning for remote sensing image scene classification

IEEE Trans. Neural Netw. Learn. Syst. (2020)

[Google Scholar](#)

[Wurm et al., 2019](#) Wurm M., Stark T., Zhu X.X., Weigand M., Taubenböck H.

Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks

ISPRS J. Photogramm. Remote Sens., 150 (2019), pp. 59–69

 [View PDF](#) [View article](#) [View in Scopus](#) [Google Scholar](#)

[Xiong et al., 2021](#) Xiong Y., Xu K., Dou Y., Zhao Y., Gao Z.

Wrmatch: Improving fixmatch with weighted nuclear-norm regularization for few-shot remote sensing scene classification

IEEE Trans. Geosci. Remote Sens., 60 (2021), pp. 1–14

[CrossRef](#) [Google Scholar](#)

[Xiong et al., 2022](#) Xiong Z., Zhang F., Wang Y., Shi Y., Zhu X.X.

Earthnets: Empowering ai in earth observation

(2022)

arXiv preprint [arXiv:2210.04936](#)

[Google Scholar](#)

[Yu et al., 2022](#) Yu T., Lin J., Mou L., Hua Y., Zhu X., Wang Z.J.

Scida: Self-correction integrated domain adaptation from single-to multi-label aerial images

IEEE Trans. Geosci. Remote Sens. (2022)

[Google Scholar](#) [🔗](#)

[Yuan et al., 2019](#) Yuan, Y., Wang, D., Wang, Q., 2019. Memory-augmented temporal dynamic learning for action recognition. In: Proceedings of the AAAI Conference on Artificial Intelligence. pp. 9167–9175.

[Google Scholar](#) [🔗](#)

[Zhang et al., 2020](#) Zhang J., Liu J., Pan B., Shi Z.

Domain adaptation based on correlation subspace dynamic distribution alignment for remote sensing image scene classification

IEEE Trans. Geosci. Remote Sens., 58 (2020), pp. 7920-7930

[CrossRef](#) [🔗](#) [View in Scopus](#) [🔗](#) [Google Scholar](#) [🔗](#)

[Zheng et al., 2021](#) Zheng X., Gong T., Li X., Lu X.

Generalized scene classification from small-scale datasets with multitask learning

IEEEトランス。地理学。リモートセンズ(2021)

[Google スカラー](#) [🔗](#)

[Zheng et al., 2022a](#) Zheng J., Wu W., Yuan S., Zhao Y., Li W., Zhang L., Dong R., Fu H.

単一ソース混合複数ターゲットドメイン適応におけるリモートセンシングシーン分類のための2段階適応ネットワーク (tsan)²メートル²t da) シナリオ

IEEEトランス。地理学。遠隔監視員(2022)

[Google スカラー](#) [🔗](#)

[鄭ら, 2022b](#) Zheng Z., Zhong Y., Su Y.

リモートセンシングのクロスシーン分類のためのタスク固有の分類子フレームワークによるドメイン適応

IEEEトランス。地理学。遠隔監視員(2022)

[Google スカラー](#) [🔗](#)

[朱他, 2021](#) Zhu S., Du B., Zhang L., Li X.

クロスシーン分類のためのアテンションベースのマルチスケール残差適応ネットワーク

IEEEトランス。地理学。Remote Sens., 60 (2021), pp. 1 - 15

[クロスリファレンス](#) [🔗](#) [Google スカラー](#) [🔗](#)

[ゾウら, 2015](#) Zou Q., Ni L., Zhang T., Wang Q.

リモートセンシングシーン分類のためのディープラーニングベースの特徴選択

IEEE 地球科学。遠隔監視員Lett., 12 (2015), 2321 - 2325ページ

[Scopus で表示](#) [🔗](#) [Google スカラー](#) [🔗](#)

引用者 (12)

深層転移学習によるリモートセンシング画像の都市間土地利用分類

2023 年、応用地球観測および地理情報国際ジャーナル

[要約を表示](#) [📄](#)

メートル解像度の衛星画像を使用して国規模の土地被覆マッピングを可能にする

2023 年、ISPRS 写真測量とリモートセンシングジャーナル

[要約を表示](#) [📄](#)

GeoMultiTaskNet: 地理座標を使用したリモートセンシングの教師なしドメイン適応 [🔗](#)

2023年、arXiv

リモートセンシング画像シーン分類のためのユニバーサルドメイン適応 [🔗](#)

2023年、arXiv

GeoMultiTaskNet: 地理座標を使用したリモート センシングの教師なしドメイン適応

2023 年、コンピューター ビジョンおよびパターン認識ワークショップに関する IEEE コンピューター協会会議

>

Scopus のすべての引用記事を表示する

- 1
- <https://pytorch.org/>
- 2
- https://github.com/VisionLearningGroup/SSDA_MME
- 3
- <https://github.com/corenel/pytorch-adda>
- 4
- <https://github.com/fungtion/DANN>
- 5
- <https://github.com/lijichang/CVPR2021-SSDA>

© 2022 著者。International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) に代わって Elsevier BV が発行。