

画像加工による異常検知モデル学習用データセットの圧縮手法の提案

Proposal of a Compression Method for Anomaly Detection Model Training Datasets by Image Processing

増田 嶺¹⁾ 最所 圭三¹⁾ 亀井 仁志¹⁾ 喜田 弘司¹⁾
 Rei Masuda Keizo Saisho Hitoshi Kamei Koji Kida

1 はじめに

機械学習は、外観検査における異常検知に活用されている。そのモデルの学習はクラウドサービス上で行われることが多く、データセットサイズは巨大なため、ストレージコストが高くなる問題が挙げられる。この解決には、機械学習における画像データが類似性のある画像集合であるため、データを圧縮する方法が有効である。以前より、複数の画像をまとめて圧縮符号化する手法は研究されており、近年では画像間距離に基づいて並び替えを行い、HEVC を用いて符号化する手法が提案されている [1]。本研究では、さらに画像加工を組合せることによって、異常検知精度の維持と高い圧縮率の両立を実現する手法を提案する。本稿では、提案した手法及びその圧縮効果、異常検知モデルの精度変化について述べる。

2 提案手法

2.1 概要

本研究では、画像データセットを HEVC によって符号化する前に、画像加工する。画像データセットに対して画像内オブジェクトの位置合わせや背景削除を行い、データ間の類似性を高め、それらを 1 つの動画ファイルに圧縮する。これにより、学習時の精度を落とすことなく圧縮効率を向上させている。また、テスト用データに対しても学習用データセットと同様の加工を施すことで、学習済モデルの精度維持を実現する。なお本稿では、剛体かつ画像間の類似性が高いと考えられる、工業製品の画像データセットを対象とした、提案手法の適用とその評価について述べる。

2.2 基準画像への位置合わせ

データセットにおける各画像 (以下、対象画像) 内に存在する単独のオブジェクトを、基準となる画像 (以下、基準画像) 内のオブジェクトに重なるよう、位置を移動させる処理を行う。なお、オブジェクトの位置をデータセットの画像のいずれかに揃えればよいため、どれを基準画像として選んでも構わない。

まず、特徴点抽出アルゴリズム AKAZE を用い、対象画像内オブジェクトの特徴量を抽出する。次に、基準画像と対象画像の特徴量を比較し、基準画像に対する対象画像の回転、拡大縮小、並行移動量をアフィン行列として推定する。最後に、推定したアフィン行列を用いて対象画像を移動させる。これを学習用データセット内の全ての画像に対して繰り返す。画像の特徴量を抽出する際、Python ライブラリ Rembg を用いて一時的な背景削除を行い、余分な情報を排除する。

2.3 背景削除

2.2 節で得た一時的背景削除済の画像群から、単一のマスク画像を作成し、切り抜きを行う。Rembg による

背景削除は、必要な領域も切り取ってしまう可能性があるため、そのまま利用しただけでは学習時の精度に影響を及ぼす恐れがある。そこで、各画像の切り抜き領域を示すマスク画像の論理和を求め、すべてのオブジェクトが収まる最小領域を持ったマスク画像を作成する (図 1)。これを用いて、全ての画像に対しマスク処理を行うことで、オブジェクトに干渉しない範囲の背景を大まかに削除できる (図 2)。

2.4 動画ファイルへの変換

加工した画像データセットを、動画ファイルへと変換する。変換は、HEVC の非可逆圧縮モードを用い、画像 1 枚を動画データの 1 フレームとした上で圧縮する。HEVC による変換では、画像間の冗長性を排除して符号化されるため、画像データセットサイズの大幅な縮小が期待される。

2.5 学習とテストに向けた前処理

モデルの学習時には、一度動画ファイルへと変換したデータセットを、再び個々の画像ファイルに復元する。これにより、従来と同様の方法で学習処理を行える。

また、テスト用データも学習用データセットと同様、位置合わせや背景削除を行い、異常検知精度の維持を目指す。これは、2.2 節及び 2.3 節の工程により、学習済モデルは、背景や位置による精度の偏りが生まれると予想されるためである。位置合わせの基準画像や、背景削除に使用するマスク画像は、学習用データセットの処理時に得たものを用いる。

3 評価

3.1 比較内容

提案手法が、モデルの精度を落とすことなく、データサイズを縮小することを確認する。また、位置合わせと背景削除それぞれの効果も確認する。

表 1 に、本稿で比較する各手法を列挙する。以後、各手法を表 1 で付与した番号を用いて呼称する。①のサイズは、オリジナルの PNG 画像 320 枚の合計サイズ、それ以外の手法のサイズは、変換した動画ファイルのサイズである。①は、①の各 AUC 値に近似するように、HEVC における圧縮パラメタを調整したものである。このパラメタは、③、⑤、⑦にも用いている。②、④、⑥、⑧は、各手法に変更は加えず圧縮パラメタのみを調

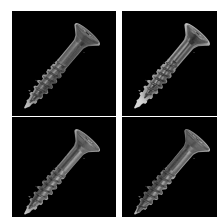
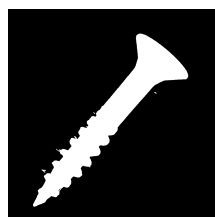


図 1 マスク画像の論理和

図 2 背景削除した画像群

1) 香川大学 Kagawa University

表 1 処理手法毎のファイルサイズと圧縮率及び学習済モデルの ROCAUC

番号	処理手法の説明	サイズ [KB]	圧縮率 [%]	image level ROCAUC	pixel level ROCAUC
①	事前処理無し	124,000	3956.6	0.844	0.983
②	動画圧縮のみ	3,134	100.0	0.846	0.982
③	①の動画化パラメタ変更	371	11.8	0.677	0.957
④	位置合わせと動画圧縮の組合せ	2,956	94.3	0.944	0.991
⑤	③の動画化パラメタ変更	285	9.1	0.626	0.949
⑥	背景削除と動画圧縮の組合せ	3,448	110.0	0.646	0.965
⑦	⑤の動画化パラメタ変更	482	15.4	0.618	0.950
⑧	位置合わせと背景削除と動画圧縮の組合せ	2,551	81.9	0.979	0.993
⑨	⑦の動画化パラメタ変更	202	6.5	0.930	0.985

整し、画質を引き下げの代わりに、使用した動画作成ツールで設定可能な最大圧縮率まで高めたものである。

本稿では、画像データセットとして、MVTec AD[2][3]の screw クラスを用いた。位置合わせをする際の基準画像は、データセットの先頭に位置するものを選択する。また、異常検知モデルの学習には PaDiM を使用する。

実験は、Windows11 上で行う。画像加工には OpenCV を、画像データセットの HEVC による動画化には FFmpeg をそれぞれ用いる。

3.2 圧縮率の比較

比較対象の各手法で生成されるデータセットファイルのサイズ及び、①の手法を基準とする圧縮率を表 1 の“圧縮率”に示す。表 1 の“圧縮率”より、同一の圧縮パラメタにおいては、位置合わせと背景削除を組合せた(⑦, ⑧)場合、最も高い圧縮率を得られたことがわかる。位置合わせのみ(③, ④)では、圧縮率はわずかな向上に留まることがわかる。また、背景削除のみでは、圧縮率の向上に寄与しないこともわかる。

3.3 学習精度の比較

各手法のデータセットで学習したモデルの精度を検証し、従来の手法(①)と同等以上の AUC 値であることを確認する。

動画化したデータセットを PNG 形式の画像データセットに復元し、オリジナルのデータセット(①)と同じ条件で学習を行う。なお、PaDiM は正常品の画像のみから画像の特徴量の分布を使って判定するため、本評価では正常品のみ画像データセットを圧縮する。

各手法で生成したデータセットで学習したモデルを用い、テスト用データで検証した場合の AUC 値を表 1 の“image level ROCAUC”と“pixel level ROCAUC”列に示す。前者は画像単位での正誤、後者はピクセル単位での正誤を値として計算する。表 1 より、①, ③, ⑦, ⑧は①と同等以上の精度であることがわかる。HEVC の圧縮パラメタ変更により限界まで圧縮した手法では、②, ④, ⑥, ⑧のうち⑧のみが①と同等以上の精度を維持できていることがわかる。

4 考察

3.3 節及び 3.2 節より、提案手法(⑧)では、モデルの精度を落とすことなく、高い圧縮率を実現できていた。一方で、②, ④, ⑥は圧縮パラメタの変更により、ファイルサイズはかなり小さくなっているものの、表 1 の AUC 値より、①を大幅に下回る精度となった。また、

従来研究に近い動画化のみの手法(①, ②)では、②のように品質を下げると精度低下を招くことから、提案手法はより高い圧縮率を実現しているといえる。

位置合わせのみで圧縮率が向上しない理由として、背景情報が残存することによって、画像全体でみた時の冗長性が変化しないことが考えられる。オブジェクトの位置合わせを行ったとしても、背景の情報は画像毎に異なるため、画像間の差分が大きくなった可能性がある。背景削除のみで圧縮率が向上しない要因としては、オブジェクトの位置が様々であることが挙げられる。オブジェクトの向きや位置、拡大率が画像間で大きく異なるため、HEVC による符号化処理に影響したものと考えられる。位置合わせと背景削除の両工程を採用することが、高い圧縮率の実現及び学習済モデルの精度維持の両立には有効であるといえる。

また、位置合わせを行った場合、AUC 値が向上することが表 1 より確認できる。これは、同一の位置及び向きの学習データが増加することにより、学習済モデルが強化されたことに起因するものと考えられる。この AUC 値向上分を、圧縮パラメタ変更による圧縮率向上に転嫁することで、提案手法ではより高い圧縮率を実現できたといえる。

5 おわりに

画像加工と HEVC を用いた動画化による学習用データセットの新しい圧縮手法を提案した。

今後の課題として、画像加工によるさらなる圧縮率の向上及び、適応範囲の拡大が挙げられる。適応範囲を拡大するために、複数クラスを含むデータセットの一括圧縮手法を検討する。

謝辞

本研究にあたり、株式会社 STNet のご協力ならびに種々のご助言を賜りましたことを感謝いたします。

参考文献

- [1] Lina Sha et al., “Novel Image Set Compression Algorithm Using Rate-Distortion Optimized Multiple Reference Image Selection”, IEEE Access, Vol.6, pp.6690–66913, (2018).
- [2] Paul Bergmann et al., “The MVTec Anomaly Detection Dataset: A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection”, International Journal of Computer Vision, Vol.129, pp.1038–1059, (2021).
- [3] Paul Bergmann et al., “MVTec AD — A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection”, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.9584–9592, (2019).