ICM手法のプライバシー保護における有効性の検証
Assessing the Effectiveness of ICM Method for Privacy Protection

進藤嵩紘 † 渡部泰樹 † 巽優衣 † 渡辺裕 †
Takahiro Shindo† Taiju Watanabe† Yui Tatsumi† Hiroshi Watanabe†
† 早稲田大学
†Waseda University

Abstract: Recent advances in image recognition technology have created new opportunities for applying these techniques. As image recognition models increasingly rely on image data, the demand for efficient image compression methods tailored for recognition tasks has grown. This area of research is known as Image Coding for Machines (ICM). In this paper, we explore the effectiveness of ICM methods in protecting privacy. With rising concerns about privacy, there is a need for image transmission and storage methods that prevent individual identification. By comparing facial features between the original images and those compressed using the ICM method, we demonstrate that the ICM approach makes it more difficult to identify individuals in the images.

1. INTRODUCTION

画像認識技術の発達により，画像認識モデルの使用機会が急増している．それに伴い，これらの使用目的のための画像圧縮技術の必要性が高まっており，Image Coding for Machines (ICM) という分野において研究が盛んになっている．また，プライバシー保護への関心も高まっており，個人の特定ができないような画像の復号手法が必要である．本稿では，ICM 手法の一つである SA-ICM[1] の復号画像における，顔画像の再現性について調査する．SA-ICM は，人物などの物体の輪郭情報を復号する一方で，画像認識に不要なテクスチャを破棄する性質を持つ．そこで，SA-ICM の入力画像と出力画像において，顔画像の類似性を評価し，プライバシーの保護における SA-ICM の有効性について検証する．

2. RELATED WORK

2.1. SA-ICM

画像認識モデルによる画像の使用量が増大し，ICM に関する研究が必要とされている．多くの ICM 手法は，特定の認識モデルのための画像圧縮手法として最適化されており，従来の圧縮規格を大きく上回る認識のための圧縮性能を達成する．しかし，他の認識タスクや認識モデルへの適用が困難であり，汎用性が低い．一方 SA-ICM は，object detection，instance segmentation，panoptic segmentation のための画像圧縮手法であり，様々な認識モデルのための画像圧縮を可能にする．Segment Anything[2] を用いて Learned Image Compression[3] モデルを学習させることで，画像中の物体を構成する輪郭部分を復号する．復号画像の一例を図 1 に示す．図 1 のように，物体の輪郭情報を再現する一方で，多くのテクスチャが破棄される．また，復号画像中の顔の情報も限られており，個人の特定が困難なように思える．本稿では，SA-ICM を画像圧縮手法として使用し，その入力画像と復号画像の特徴量を比較する．

Fig. 1. Example of coded images using SA-ICM.

2.2. DeepFace

DeepFace[4] は，顔画像に関する様々な認識タスクを行うための， pytorch で実装されたフレームワークである．本稿では，顔認識のためのツールとして，顔の検出と，顔画像同士の類似度を測定するために使用する．顔検出モデルとしては，mediapipe，mtcnn[5]，retinaface，yolov8 などが実装されている．類似度測定のための特徴抽出器としては，VGG-Face[7]，Facenet[8]，ArcFace[9] などが実装されている．これらのモデルは，使用用途に合わせて自由に選択し，組み合わせることができる．実験において，検出モデルには mtcnn を使用し，特徴抽出器には，複数種類のモデル使用を試みる．

Table 1. Face Recognition Accuracy in VVC and SA-ICM Compressed Images.

Method	Bitrate [bpp] (↓)	Detection accuracy [%] (↑)	Cosine similarity (↓)					
			VGG-Face	FaceNet	ArcFace	OpenFace	SFace	GhostFaceNet
SA-ICM	0.198	**41.5**	**0.27**	**0.28**	**0.28**	**0.62**	**0.15**	**0.26**
VVC (QP=32)	0.524	41.5	0.89	0.94	0.89	0.97	0.87	0.87
VVC (QP=35)	0.357	39.4	0.80	0.86	0.78	0.94	0.79	0.78
VVC (QP=37)	0.271	37.4	0.74	0.79	0.71	0.91	0.72	0.72
VVC (QP=40)	0.172	33.2	0.58	0.67	0.54	0.88	0.57	0.59

3. EXPERIMENTAL METHOD

SA-ICM による復号画像と，圧縮前の元の画像における，顔画像の特徴量を比較する．我々が用意した画像処理プロセスを図 2 に示す．まず COCO dataset[6] より，人物の顔を含む画像を 15 枚用意し，SA-ICM を用いて圧縮する．用意した符号化画像と，元画像における顔の特徴量を比較するため，DeepFace を用いて，元画像から顔の矩形検出を行う．このとき検出モデルには，mtcnn[5] を用いる．次に SA-ICM 符号化画像と元画像より，この矩形領域を取り出し，それらの類似度を計測する．類似度測定には DeepFace を使用し，特徴抽出器には ArcFace[9] など 6 種類の手法を用いる．これらの特徴量の類似度をコサイン類似度により計測する．

4. EXPERIMENTAL RESULTS

表 1 に，SA-ICM 圧縮画像におけるビットレートと，物体検出精度，元画像との顔の類似度を示す．比較手法として，VVC (Versatile Video Coding)[10] 符号化画像を使用する．表 1 より，SA-ICM は物体の輪郭情報のみを復号するため，画像認識のための画像圧縮手法として優れていることが分かる．また，元画像の顔と符号化画像の顔の類似度を判定した結果，SA-ICM による圧縮画像において，最も類似度が低いことを確認した．この結果は，プライバシー保護の観点から，SA-ICM の有効性を示している．

5. CONCLUSION

本稿では，プライバシー保護における，SA-ICM の有効性について検証した．実験において，SA-ICM の入力画像と出力画像における，顔の特徴量を比較した．これにより，SA-ICM を用いた画像の圧縮は，個人を特定するための顔認識を妨げることを確認した．今後の研究において，顔画像のサンプル数を増やし，汎用性を確認する．また，他の ICM 手法についても，プライバシーの保護における有効性を検証する必要がある．

6. ACKNOWLEDGMENT

The results of this research were obtained from the commissioned research (JPJ012368C05101) by National Institute of Information and Communications Technology (NICT).

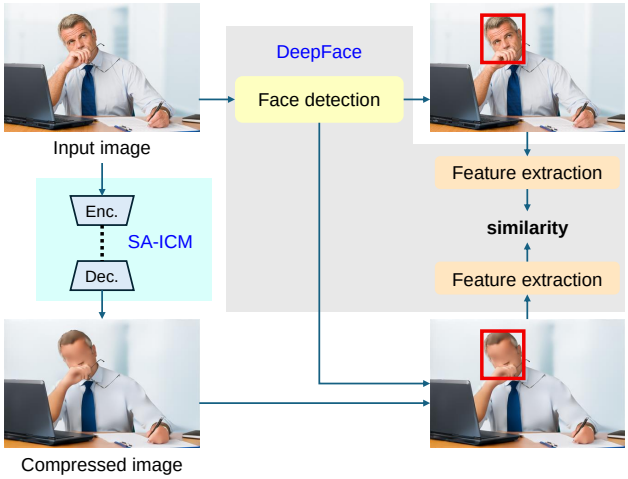

Fig. 2. Image processing flow.

7. REFERENCES

[1] T. Shindo *et al.*, "Image Coding For Machines With Edge Information Learning Using Segment Anything," 2024 IEEE International Conference on Image Processing (ICIP), 2024, pp. 3702-3708.

[2] A. Kirillov *et al.*, " Segment Anything, " Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023, pp. 4015-4026.

[3] J. Liu *et al.*, "Learned Image Compression with Mixed Transformer-CNN Architectures," 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 14388-14397.

[4] S. I. Serengil *et al.*, "LightFace: A Hybrid Deep Face Recognition Framework," 2020 Innovations in Intelligent Systems and Applications Conference (ASYU), 2020, pp. 1-5.

[5] K. Zhang *et al.*, "Joint face detection and alignment using multitask cascaded convolutional networks." IEEE signal processing letters 23.10 (2016): 1499-1503.

[6] T. Y. Lin *et al.*, "Microsoft COCO: Common Objects in Context," Computer Vision - ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8693. 2014, pp 740-755.

[7] O. M. Parkhi *et al.*, " Deep Face Recognition, " 2015, pp. 41.1-41.12, doi: 10.5244/c.29.41.

[8] F. Schroff *et al.*," FaceNet: A unified embedding for face recognition and clustering, " in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015, pp. 815-823.

[9] J. Deng *et al.*, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 4685-4694.

[10] Versatile Video Coding, Standard ISO/IEC 23090-3, ISO/IEC JTC 1, Jul. 2020.

早稲田大学大学院 基幹理工学研究科 情報理工・情報通信専攻
〒*162-0072* 東京都新宿区大久保 *3-14-9* 早大 *66-401*
Phone: 03-5286-2509, E-mail: taka_s0265@ruri.waseda.jp