IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 11, No. 1, March 2022, pp. 254~264
ISSN: 2252-8938, DOI: 10.11591/ijai.v11.i1.pp254-264

Privacy preserving human activity recognition framework using an optimized prediction algorithm

Kambala Vijaya Kumar, Jonnadula Harikiran
School of Computer Science and Engineering, VIT-AP University, Amaravathi, India

Article Info

Article history:

Received Jul 23, 2021
Revised Dec 22, 2021
Accepted Jan 3, 2022

Keywords:

Aadaptive privacy model
Adversarial learning
Deep neural networks
Human action recognition
Visual privacy

ABSTRACT

Human activity recognition, in computer vision research, is the area of growing interest as it has plethora of real-world applications. Inferring actions from one or more persons captured through a live video has its immense utility in the contemporary era. Same time, protecting privacy of humans is to be given paramount importance. Many researchers contributed towards this end leading to privacy preserving action recognition systems. However, having an optimized model that can withstand any adversary models that strives to disclose privacy information. To address this problem, we proposed an algorithm known optimized prediction algorithm for privacy preserving activity recognition (OPA-PPAR) based on deep neural networks. It anonymizes video content to have adaptive privacy model that defeats attacks from adversaries. The privacy model enhances the privacy of humans while permitting highly accurate approach towards action recognition. The algorithm is implemented to realize privacy preserving human activity recognition framework (PPHARF). The visual recognition of human actions is made using an underlying adversarial learning process where the anonymization is optimized to have an adaptive privacy model. A dataset named human metabolome database (HMDB51) is used for empirical study. Our experiments with using Python data science platform reveal that the OPA-PPAR outperforms existing methods.

Corresponding Author:

Vijaya Kumar Kambala
School of Computer Science and Engineering, Vellore Institute of Technology, VIT-AP University
Amaravathi, Vijayawada, Andhra Pradesh, India
Email: kvkumar@pvpsiddhartha.ac.in

1. INTRODUCTION

Video based surveillance has become an important computer vision application. It has plenty of applications in the real world. While video based surveillance in different domains is useful, it has potential risk in terms of privacy leakage. Therefore, many researchers contributed towards privacy preserving action recognition. Human action recognition is an important research area with rich set of methods with machine learning, deep learning and generative adversarial network (GAN) based models. Action recognition using deep learning, often supported by privacy preserving method, are explored in [1]–[6]. Lyu *et al.* [1] proposed a deep learning based method for privacy preserving framework with fair and decentralized approach. Rasim *et al.* [2] proposed a deep learning based model for privacy preserving approach to protect personal data. Weng *et al.* [3] proposed a deep learning model with blockchain for privacy protection. Lyu *et al.* [4] studied federated cloud models to achieve fair and privacy preserving approaches to solve problems. Kumar *et al.* [5] explored deep learning algorithms and resolution images besides spatial relationships to recognize human actions. Rajpur *et al*. [6] proposed a cloud-based service to achieve privacy preserving action recognition using deep convolution neural network (CNN) model.

There are many adversarial models that paved way for human action recognition. They are found in [7]–[12] to mention few. Wu *et al*. [7] proposed a privacy-protective-generative adversarial network (PP-GAN) with modules such as regulator and verificator. It ensures protection of privacy, structure similarity and utility of the approach. Debie *et al*. [8] proposed a privacy preserving GAN for classification of ECG data. Maximov *et al*. [9] proposed a GAN based system known as conditional identity anonymization generative adversarial network (CIAGAN) which supports anonymization and recognition of actions in image and video. In future, they intend to enhance it with full image anonymization. Martinsson *et al*. [10] proposed an adversarial representation learning model with efficient management of learnable parameters. Li *et al*. [11] used a pre-trained GAN based model for privacy protection. Shirai and Whitehill [12] proposed a GAN based model for recognition of faces.

From the literature, it is understood that there are plenty of deep learning based methods for action recognition. Similarly, there are many GAN based approaches used for human activity recognition. Many of the deep learning and GAN based methods are equipped with privacy preserving approaches to protect data. However, there is need for optimization of action recognition method with privacy budget optimization. To address this problem, we proposed an algorithm known as optimized prediction algorithm for privacy preserving activity recognition (OPA-PPAR) based on deep neural networks. It anonymizes video content to have adaptive privacy model that defeats attacks from adversaries. The privacy model enhances the privacy of humans while permitting highly accurate approach towards action recognition. The algorithm is implemented to realize privacy preserving human activity recognition framework (PPHARF). The visual recognition of human actions is made using an underlying adversarial learning process where the anonymization is optimized to have an adaptive privacy model. A dataset named HMDB51is used for empirical study. Our contributions in this paper are: i) we proposed a framework known as PPHARF that leverages action recognition model, privacy budget model and anonymization model for privacy preserving with adversarial setting; ii) we proposed an algorithm known as OPA-PPAR based on deep neural networks; and iii) we built an application to evaluate the PPHARF and the underlying OPA-PPAR algorithm using HMDB51 dataset.

The remainder of the paper is structured in: section 2 review different kinds of methods used for action recognition and privacy preservation. Section 3 presents the proposed method with underlying algorithm. Section 4 presents experimental results and evaluates the same. Section 5 concludes the paper and gives suggestions for future work.

2. RELATED WORK

Human action recognition is an important research area with rich set of methods with machine learning, deep learning and generative adversarial network (GAN) based models. Many privacy preserving deep learning techniques are explored by Boulemtafes *et al*. [13]. Malekzadeh *et al*. [14] proposed privacy preserving based approach that makes use of deep autoencoder. Lyu *et al*. [1] proposed a deep learning-based method for privacy preserving framework with fair and decentralized approach. Rasim *et al*. [2] proposed a deep learning-based model for privacy preserving approach to protect personal data. Weng *et al*. [3] proposed a deep learning model with blockchain for privacy protection. Yonetani *et al*. [15] investigated on security using doubly permuted homomorphic encryption (DPHE) which is meant for protecting high-dimensional data. Lyu *et al*. [4] studied federated cloud models to achieve fair and privacy preserving approaches to solve problems. Yang *et al*. [16] employed machine learning (ML) models for hyperspectral image classification. Du *et al*. [17] proposed deep learning models with privacy preserving and also approximate approach in computing. Jhonson *et al*. [18] focused on the real time style transfer using perception loss and super-resolution. He *et al*. [19] proposed a method for image recognition based on deep residual learning. Kuehne *et al*. [20] worked on the video database known as HMDB that is used for human action recognition.

Yun *et al*. [21] focused on human activity recognition using multiple instance learning and body pose features. He *et al*. [22] exploited deep residual networks with identity mapping. Szegedy *et al*. [23] investigated on deep convolutional networks with action recognition using pre-recorded videos. Leenes *et al*. [24] studied on the privacy issues associated with data protection Dai *et al*. [25] proposed a novel method towards human action recognition with privacy preserved. Kumar *et al*. [5] explored deep learning algorithms and resolution images besides spatial relationships to recognize human actions. Orekondy *et al*. [26] proposed a model for visual privacy advisor that improves privacy of the system. Pittaluga *et al*. [27] focused on motion reconstruction of videos by using different image descriptors. Dai *et al*. [28] used spatial resolution cameras and extremely low temporal resolutions for activity recognition and preserving privacy. Dosovitskiy and Brox [29] investigated on convolutional networks for inverting of visual representations. Lyu *et al*. [30] proposed collaborative deep learning models for human activity recognition. Weinzaepfel *et al*. [31] exploited local descriptors in images to arrive at reconstruction of images for visual quality. Ryoo *et al*. [32] used superstitious video recordings in order to recognize human actions from extreme low-resolution videos. Mahendran and Vedaldi [33] explored on the visualization of CNNs by using natural pre-images. Wang *et al*. [34] used coded aperture videos for human activity recognition with privacy preserved.

Machot *et al*. [35] investigated on sensor data in order to discover unseen activities associated with human action recognition. Pittaluga and Koppal [36] used miniature vision sensors proposed privacy preserving optics to strike balance between utility of videos and privacy. It has many applications like motion tracking, depth sensing and blob detection. Pittaluga *et al*. [37] did similar kind of work. Zhang *et al*. [38] proposed a methodology to identify human activities associated with fall detection of elderly people. Sur *et al*. [39] on the other hand proposed a technique to characterize given target using MIMO radar. Rajpur *et al*. [40] proposed a cloud-based service to achieve privacy preserving action recognition using deep CNN model. Cheng *et al*. [41] used a deep learning approach for emotion recognition. Riboni and Bettini [42] provided an ontology-based approach towards context aware activity recognition supported by hybrid reasoning. Xu *et al*. [43] defined an architecture for human activity recognition with two-stream spatiotemporal networks fully coupled.

Zolfaghari *et al*. [44] proposed smart activity recognition framework (SARF) that helps in monitoring humans that promote ambient assisted living (AAL). Youn *et al*. [45] focused on prognostics and health management that involves sensing functions, reasoning, prognostics, and health management. Ciliberto *et al*. [46] proposed a 3D model to have action recognition with privacy preserved. Cippitelli *et al*. [47] used skeletal data collected from sensors to detect human actions. Wang *et al*. [48] studied on gender bias elimination while making deep image representations.

Wu *et al*. [7] proposed a privacy-protective-GAN (PP-GAN) with modules such as regulator and verificator. It ensures protection of privacy, structure similarity and utility of the approach. It has issues with different head poses of humans in terms of face recognition that needs further improvement. Debie *et al*. [8] proposed a privacy preserving GAN for classification of ECG data. Maximov *et al*. [9] proposed a GAN based system known as CIAGAN which supports anonymization and recognition of actions in image and video. In future, they intend to enhance it with full image anonymization. Martinsson *et al*. [10] proposed an adversarial representation learning model with efficient management of learnable parameters. Tseng and Wu [49] proposed GAN known as "privacy generative adversarial network (CPGAN)" which is a learning framework with adversarial settings. Jin *et al*. [50] proposed Asynchronous Interactive GAN while Li *et al*. [11] used a pre-trained GAN based model for privacy protection. Ma *et al*. [51] defined yet another GAN model known as fusion GAN which makes use of a game between generator and discriminator. Shirai and Whitehill [12] proposed a GAN based model for recognition of faces. Liu *et al*. [52] explored adversarial networks for accuracy enhancement and privacy quantification.

Wu *et al*. [53] proposed GAN model for visual recognition while preserving privacy. They used the concept of restarting and ensemble approaches to leverage performance. Roy and Boddeti [54] proposed a non-zero sum game with adversarial settings. Zhang *et al*. [55] used adversarial learning mechanism to reduce unwanted biases in ML applications. Cheid *et al*. [56] proposed a protocol named multi-party classification that helps in human action recognition with privacy preserved. Cheid and Challal [57] investigated on human activity based on sensor based on sensors and privacy preserving protocols. Oh *et al*. [58] proposed a faceless person recognition and investigated on its implications. From the literature, it is understood that there are plenty of deep learning-based methods for action recognition. Similarly, there are many GAN based approaches used for human activity recognition. Many of the deep learning and GAN based methods are equipped with privacy preserving approaches to protect data. However, there is need for optimization of action recognition method with privacy budget optimization. Towards this end a framework is proposed in this paper.

3. MATERIALS AND METHOD

3.1. Problem definition

Given a video dataset (raw videos captured), denoted as *X*, which is subjected to action recognition task *T* with a privacy budget. The dataset *X* has set of class labels denoted by y_T and the performance of task is evaluated using a cost function denoted as L_T. An existing supervised learning method for prediction of actions is denoted as f_T which is enhanced to support J_B which is a cost function for budget associated with privacy leakage and used to find privacy leakage. Smaller value of J_B indicates that the input data has less private information associated with it. Table 1 shows the notations used in the paper.

Provided *X*, define an anonymization function f_A^* which transforms *X* into anonymized *X* denoted as $f_A^*(X)$ and a new deep learning based action recognition model, denoted as f_T^* is derived. In the process, care is taken to ensure that the function of $f_T$is affected minimally. This dual goal is to be achieved is considered as an optimization problem expressed in (1). The cost function of privacy budget is dynamic in nature as it depends on the runtime task. Therefore, (1) is redefined and expressed as in (2). A fixed structure neural network, denoted as f_B, is defined in order to have finite search space to solve the problem with ease. This modification is expressed in (3). In order to enhance performance of the deep learning model, we proposed an ensemble approach.

Table 1. Notations used in the paper

Notation	Description
f_A^*	Anonymization function optimized
f_T^*	The new or derived deep learning method
f_T	An existing deep learning method for prediction
L_T	Cost function
f_A	Anonymization function
X	Raw video dataset
y_T	Set of labels of X
J_B	Cost function for privacy budget to find privacy leakage
$f_A^*(X)$	Anonymized input data X
f_B	A privacy budget model. It is a fixed structure neural network
θ_A	Represents learnable parameters of f_A
θ_B	Represents learnable parameters of f_B
θ_T	Represents learnable parameters of f_T
H_B	Negative entropy

$$f_A^*, f_T^* = argmin\,\left[L_T\left(f_T\left(f_A(X)\right), y_T\right) + \gamma_B^J\left(f_A(X)\right)\right] \tag{1}$$

$$f_A^*, f_T^* = argmin_{(f_A, f_T)}\left[L_T\left(f_T\left(f_A(X)\right), y_T\right) + \gamma sup\ _{fB\epsilon p} J_B\left(f_B\left(f_A(X)\right), y_B\right)\right] \tag{2}$$

$$f_A^*, f_T^* = argmin_{(f_A, f_T)}\left[L_T\left(f_T\left(f_A(X)\right), y_T\right) + \gamma max\ _{fB} J_B\left(f_B\left(f_A(X)\right), y_B\right)\right] \tag{3}$$

3.2. Proposed framework

We proposed a framework named PPHARF which is crucial for accommodating the underlying mechanisms and algorithms to achieve the desired dual goal of the system which enhances the action recognition with privacy leakage prevention and keeps the capabilities of prediction algorithm maximal. The framework has different models involved. They are known as the action recognition model f_T^*, an optimised anonymization function f_A^* and a privacy budget model denoted as f_B. These models are implemented as deep neural networks with learnable parameters. The training of the entire model is made with combination of two loss functions namely L_T and J_B. The underlying training in the framework has a dual goal consisting of achieving optimized anonymization function f_A^* which filters private information prior to the actual task and also ensures that $f_A^*(X)$ is achieved without limiting functionality of action recognition model.

As presented in Figure 1, the learned anonymization module takes X as input and transforms it into anonymized video content that filters out private information and modifies it so as to ensure that the video content is useful for action recognition, but unique human identity cannot be achieved. The anonymized video is subjected to action recognition model which is denoted as f_T. It has its cost function denoted as L_T. In the same fashion, the anonymized video content is subjected to privacy budget module $f_B$where another cost function denoted as J_B. When both cost functions are combined to form of a loss function which controls the iterative process of the framework and ensures optimization of action recognition while preserving privacy. The anonymization model f_A is implemented as a frame level filter which is based on 2D convolutional network. The action recognition module is taken from [59] and reused it. The privacy budget model $f_B$is made up of ResNet. For the same of action recognition, the video is divided into number of frames (video clips) and each frame is uniquely identified.

A minimax problem associated with (3) is solved by considering different learnable parameters of the three models used in the framework. The learnable parameters θ_A, θ_B, and θ_T are associated with f_A, f_B and f_T respectively. In order to solve the minimax problem, we considered the notion of alternative minimization found in [60]. It is expressed as in (4)-(6). Then the two loss functions are optimized to solve the optimization problems expressed in (7) and (8). The (7) is the minimization problem while (8) is minimax problem. The former is used to have training of f_A and f_T while the latter is used to keep track of different parameters of privacy budget model. In order to solve two loss functions two parameter update rules are expressed in (9)-(12).

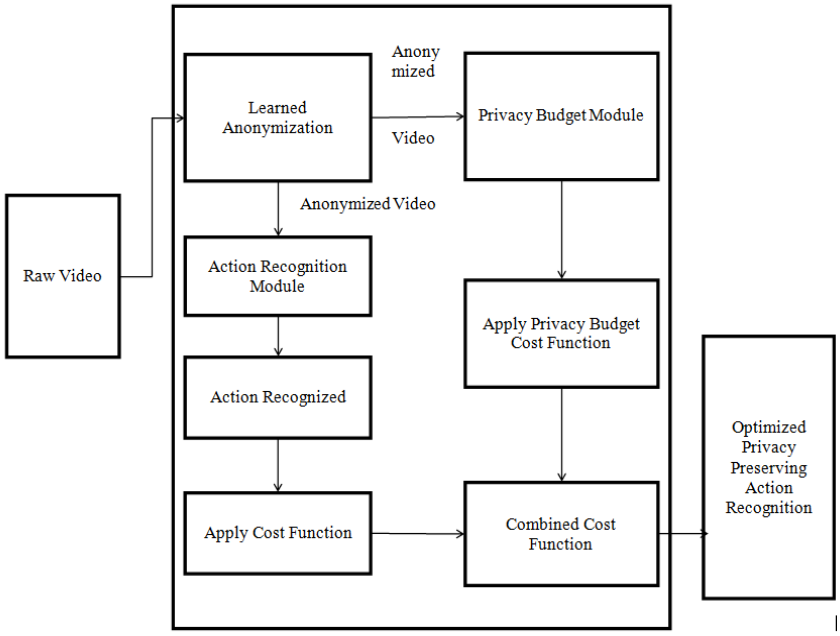

Figure 1. Proposed privacy preserving human activity recognition framework with adversarial setting

$$\theta_A \leftarrow \theta_A - \propto_A \nabla_{\theta_A}\big(L_T(\theta_A,\theta_T) - \gamma L_B(\theta_A,\theta_B)\big) \tag{4}$$

$$\theta_T \leftarrow \theta_T - \propto_T \nabla_{\theta_A} L_T(\theta_A,\theta_T) \tag{5}$$

$$\theta_B \leftarrow \theta_B - \propto_B \nabla_{\theta_B} L_B(\theta_A,\theta_B) \tag{6}$$

$$Q_A^*, Q_T^* = argmin_{(Q_A,Q_T)} L_T(\theta_A,\theta_T) \tag{7}$$

$$Q_B^*, Q_A^* = argmin_{\theta_B} argmax_{\ \theta_A} L_B(\theta_A,\theta_B) \tag{8}$$

$$\theta_A \leftarrow \theta_T \leftarrow \theta_A, \theta_T - \propto_T \nabla_{(Q_T,Q_A)} L_T(\theta_A,\theta_T) \tag{9}$$

$$j \leftarrow argmin_{i\in\{1,\ldots K\}} L_B\big(\theta_A,\theta_B^i\big) \tag{10}$$

$$\theta_A, \theta_A + \propto_A \nabla_{\theta_A} L_B\big(\theta_A,\theta_B^j\big) \tag{11}$$

$$\theta_B^i, \theta_B^i - \propto_B \nabla_{\theta_B^i} L_B\big(\theta_A,\theta_B^j\big), \forall i \in \{1,\ldots K\} \tag{12}$$

We found that (4) is instable which can be solved by considering negative entropy which is incorporated to have a new scheme as expressed in (13)-(15). With these optimizations, there is possibility of maximizing entropy that leverages performance. The (2) is further optimized with ensemble approach in the training process to improve model accuracy as expressed in (16). The ensemble model and optimized parameter settings are further improved with a scheme expressed in (17)-(19).

$$\theta_A \leftarrow \theta_A - \propto_A \nabla_{\theta_A}\big(L_T(\theta_A,\theta_T) - \gamma H_B(\theta_A,\theta_B)\big) \tag{13}$$

$$\theta_T, \theta_A \leftarrow \theta_T, \theta_A - \propto_T \nabla_{\theta_T,\theta_A} L_T(\theta_A,\theta_T) \tag{14}$$

$$\theta_B \leftarrow \theta_B - \propto_B \nabla_{\theta_B} L_B(\theta_A,\theta_B) \tag{15}$$

$$f_A^*, f_T^* = argmin_{(f_A,f_T)} \left[L_T\big(f_T(f_A(x)), y_T\big) + \gamma max_{f_B^i \in Pt} J_B\big(f_B^i(f_A(x)), y_B\big]\right. \tag{16}$$

$$\theta_A \leftarrow \theta_A - \propto_A \nabla_{\theta_A}\left(L_T + \gamma max_{\theta_B^i \in Pt} - H_B\left(\theta_A, \theta_B^j\right)\right) \tag{17}$$

$$\theta_A \leftarrow \theta_T \leftarrow \theta_A \leftarrow \theta_T - \propto_T \nabla_{(Q_T, Q_A)} L_T(\theta_A, \theta_T) \tag{18}$$

$$\theta_B^i, \theta_B^i - \propto_B \nabla_{\theta_B^i} L_B\left(\theta_A, \theta_B^j\right), \forall i \in \{1, \dots M\} \tag{19}$$

With these optimizations, the proposed framework PPHARF is made more sophisticated in terms of human action recognition and preserving privacy that ensures non-disclosure of identity. With different modules in place, the framework operates in an iterative model in order to have better performance. With combined loss function it can realize the dual goal of the framework aforementioned.

3.3. The proposed algorithm

We proposed an algorithm known OPA-PPAR based on deep neural networks. It anonymizes video content to have adaptive privacy model that defeats attacks from adversaries. The privacy model enhances the privacy of humans while permitting highly accurate approach towards action recognition. The algorithm is implemented to realize PPHARF. The visual recognition of human actions is made using an underlying adversarial learning process where the anonymization is optimized to have an adaptive privacy model. A dataset named HMDB51 is used for empirical study.

Algorithm 1. Optimized prediction algorithm for privacy preserving activity recognition

Inputs: X, model learnable parameters such as θ_A, θ_B and θ_T
Output: Updated recognized actions map with privacy preserved
1. Initialize frames vector F
2. Initialize actions map R
3. F🡐SplitVideo(X)
4. For each frame f in F
5.   Repeat
6.   Apply learned anonymization model f_A on f
7.   Apply privacy budget model $f_B$on f
8.   Apply action recognition model f_T^*optimized by $f_A$and f_B
9. L_T🡐ComputeCostFunctionOfActionRecognition()
10. J_B🡐ComputeCostFunctionOfPrivacyLeakage()
11.   loss function L 🡐L_T+J_B
12.   Use learnable parameters θ_A, θ_B and θ_T
13.   Get feedback for three models
14.   Until Convergence
15.   Update R
16. End For
17. Return R

As presented in Algorithm 1, it takes X and model learnable parameters such as θ_A, θ_B, and θ_T as inputs and produces an updated recognized action map with privacy preserved. In step 1, it initialized frames vector named F which holds frames (nothing but split films of video). Step 2 initializes actions map that will be updated iteratively and retuned on convergence. Step 3 splits given raw video into some frames. An iterative process is expressed in steps 4 through step 16. For each frame again, there is an iterative process that applies the two modules as given in step 6, step 7, and step 8 respectively. Two kinds of cost functions are computed in step 9 and step 10 respectively. These two cost functions are combined in step 11 to arrive at a combined loss function that is used in the training of the models in order to give feedback and continue process until convergence. Step 12 uses learnable parameters and step 13 gives feedback needed in the adversarial setting of the proposed framework. Step 7 returns final results that are obtained with privacy preserved.

4. EXPERIMENTAL RESULTS

We proposed an algorithm known OPA-PPAR is evaluated using HMDB51 dataset. The results of OPA-PPAR is compared with that of our prior work named multi-task learning based hybrid prediction algorithm (MTL-HPA) and the state of the art method named gradient reversal layer (GRL) [61]. As shown in Figure 2, there are 51 action samples in HMDB51 dataset. Out of them 100 samples are used for empirical study in this paper. However, results are presented in this paper for 10 actions. They include climb, eat, jump, kiss, push, pushup, run, sit, smile and walk. As presented in Figure 3, the input images or frames are shown in left column and the action recognized and anonymized frame are shown in second and third columns respectively. The experimental results are evaluated in terms of precision, recall and F1-Score. The performance values are obtained with human study on anonymized samples. The ground truth and prediction results of the action recognition methods are subjected to evaluation in terms of the measures.

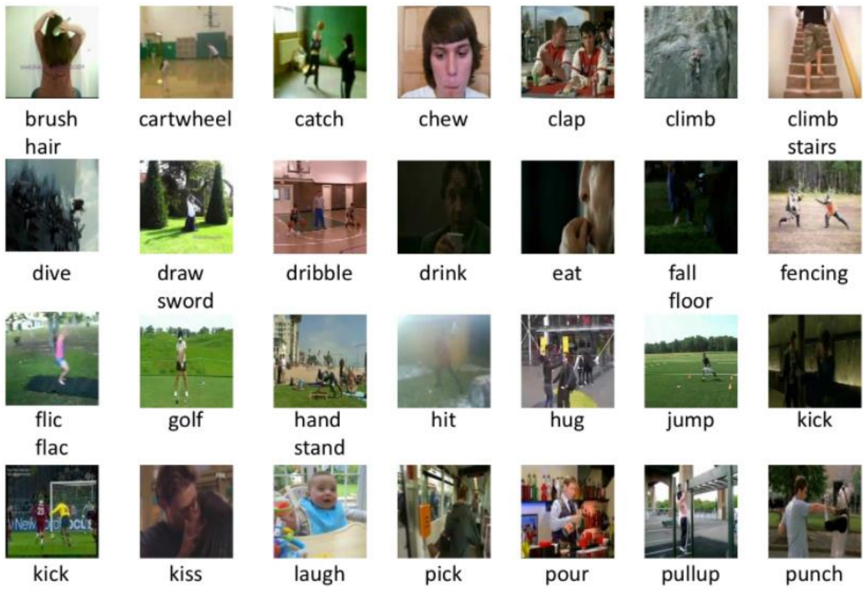

Figure 2. Some human action samples present in HMDB51 dataset [39]

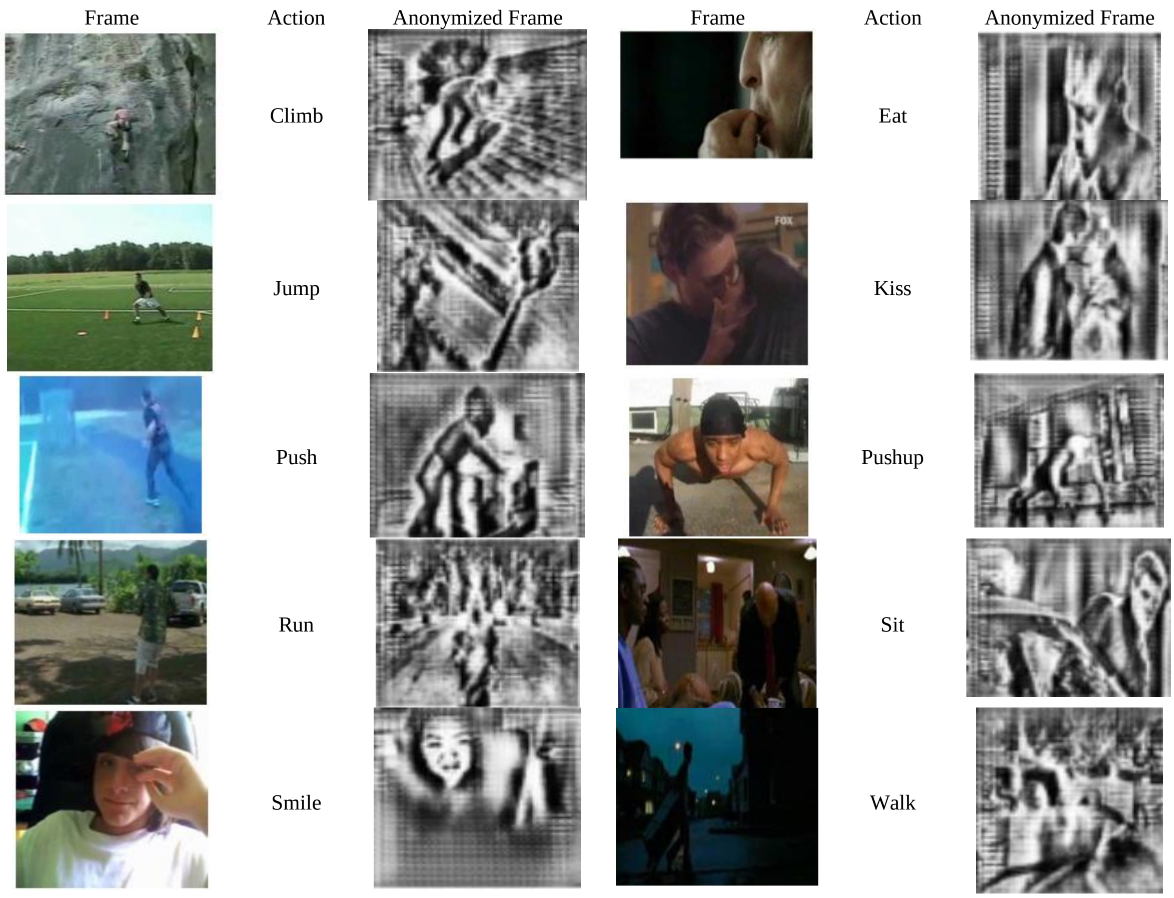

Figure 3. Experimental results for the selection actions

Figure 4 and Figure 5 show the performance comparison between the GRL vs. proposed method and MTL-HPA vs. the proposed method. In both the cases, the action recognition models are presented in horizontal axis and vertical axis shows the performance (%). Observations are made with 10 human actions. For each human action 100 experiments are made with the prototype made to demonstrate proof of the concept. Precision, recall and F1-score are computed based on ground truth and the results of the action recognition models. The final evaluation results are obtained with human study. The results revealed that the proposed method OPA-PPAR outperforms the existing methods known as GRL and MTL-HPA. The MTL-HPA showed significantly better performance over the baseline GRL method. The experimental results revealed that the proposed action recognition method not only preserves privacy and recognizes human actions but also has optimizations in terms of privacy budget and a combined loss function to guide the recognition process associated with the proposed framework. As presented in Table 2 and Table 3, the performance of the proposed method is compared with that of GRL and MTL-HPA.

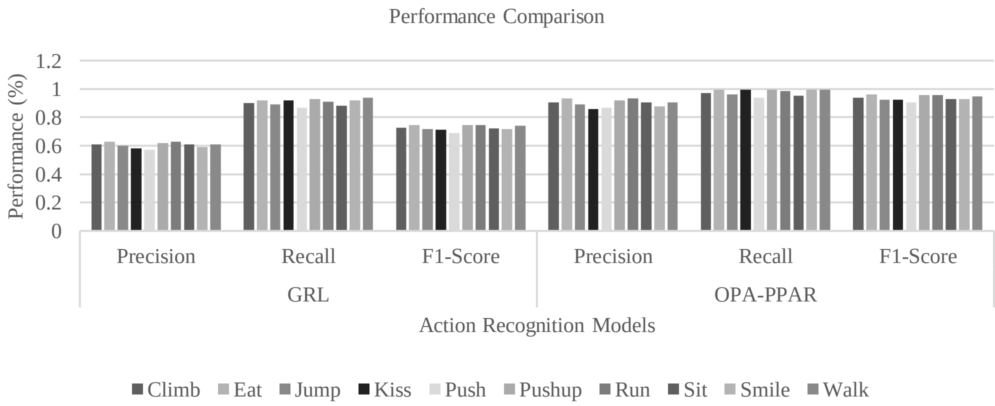

Figure 4. Performance comparison of action recognition models GRL and OPA-PPAR

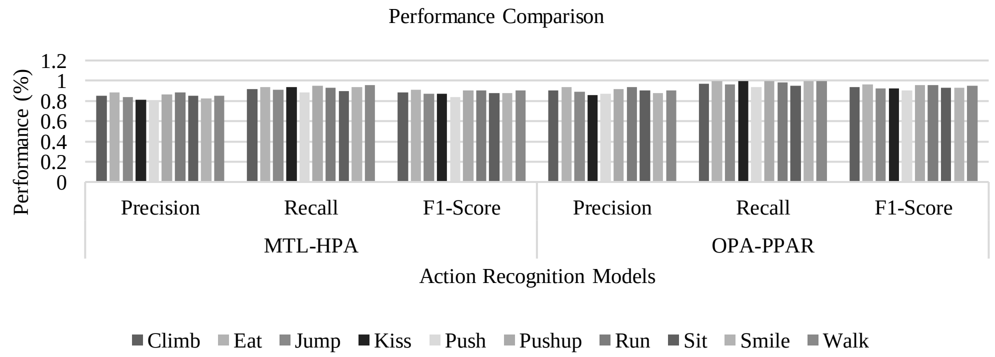

Figure 5. Performance comparison of action recognition models MTL-HPA and OPA-PPAR

Table 2. Results of the proposed method compared with that of GRL

Action	GRL Performance			OPA-PPAR Performance		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Climb	0.61	0.9	0.727152	0.90524	0.97308	0.937935
Eat	0.63	0.92	0.747871	0.93492	0.994704	0.963886
Jump	0.6	0.89	0.716779	0.8904	0.962268	0.92494
Kiss	0.58	0.92	0.711467	0.86072	0.994704	0.922874
Push	0.57	0.87	0.68875	0.86982	0.940644	0.903847
Pushup	0.62	0.93	0.744	0.92008	0.99603	0.95655
Run	0.63	0.91	0.744545	0.93492	0.983892	0.958781
Sit	0.61	0.88	0.720537	0.90524	0.951456	0.927773
Smile	0.59	0.92	0.71894	0.87556	0.994704	0.931337
Walk	0.61	0.94	0.739871	0.90524	0.997152	0.948976

Table 3. Results of the proposed method compared with that of our prior method MTL-HPA

Actions	MTL-HPA			OPA-PPAR		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Climb	0.854	0.918	0.884844	0.90524	0.97308	0.937935
Eat	0.882	0.9384	0.909326	0.93492	0.994704	0.963886
Jump	0.84	0.9078	0.872585	0.8904	0.962268	0.92494
Kiss	0.812	0.9384	0.870636	0.86072	0.994704	0.922874
Push	0.798	0.8874	0.840329	0.86982	0.940644	0.903847
Pushup	0.868	0.9486	0.906512	0.92008	0.99603	0.95655
Run	0.882	0.9282	0.90451	0.93492	0.983892	0.958781
Sit	0.854	0.8976	0.875257	0.90524	0.951456	0.927773
Smile	0.826	0.9384	0.87862	0.87556	0.994704	0.931337
Walk	0.854	0.9588	0.903371	0.90524	0.997152	0.948976

5. CONCLUSION AND FUTURE WORK

We proposed an algorithm known OPA-PPAR based on deep neural networks. It anonymizes video content to have adaptive privacy model that defeats attacks from adversaries. The privacy model enhances the privacy of humans while permitting highly accurate approach towards action recognition. The algorithm is implemented to realize PPHARF. The visual recognition of human actions is made using an underlying adversarial learning process where the anonymization is optimized to have an adaptive privacy model. A dataset named HMDB51 is used for empirical study. Our experiments with using Python data science platform reveal that the OPA-PPAR outperforms existing methods. It can be used in real world applications where PPHARF can fit seamlessly. The experimental results revealed that the proposed action recognition method not only preserves privacy and recognizes human actions but also has optimizations in terms of privacy budget and a combined loss function to guide the recognition process associated with the proposed framework. The proposed method paves way for further investigations in terms of optimizing the three models involved in the system.

REFERENCES

[1] L. Lyu *et al.*, "Towards fair and decentralized privacy-preserving deep learning with blockchain," *CoRR*, pp. 1–4, Jun. 2019, [Online]. Available: http://arxiv.org/abs/1906.01167.

[2] R. M. Alguliyev, R. M. Aliguliyev, and F. J. Abdullayeva, "Privacy-preserving deep learning algorithm for big personal data analysis," *Journal of Industrial Information Integration*, vol. 15, pp. 1–14, Sep. 2019, doi: 10.1016/j.jii.2019.07.002.

[3] J. Weng, J. Weng, J. Zhang, M. Li, Y. Zhang, and W. Luo, "DeepChain: auditable and privacy-preserving deep learning with blockchain-based incentive," *IEEE Transactions on Dependable and Secure Computing*, pp. 1–1, 2019, doi: 10.1109/TDSC.2019.2952332.

[4] L. Lyu *et al.*, "Towards fair and privacy-preserving federated deep models," *IEEE Transactions on Parallel and Distributed Systems*, vol. 31, no. 11, pp. 2524–2541, Nov. 2020, doi: 10.1109/TPDS.2020.2996273.

[5] K. V. Kumar, D. J. Harikiran, M. A. R. Prasad, and U. Sirisha, "Privacy-preserving human activity recognition and resolution image using deep learning algorithms spatial relationship and increasing the attribute value in OpenCV," *International Journal of Advanced Science and Technology*, vol. 29, no. 7, pp. 514–523, 2020.

[6] A. S. Rajput, B. Raman, and J. Imran, "Privacy-preserving human action recognition as a remote cloud service using RGB-D sensors and deep CNN," *Expert Systems with Applications*, vol. 152, Aug. 2020, doi: 10.1016/j.eswa.2020.113349.

[7] Y. Wu, F. Yang, Y. Xu, and H. Ling, "Privacy-protective-GAN for privacy preserving face de-identification," *Journal of Computer Science and Technology*, vol. 34, no. 1, pp. 47–60, Jan. 2019, doi: 10.1007/s11390-019-1898-8.

[8] E. Debie, N. Moustafa, and M. T. Whitty, "A privacy-preserving generative adversarial network method for securing EEG brain signals," Jul. 2020, doi: 10.1109/IJCNN48605.2020.9206683.

[9] M. Maximov, I. Elezi, and L. Leal-Taixe, "CIAGAN: conditional identity anonymization generative adversarial networks," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 5446–5455, doi: 10.1109/CVPR42600.2020.00549.

[10] J. Martinsson, E. L. Zec, D. Gillblad, and O. Mogren, "Adversarial representation learning for synthetic replacement of private attributes," Jun. 2020, [Online]. Available: http://arxiv.org/abs/2006.08039.

[11] Q. Li, Z. Zheng, F. Wu, and G. Chen, "Generative adversarial networks-based privacy-preserving 3D reconstruction," in *2020 IEEE/ACM 28th International Symposium on Quality of Service (IWQoS)*, Jun. 2020, pp. 1–10, doi: 10.1109/IWQoS49365.2020.9213037.

[12] S. Shirai and J. Whitehill, "Privacy-preserving annotation of face images through attribute-preserving face synthesis," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jun. 2019, vol. 2019-June, pp. 21–29, doi: 10.1109/CVPRW.2019.00009.

[13] A. Boulemtafes, A. Derhab, and Y. Challal, "A review of privacy-preserving techniques for deep learning," *Neurocomputing*, vol. 384, pp. 21–45, Apr. 2020, doi: 10.1016/j.neucom.2019.11.041.

[14] M. Malekzadeh, R. G. Clegg, and H. Haddadi, "Replacement autoencoder: a privacy-preserving algorithm for sensory data analysis," *Proceedings - ACM/IEEE International Conference on Internet of Things Design and Implementation, IoTDI 2018*, pp. 165–176, Oct. 2017, doi: 10.1109/IoTDI.2018.00025.

[15] R. Yonetani, V. N. Boddeti, K. M. Kitani, and Y. Sato, "Privacy-preserving visual learning using doubly permuted homomorphic encryption," in *Proceedings of the IEEE International Conference on Computer Vision*, Oct. 2017, vol. 2017-Octob, pp. 2059–2069, doi: 10.1109/ICCV.2017.225.

[16] M. Der Yang, K. S. Huang, Y. F. Yang, L. Y. Lu, Z. Y. Feng, and H. P. Tsai, "Hyperspectral image classification using fast and

adaptive bidimensional empirical mode decomposition with minimum noise fraction," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1950–1954, Dec. 2016, doi: 10.1109/LGRS.2016.2618930.
[17] W. Du *et al.*, "Approximate to be great: communication efficient and privacy-preserving large-scale distributed deep learning in internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 12, pp. 11678–11692, Dec. 2020, doi: 10.1109/JIOT.2020.2999594.
[18] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9906, Springer International Publishing, 2016, pp. 694–711.
[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
[20] H. Kuehne, H. Jhuang, E. Garrote, T. Poggio, and T. Serre, "HMDB: a large video database for human motion recognition," in *2011 International Conference on Computer Vision*, Nov. 2011, pp. 2556–2563, doi: 10.1109/ICCV.2011.6126543.
[21] K. Yun, J. Honorio, D. Chattopadhyay, T. L. Berg, and D. Samaras, "Two-person interaction detection using body-pose features and multiple instance learning," in *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Jun. 2012, pp. 28–35, doi: 10.1109/CVPRW.2012.6239234.
[22] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9908, pp. 630–645, Mar. 2016, doi: 10.1007/978-3-319-46493-0_38.
[23] C. Szegedy *et al.*, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.
[24] R. Leenes, R. van Brakel, S. Gutwirth, and P. De Hert, *Data Protection and Privacy: (In)visibilities and Infrastructures*, vol. 36. Cham: Springer International Publishing, 2017.
[25] J. Dai, B. Saghafi, J. Wu, J. Konrad, and P. Ishwar, "Towards privacy-preserving recognition of human activities," in *2015 IEEE International Conference on Image Processing (ICIP)*, Sep. 2015, pp. 4238–4242, doi: 10.1109/ICIP.2015.7351605.
[26] T. Orekondy, B. Schiele, and M. Fritz, "Towards a visual privacy advisor: understanding and predicting privacy risks in images," in *Proceedings of the IEEE International Conference on Computer Vision*, Oct. 2017, vol. 2017-Octob, pp. 3706–3715, doi: 10.1109/ICCV.2017.398.
[27] F. Pittaluga, S. J. Koppal, S. B. Kang, and S. N. Sinha, "Revealing scenes by inverting structure from motion reconstructions," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 145–154, Apr. 2019, doi: 10.1109/CVPR.2019.00023.
[28] J. Dai, J. Wu, B. Saghafi, J. Konrad, and P. Ishwar, "Towards privacy-preserving activity recognition using extremely low temporal and spatial resolution cameras," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Jun. 2015, pp. 68–76, doi: 10.1109/CVPRW.2015.7301356.
[29] A. Dosovitskiy and T. Brox, "Inverting visual representations with convolutional networks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2016, pp. 4829–4837, doi: 10.1109/CVPR.2016.522.
[30] L. Lyu, X. He, Y. W. Law, and M. Palaniswami, "Privacy-preserving collaborative deep learning with application to human activity recognition," in *International Conference on Information and Knowledge Management, Proceedings*, Nov. 2017, pp. 1219–1228, doi: 10.1145/3132847.3132990.
[31] P. Weinzaepfel, H. Jégou, and P. Pérez, "Reconstructing an image from its local descriptors," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2011, pp. 337–344, doi: 10.1109/CVPR.2011.5995616.
[32] M. S. Ryoo, B. Rothrock, C. Fleming, and H. J. Yang, "Privacy-preserving human activity recognition from extreme low resolution," Apr. 2016, [Online]. Available: http://arxiv.org/abs/1604.03196.
[33] A. Mahendran and A. Vedaldi, "Visualizing deep convolutional neural networks using natural pre-images," *International Journal of Computer Vision*, vol. 120, no. 3, pp. 233–255, May 2016, doi: 10.1007/s11263-016-0911-8.
[34] Z. W. Wang, V. Vineet, F. Pittaluga, S. N. Sinha, O. Cossairt, and S. B. Kang, "Privacy-preserving action recognition using coded aperture videos," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Jun. 2019, pp. 1–10, doi: 10.1109/CVPRW.2019.00007.
[35] F. Al Machot, M. R. Elkobaisi, and K. Kyamakya, "Zero-shot human activity recognition using non-visual sensors," *Sensors*, vol. 20, no. 3, p. 825, Feb. 2020, doi: 10.3390/s20030825.
[36] F. Pittaluga and S. J. Koppal, "Privacy preserving optics for miniature vision sensors," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2015, pp. 314–324, doi: 10.1109/CVPR.2015.7298628.
[37] F. Pittaluga and S. J. Koppal, "Pre-capture privacy for small vision sensors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 11, pp. 2215–2226, Nov. 2017, doi: 10.1109/TPAMI.2016.2637354.
[38] C. Zhang, Y. Tian, and E. Capezuti, "Privacy preserving automatic fall detection for elderly using RGBD cameras," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7382, no. 1, Springer Berlin Heidelberg, 2012, pp. 625–633.
[39] S. N. Sur, S. Bera, S. Shome, R. Bera, and B. Maji, "Target characterization using MIMO radar," *International Journal on Smart Sensing and Intelligent Systems*, vol. 13, no. 1, pp. 1–8, 2020, doi: 10.21307/ijssis-2019-013.
[40] Miran Kim, Xiaoqian Jiang, "HEAR: Human Action Recognition via Neural Netwroks on Homomorphically Encrypted Data", arXiv:2104.0916v1 [cs.CR].
[41] B. Cheng *et al.*, "Robust emotion recognition from low quality and low bit rate video: a deep learning approach," in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, Oct. 2017, pp. 65–70, doi: 10.1109/ACII.2017.8273580.
[42] D. Riboni and C. Bettini, "COSAR: hybrid reasoning for context-aware activity recognition," *Personal and Ubiquitous Computing*, vol. 15, no. 3, pp. 271–289, Mar. 2011, doi: 10.1007/s00779-010-0331-7.
[43] M. Xu, A. Sharghi, X. Chen, and D. J. Crandall, "Fully-coupled two-stream spatiotemporal networks for extremely low resolution action recognition," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Mar. 2018, pp. 1607–1615, doi: 10.1109/WACV.2018.00178.
[44] S. Zolfaghari and M. R. Keyvanpour, "SARF: smart activity recognition framework in ambient assisted living," in *Proceedings of the 2016 Federated Conference on Computer Science and Information Systems, FedCSIS 2016*, Oct. 2016, pp. 1435–1443, doi: 10.15439/2016F132.
[45] B. D. Youn *et al.*, "Statistical health reasoning of water-cooled power generator stator bars against moisture absorption," *IEEE Transactions on Energy Conversion*, vol. 30, no. 4, pp. 1376–1385, Dec. 2015, doi: 10.1109/TEC.2015.2444873.

[46] M. Ciliberto, D. Roggen, and F. J. O. Morales, “Exploring human activity annotation using a privacy preserving 3D model,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, Sep. 2016, pp. 803–812, doi: 10.1145/2968219.2968290.

[47] E. Cippitelli, S. Gasparrini, E. Gambi, and S. Spinsante, “A human activity recognition system using skeleton data from RGBD sensor,” *Computational Intelligence and Neuroscience*, vol. 2016, pp. 1–14, 2016, doi: 10.1155/2016/4351435.

[48] T. Wang, J. Zhao, M. Yatskar, K.-W. Chang, and V. Ordonez, “Balanced datasets are not enough: estimating and mitigating gender bias in deep image representations,” in *Proceedings of the IEEE International Conference on Computer Vision*, Nov. 2018, pp. 5309–5318, doi: 10.1109/ICCV.2019.00541.

[49] B. W. Tseng and P. Y. Wu, “Compressive privacy generative adversarial network,” *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2499–2513, 2020, doi: 10.1109/TIFS.2020.2968188.

[50] X. Jin, Z. Chen, and W. Li, “AI-GAN: asynchronous interactive generative adversarial network for single image rain removal,” *Pattern Recognition*, vol. 100, Apr. 2020, doi: 10.1016/j.patcog.2019.107143.

[51] J. Ma, W. Yu, P. Liang, C. Li, and J. Jiang, “FusionGAN: a generative adversarial network for infrared and visible image fusion,” *Information Fusion*, vol. 48, pp. 11–26, Aug. 2019, doi: 10.1016/j.inffus.2018.09.004.

[52] S. Liu, A. Shrivastava, J. Du, and L. Zhong, “Better accuracy with quantified privacy: representations learned via reconstructive adversarial network,” Jan. 2019, [Online]. Available: http://arxiv.org/abs/1901.08730.

[53] Z. Wu, Z. Wang, Z. Wang, and H. Jin, “Towards privacy-preserving visual recognition via adversarial training: a pilot study,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11220 LNCS, Springer International Publishing, 2018, pp. 627–645.

[54] P. C. Roy and V. N. Boddeti, “Mitigating information leakage in image representations: a maximum entropy approach,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 2581–2589, doi: 10.1109/CVPR.2019.00269.

[55] B. H. Zhang, B. Lemoine, and M. Mitchell, “Mitigating unwanted biases with adversarial learning,” in *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, Dec. 2018, pp. 335–340, doi: 10.1145/3278721.3278779.

[56] Z. Gheid, Y. Challal, X. Yi, and A. Derhab, “Efficient and privacy-aware multi-party classification protocol for human activity recognition,” *Journal of Network and Computer Applications*, vol. 98, pp. 84–96, Nov. 2017, doi: 10.1016/j.jnca.2017.09.005.

[57] Z. Gheid and Y. Challal, “Novel efficient and privacy-preserving protocols for sensor-based human activity recognition,” in *Proceedings - 13th IEEE International Conference on Ubiquitous Intelligence and Computing, 13th IEEE International Conference on Advanced and Trusted Computing, 16th IEEE International Conference on Scalable Computing and Communications, IEEE International*, Jul. 2017, pp. 301–308, doi: 10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0062.

[58] S. J. Oh, R. Benenson, M. Fritz, and B. Schiele, “Faceless person recognition: privacy implications in social media,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9907, Springer International Publishing, 2016, pp. 19–35.

[59] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, “Learning spatiotemporal features with 3D convolutional networks,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, pp. 4489–4497, doi: 10.1109/ICCV.2015.510.

[60] Y. Ganin and V. Lempitsky, “Unsupervised domain adaptation by backpropagation,” *32nd International Conference on Machine Learning, ICML 2015*, vol. 2, pp. 1180–1189, 2015.

[61] D.-Z. Du, *Minimax and its applications*. Springer US, 1995.

BIOGRAPHIES OF AUTHORS

Vijaya Kumar Kambala working as Part Time Research Scholar in School of CSE VIT-AP. he is working as Assistant Professor, in PVP Siddhartha Institute of Technology, Vijayawada. His research interest includes Image Processing, Video Analysis, human activity recognition, privacy preserving human activity recognition. He can be reached at email: vijaya.18phd7017@vitap.ac.in.

Jonnadula Harikiran working as Associate professor, School of CSE, Vellore Institute of Technology, VIT-AP. His reserach interest include microarray image analysis, Hyperspectral Imaging and Video Analytics. He can be reached at email: harikiran.j@vitap.ac.in.