DOI: 10.15849/IJASCA.230320.04

Int. J. Advance Soft Compu. Appl, Vol. 15, No. 1, March 2023
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Novel deep learning system for person re-identification using sequence frames of motion

Nisreen Abd Alhadi Jabr

Branch of Theoretical Sciences, College of Physical Education and Sports Science,
Mustansiriyah University, Baghdad – Iraq
nisreen1987@uomustansiriyah.edu.iq

Abstract

Person re-identification (Re-ID) is in fact a difficult and modern task because of its application and to crucial difficulties of variation of human position, occlusion of human body, variation of camera view, etc. To address this topic, a novel deep learning method is proposed in this paper, a Convolution Neural Network (CNN) model was implemented for Re-ID problem. The main approach is the structure of the developed CNN, this structure performed high accuracy in Re-ID. Many datasets were used for validation and cumulative match characteristic (CMC) curve was adopted for accuracy. Results were compared with other studies in this field. The validation was done using familiar datasets and python language. The results have outperformed other studies in this field.

Keywords*: Deep Learning, Convolution Neural Network (CNN), Person re-identification (Re-ID), Average precision, Cumulative Match Characteristic, Relevance Score.*

1. Introduction

Re-identification provides an advantageous tool for non-invasive biometric validation, monitoring, and human-robot interaction in a variety of applications, from crowd traffic management to individualized healthcare. It aims to visually match pedestrian photos captured from various camera angles. To identify the identity of the query image, the aim is to match one or more sets of query photographs with images of several potential people in the gallery (set) [1]. A computer vision task related to video surveillance, which can be useful for forensic investigations. It consists in automatically analyzing a given set of recorded videos (e.g... from a CCTV system) to search for a given (suspect) individual, starting from an image ("query") which is manually selected on one of the videos. The main problem in person re-identification is how to evaluate the similarity between two images of people. One of the approaches in unconstrained videos (like the ones of most CCTV systems) is to use clothing appearance (color and texture) [2].

This research was implemented under artificial intelligence field in general and deep learning especially; the main target is to implement a person re-identification system. The validation was done using familiar datasets and python language, Validation part was done by using CMC: Cumulative Match Curve.

Received 17 December 2022; Accepted 2 February 2023

Deeply trained models are frequently used in modern models in order to extract features and obtain good performance. It considers a person's entire appearance, including their attire, and can identify someone even if their face is completely hidden [3]. We with aid of re-identification and convert identification to a system of tracking can follow the path of identified person and then check if there is anything that is unlawful or improper is carried out. Vehicles and other items might also be tracked. This allows for the analysis and improvement of the traffic situation. Fig. 1**Error! Reference source not found.** shows the motion of person in simple path.

Fig. 1: Capturing during motion [4].

In our work, a novel CNN was developed for person re-identification, the final results were validated using specific metrics and the research results outperformed the existing studies in the same field, the accuracy on first dataset (DukeMTMC) was 97.96% and on the second dataset (Market dataset) was 98.6%. The final structure can be applied in same projects such as recognition in walking states and other project that depends on images of people in walking mode.

1.1 Person Re-Identification

Re-identification is an important tool for many applications such as field of surveillance, biometric validation, health care, crowded traffic and some-robot interactions application [5]. Person re-id is related to many fields:

- Person recognition
- Gait model features
- Pedestrian application

Person re-identification has significant part in surveillance applications. It is a difficult task due to that identified pedestrian undergo considerable variations across different cameras and various views. In addition, there are a lot of pedestrians to be re-identified in surveillance videos [6].

Application areas for person re-identification in video surveillance include behavior analysis, multi-camera monitoring, and pedestrian search. Visual feature matching and spatial and temporal reasoning can be combined in multi-camera object tracking scenarios. Despite the positive developments in person re-identification, this process is still exceedingly difficult and has numerous problems that have not yet been resolved [7]:

- Variations of camera setting
- Changes of human pose
- Illumination variations as well as background clutter and occlusions
- Images rotation
- Moreover, people might have similar visual appearance

Those make the re-id problem more challenging. Fig. 2 shows how to match a target person's probe image to a collection of gallery images and various image variations [8].

Fig. 2: Matching a probe image [9].

(a) Comparing gallery photos to a probe image. (b) Image variations (from left to right): poor resolution, people with similar appearances, changing camera views, occlusions, and differences in illumination. The re-id system just needs a first image of the subject, which it then thoroughly examines. People can be identified from a variety of camera perspectives, including the front, side, and even the back. Because it considers a person's entire appearance, including clothing, it is feasible to identify someone even when their face is completely hidden [10].
Fig. 3 shows the matching.

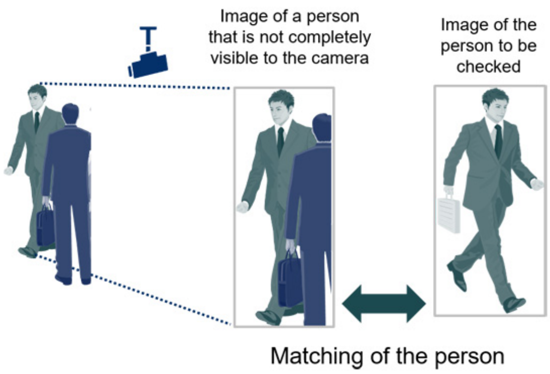


Fig. 3: Matching.

It is utilized most frequently in situations involving video monitoring. When there are numerous cameras installed all around a campus, retail mall, parking lot, or any other place and we want to ensure security.
Here, we define some related terms,
Cameras with disjoint views: assume two cameras A and B are sequentially positioned next to a walkway. Let's say that after a specific amount of time, those who were initially seen entering camera A and going along the path in one direction are also seen doing so. People traveling in the other direction in camera A, however, might not be seen later in camera B. As a result, correspondences can be limited by taking into account the typical

locations of exits and entry between cameras, the direction of movement, and the typical travel duration between A and B [11].
Probe image: is a test image to be matched against the gallery images. *Gallery:* set of images of persons that we have trained the system using their images. *Matching:* fetch probe image in gallery. *Ranked list:* list of ranks that refers to images with top ranks; top ranks mean correct prediction of probe images.

2. Related Work

Authors in [12], addressed RE-ID problem by suggesting an unsupervised approach of re-id deep learning. Their proposed model was able to use the discriminative details of pictures information from automatically generated person track let data from videos. In order to maximize the discovery of the most likely tracklet relations across camera views, they developed a Tracklet Association Unsupervised Deep Learning (TAUDL) framework, which is featured by combining learning by cameras with labels from tracklet association. They worked on 6 re-id datasets, extensive tests show that the proposed system outperformed state-of-the-art unsupervised and domain adaption re-id approaches. The used datasets were CUHK03, Market-1501 and DukeMTMC.

In [13], The Re-ID between two types of images RGB and infrared was studied by authors. When employing infrared, it is challenging to acquire a reliable measure for such a large-scale cross-modality retrieval because there aren't enough discriminative details to re-identify the same person between infrared and RGB modalities. The authors developed a brand-new cross-modality generative adversarial network to address these problems (named cmGAN). They created a cutting-edge generative adversarial training based discriminator to learn discriminative feature representation from diverse modalities in order to address the issue of insufficient discriminative information. In order to reduce inter-class ambiguity and increase cross-modal similarity between instances, they combined identification loss and cross-modality triplet loss. The standard deep neural network framework allows for end-to-end training of the entire cmGAN. Results showed a superior performance using Cumulative Match Characteristic curve (CMC) and Mean Average Precision (MAP) as validation metrics. The used dataset was SYSU RGB-IR Re-ID dataset.

According to authors in [14], the majority of current models estimate the similarity of various probe and gallery picture pairings individually while ignoring the relationship data between various probe-gallery pairs. This will have the effect of making some hard sample similarity estimates less accurate. To get over these restrictions, authors presented the Similarity-Guided Graph Neural Network (SGGNN), and unique deep learning framework. Given a probe image and a number of gallery images, SGGNN generates a graph to describe the relation between probe gallery pairs (nodes) in order to edit the probe gallery relation features end-to-end. When updated probe-gallery relation features are used for prediction, accurate similarity estimate may be accomplished. The related properties of various probe-gallery image pairings serve as the input features for nodes on the graph. The messages travelling in SGGNN then update the probe-gallery relation feature, which estimates similarity by taking into account information from other nodes. SGGNN directly use gallery labels of instance pairing to learn the edge weights, which allows relation fusion with more exact information, in contrast to standard GNN techniques. The effectiveness of their suggested method is to validate on three known re-id datasets CUHK03, Market-1501 and DukeMTMC.

Authors in [15, 16] have worked with same adopted datasets and worked with deep learning model.

In [17], authors suggested a technique for learning features and a matching similarity metric for re-identification of people. An advanced convolutional architecture with layers specifically created to deal with the re-identification issue. Their network generates a similarity value from a pair of input photos that indicates whether they both show the same individual. A layer that computes cross-input neighborhood differences, which capture local correlations between the two input images based on midlevel features from each input image, is one of our architecture's novel features. A layer of patch summary characteristics, which are later spatially integrated, computes a high-level summary of the outputs of this layer. On a big data set (CUHK03) as well as a medium-sized data set (CUHK01), our technique performs noticeably better than the state of the art and is resistant to overfitting. We also show that even on a tiny data set, our network can produce results comparable to the state of the art by first training on an unrelated huge data set before fine-tuning on it (VIPeR).

3. Proposed method

Firstly, dataset might be divided into train part and test part as shown in Fig. 4

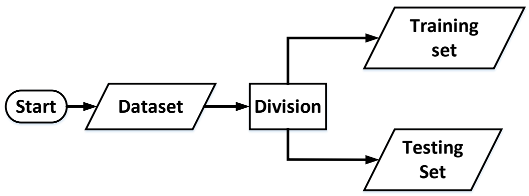


Fig. 4: Dataset division.

All images must be labelled with known label for each dataset as shown in Fig. 5

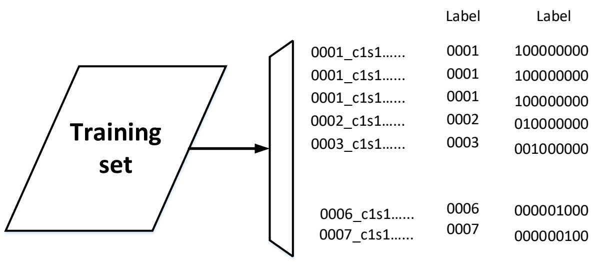


Fig. 5: Image labelling.

The limitations of our work can be summarized into some points:

- The dataset images have small size, since datasets with big size of images will need developed hardware to be used. The existence datasets have small size of images
- The dataset should have a sequence of images of the person or images from different cameras located in different angles.

Many studies have been done in this field, but we made a structure and the results showed its effectiveness, it is outperformed other studies. The developed approach can be summarized as follows:

- The structure of layers

- Batch normalization layers
- Parameters of layers
- The selection of parameters values

3.1Validation Method

CMC curve: is explained in previous section.

mAP: mean Average precision, the average of AP. compute the AP for each class and then average them. Another definition: is the area under the precision-recall curve. In person re-ID, the gap between every recall value.

Performance of an identifying system is gauged using the Cumulative Match Curve (CMC). It rates an identifying system's capacity for ranking. The ROC curve of a verification system, or relative operating characteristic curve. It depends on how far apart the true class and the forecast are from each other. Fig. 6 shows an example of CMC curve.

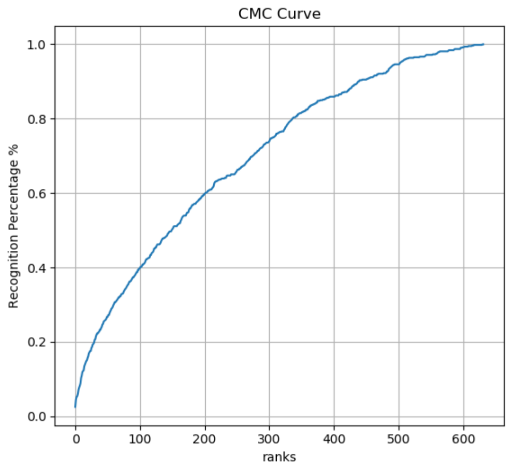


Fig. 6: CMC example.

Main Steps:

Define RANK vector containing zeroes with a shape equal to m where m is the number of query images;

For each image should store its ID (every image has an ID, e.g. 111_45.bpm the ID is "111")

1. Each query image Qi is compared (compute distance) against all template gallery.

2. Results should be sorted by distance.

3. These distances should be analyze to find the index of the correct identity (ID_Q==ID_T) among all template.

4. At that index the RANK, increase one.

The steps 1 - 4 should be executed for each probe.

CMC is a normalized cumulative sum of RANK (RANK/m) Consider five courses and five tests (Test: T) with Euclidian distance in order to comprehend it. The less resemblance is obtained, presume findings, the further a test is from a class (Class: C).

T1 VS C1= .96
T1 VS C2= .71
T1 VS C3= .91
T1 VS C4= .72
T1 VS C5= .32

T1 and C1 are more comparable than the other classes. Test 1 is thus appropriately identified as being in the top rank. Imagine the outcomes:

T1 VS C1= .91
T1 VS C2= .96
T1 VS C3= .39
T1 VS C4= .75
T1 VS C5= .26

Test 1 is more like class 2 than it is like class 1. In other words, test 1 is acknowledged as one of the top "two matches," rather than the top match. This is what rank two recognition means!

- If all tests have top rank, so CMC will be 100% for all classes.

Consider that tests 1, 2, and 3 are acknowledged in the first rank, followed by tests 4, and 5, which are recognized in the second and fourth ranks, respectively.

Rank 1: 60% (tests 1, 2, and 3 are some of the best matches for N=1).

Rank 2: 80% (the first four tests are among the top N=2 matches).

Rank 3: 80% (the first four tests are among the top N=3 matches).

Rank 4: 100% (all test are among the top N=4 matches).

Rank 5: 100% (all test are among the top N=5 matches).

3.2 Pedestrian descriptors and similarity metrics

3.2.1 CMC curve

For each rank R the value of the CMC curve is computed by answering this question: what is the fraction of query images for which the same identity in the template gallery is ranked lower or equal to R?

3.2.2 RS metric

Relevance Score is to associate to each input image of the train set onto the images to which they belong. Suppose for person Michel; train set has 10 images for TOM. For his query image; suppose prediction vector is

V=[1,1,2,3,4,5,10,1,2,7,8,123,32,43,14,96,47,1,2,32,54,T,T,T,T,T,T,T,T,3,4,5,6,2,T,T]

In first 20 predictions, there is no correct prediction, so relevant score for Michel is zero in first 20 ranks. In first 30 predictions, there eight correct predictions, so relevant score for Michel is eight in first 30 ranks. Here is an additional explanation of average precision. Imagine if you type something into Google and it returns 10 results. It would probably be ideal if they were all pertinent. It is much preferable if the relevant ones are shown first if only a small number, let's say five, are. If the first five were unimportant and the good ones only began after the sixth, that would be horrible. Therefore, if mAP starts with high values it means the relevant images are predicated first, but if it starts with low value that means wrong persons are shown first and that behavior is undesirable.

For example above:

V=[1,1,2,3,4,5,10,1,2,7,8,123,32,43,14,96,47,1,2,32,54,T,T,T,T,T,T,T,T,3,4,5,6,2,T,T]

In first 20 predictions, there is no correct prediction. This is worse performance. In first 30 predictions, there eight correct predictions, so it is bad performance but it is not the worse. Best performance is when the predicted vector is:

V=[T,T,T,T,T,T,T,T,T,T,2,3,4,5,10,1,2,7,8,123,32,43,14,96,47,1,2,32,54 ,3,4,5,6,2,1,2]
Good performance is when the predicted vector might be like this:

V=[T,T,T,T,T,1,2,3,T,T,T,T,T,2,3,4,5,10,1,2,7,8,123,32,43,14,96,47,1,2,32,54 ,3,4,5,6,2,1,2]

In addition, performance is measured according to search set or top ranks that we calculate first 20 ranks, first 30 ranks, first rank or all ranks.

Relevance score is related directly to validation metrics such as mAP mean average precision or only AP; AP is a used metric of accuracy on application of object detection, classification, identification or recognition tasks.

Using Method RS: For image with person 3: the result was Fig. 7:

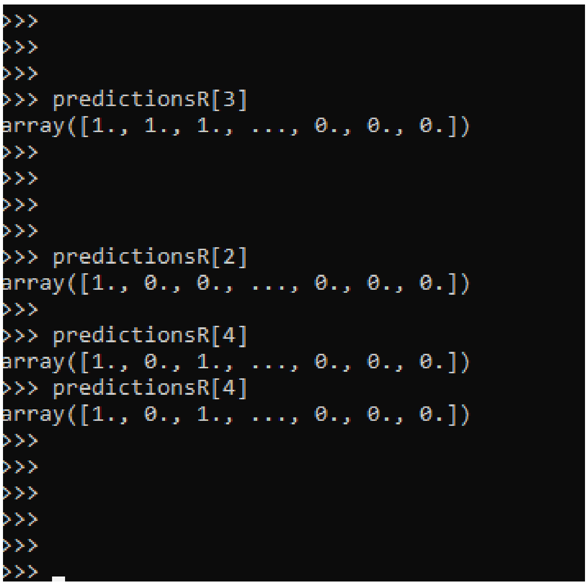


Fig. 7: Example to Calculate AP.

Where one refers to Relevant and zero to Non-relevant. Calculating the AP parameter with the equation [20]:

$$AP_i = \frac{1}{Q_i}\sum_{n=1}^{N}\frac{R_i^n}{n}t_n^i$$

where Q_i is the number of images that are relevant to the i_th query, N is the total number of images in the search set, R_i^n in is the number of images that are relevant to the t_n^i query that were retrieved from the top n results, and t ni indicates whether the n_th image that was retrieved is relevant (=1) or not (=0) for the i_th query.

We have calculated accumulate AP (AAP) as follows then normalize it:

$$AP_i = AP_{i-1} + \frac{1}{Q_i} \sum_{n=1}^{N} \frac{R_i^n}{n} t_n^i$$

The calculated AP is for n ranks (N = all ranks).

The AP equals to one for all images; that is the best result and that means all relevant images for all query images in the first ranks.

For first 50 images the equation will be:

$$AP_i = \frac{1}{Q_i} \sum_{n=1}^{50} \frac{R_i^n}{n} t_n^i$$

3.3 CNN Model

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm that deals with 2D data as input and try to extract features from the input. It has ability to distinguish between two inputs. The general design of a ConvNet is similar to the connecting neurons as in human body. Individual neurons only respond to stimulate in this restricted region of the visual field, known as the Receptive Field. A combination of these fields covers the full visual field [21, 22].
The general design is shown in Fig. 8:

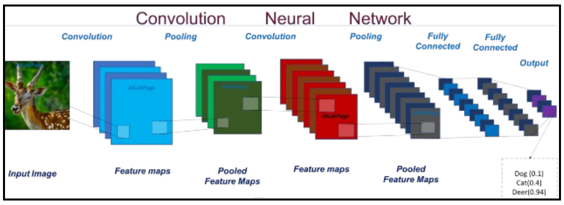


Fig. 8 General design of CNN.

The standard deep CNN used in this study consists of: Convolution layer, pooling layer, fully connected layer (FC), input and output. Every component has unique parameters, and the convolution layer has:

- Filter size: how many filters of convolution are existed.
- Strides: Convolution operation steps.
- Padding simply involves adding extra rows or columns of zeros to the input images to solve the size issue that can arise after convolution.
- Activation: apply activation formula to the data (many types of activation function are existed)
- Regularizer: is used to reduce over-fitting.

The following factors affect the pooling layer: (kernal size, Strides: The operation's pooling strides, Padding is identical to the convolution layer. Fully connected layer is named FC and it is a simple feed forward neural networks. It is the ast Layer in the network. The output from the last pooling or convolutional layer is flattened before being fed into the fully connected layer as the input. It has (Activation function and Neurons)

There is also a layer of regression that has:

- Loss function: it is important in any ML model because it specifies an objective by that the model's effectiveness is judged.
- Learning Rate: aids in accelerating convergence. The output can be significantly impacted by selecting the incorrect learning rate.
- Optimizer: is used to reduce the given loss function.

Dropout layer can be added between layers. It has a term Keep Probability, which its value determine how many neurons will be deleted or kept for the over/under fitting issues. After get trained model, we should be able to calculate similarity between two images. How much each two images are correlated. Fig. 9 shows the usage of CNN model.

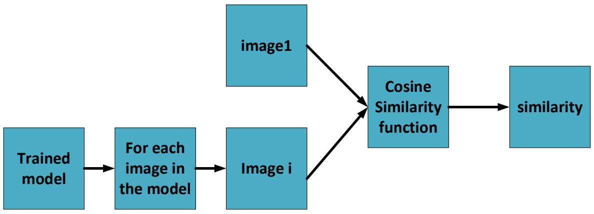


Fig. 9: CNN Model usage.

Figures below show stages of CNN, Fig. 10 shows convolution stage

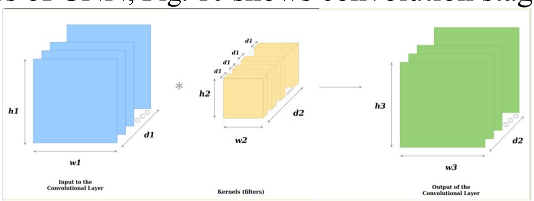


Fig. 10: Convolution stage.

The output has dimensions according to the input and the kernal. There are equations for calculating the output dimensions, but the depth will inevitably be like the kernel. Figure below shows how convolution moves; depth is K, Fig. 11 shows convolution motion

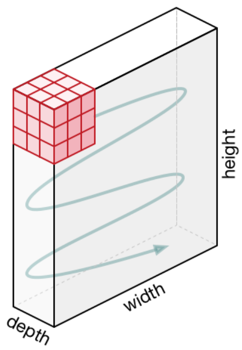


Fig. 11: Convolution motion.

The Pooling layer reduces the spatial size of the Convolved Feature. The amount of processing resources required to process the data will be decreased through dimensionality reduction. Additionally, it aids in the extraction of rotational and positional invariant dominant traits, maintaining the efficiency of the model training procedure.

There are two types of pooling average and max as shown in Fig. 12.

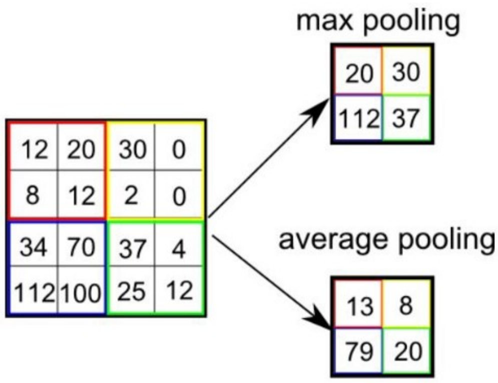


Fig. 12: Max vs Average pooling.

By adding a Fully Connected layer (as shown in Fig. 13), it is possible to learn non-linear combinations of the high-level features that are represented by the output of the convolutional layer at a relatively low cost. The Fully Connected layer in that region is now learning a function that may not be linear.

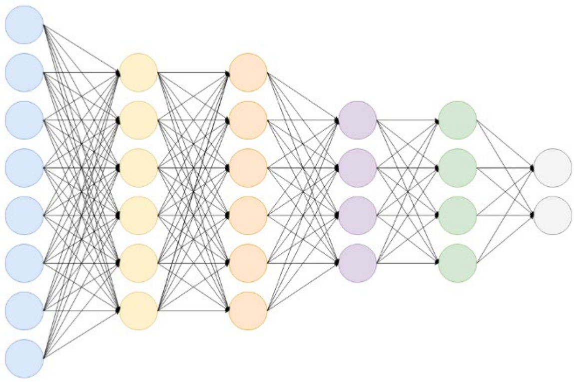


Fig. 13: Normal FC.

Final design is shown in Fig. 14:

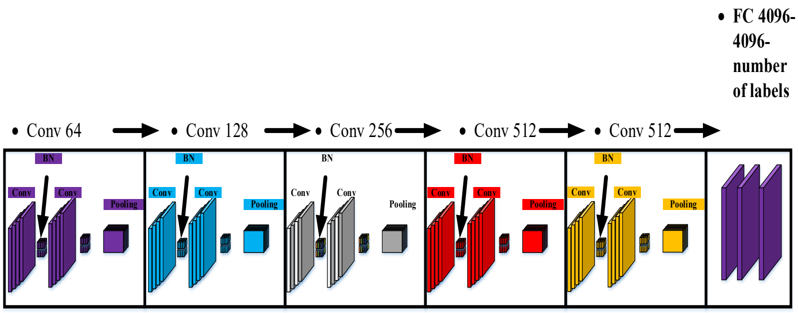


Fig. 14: Final CNN design.

This design is chosen after many researches and after a comprehensive reading about components of CNN. The injection of Batch normalization between two convolution layers. The number of filters in each convolution layer is chosen carefully. It is all come up with final developed cnn system with high performance

4. Datasets

There are many datasets in this field of research, three are mentioned here, and they are the most used by researchers. They are chosen carefully to be able to compare the results with other researchers

4.1 Market 1501 dataset

The following photographs were taken in front of a supermarket by Tsinghua University and are included in this dataset [18]:

- 6 cameras (one low resolution camera and 5 high resolution
- Environment in general (in front of supermarket)
- The image name includes:
  - Camera id, frame sequence and person number.
  - Examples person number 0001, camera c1, and frame s1 0001_c1_s1_002301.
- More than 12900 photos of 1501 identities
- Every person is present in at least two photographs from two cameras.

This dataset has 1,501 IDs and 32,668 bounding boxes with annotations. Cross-camera searching is made possible by each annotated identification appearing in at least two cameras.

4.2 Duke MTMC dataset

(Multi-Target, Multi-Camera) is a dataset of security video footage from the Duke University campus. It was photographed in 2014 and is used to create low-resolution facial recognition software, video tracking systems, and re-identification of individuals. The collection includes more than 2 million frames of 2,000 students walking to and from courses over the course of more than 14 hours of synchronized surveillance video from eight cameras at 1080p and 60 frames per second. In order to catch students "during moments between lectures, when pedestrian traffic is heavy," the 8 surveillance cameras installed on campus were intentionally set up to do so.

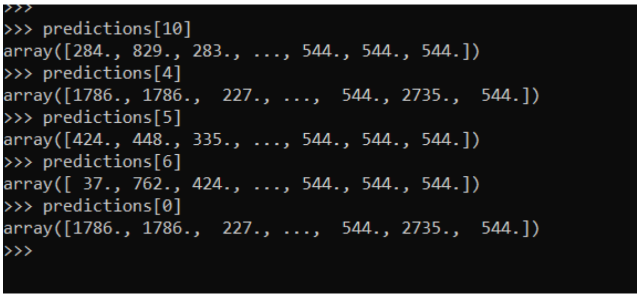
Included: Bounding box train: 16522 pic & bounding box test: 17661 pic.

4.3 CUHK01

The dataset contains images that are captured by two images of 971 persons in campus environment, and each person has two images. Included: 3884 images manually cropped [19].

5. Experimental results

Here are all tests that we have made for all the studied methods, the output of each method is a vector that contains distances or similarities between query images and images from train set, output without CMC is like below (in Fig. 15) a vector contains distances from image ex:10 to all images in the train set.



```
>>>
>>> predictions[10]
array([284., 829., 283., ..., 544., 544., 544.])
>>> predictions[4]
array([1786., 1786.,  227., ...,  544., 2735.,  544.])
>>> predictions[5]
array([424., 448., 335., ..., 544., 544., 544.])
>>> predictions[6]
array([ 37., 762., 424., ..., 544., 544., 544.])
>>> predictions[0]
array([1786., 1786.,  227., ...,  544., 2735.,  544.])
>>>
```

Fig. 15: Output without CMC.

5.1 Results on Market dataset

Market dataset, which contains 3541 images of 199 persons, is split into train and test parts (test part was 475 images). Fig. 16 shows AP Curve on Market dataset

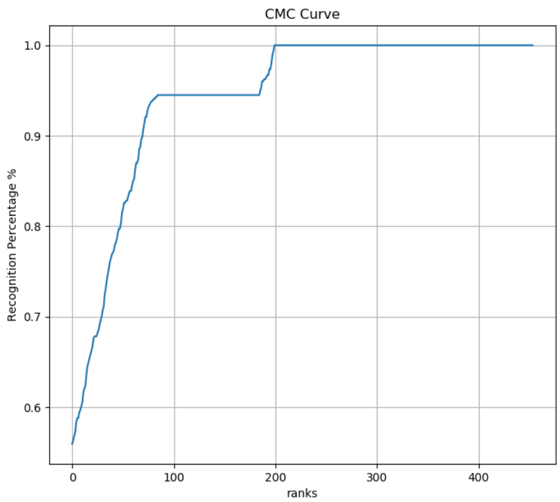


Fig. 16: AP Curve, Market dataset.

This tells that the model might identify first person right with accuracy 55%. But the person 50 will identify him with accuracy 80%.

5.2 Results on DukeMTMC dataset

Training images: 5011 images. DukeMTMC. Testing images: 600 images from DukeMTMC. Fig. 17 shows AP Curve on DukeMTMC

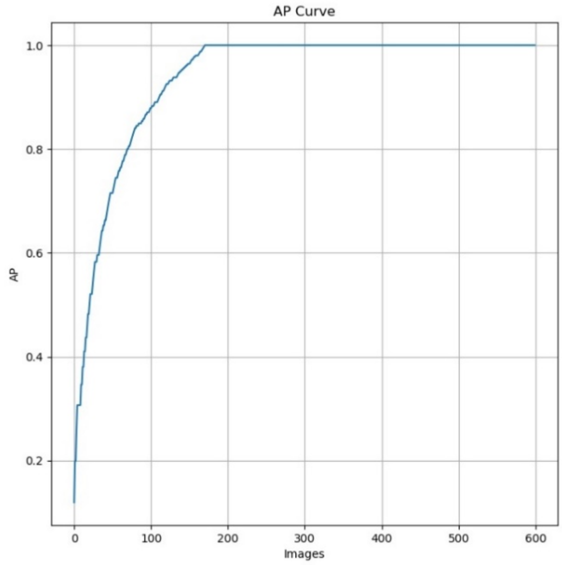


Fig. 17: AP Curve, DukeMTMC dataset.

This tells that the model might identify first person right with accuracy 10%. But the person 50 will identify him with accuracy 60%.

5.3 Results on CUHK dataset

Training images: 3000 images. Testing images: 400 images. Fig. 18 shows AP Curve on

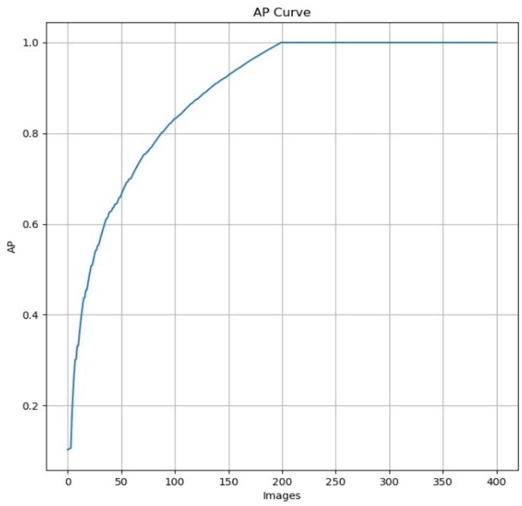


Fig. 18: AP Curve, CHUK dataset.

This tells that the model might identify first person right with accuracy 7%. But the person 50 will identify him with accuracy 65%.

5.4 Relevance Score of different ranks for all datasets

- *DukeMTMC:* Fig. 19 shows top 100 ranks of Dukemtmc dataset. Fig. 20 shows top 25 ranks of Dukemtmc dataset

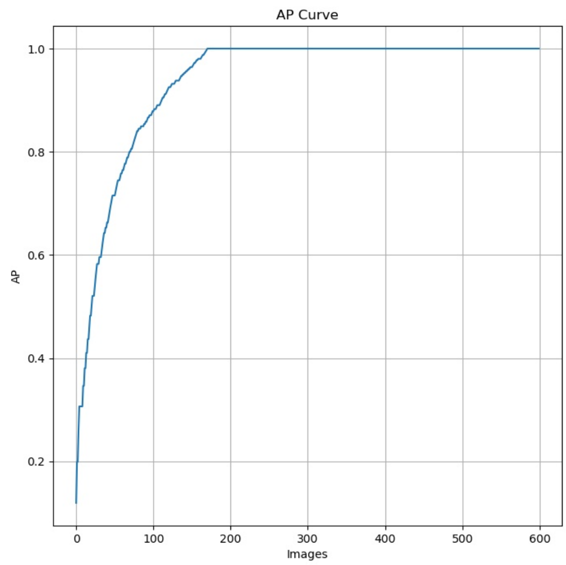


Fig. 19: Top 100 ranks.

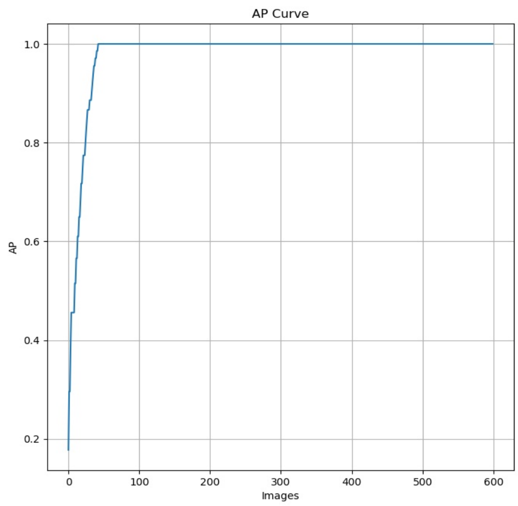


Fig. 20: Top 25 ranks.

- *CUHK*: Fig. 21 shows top 100 ranks of Dukemtmc dataset. Fig. 22 shows top 25 ranks of Dukemtmc dataset

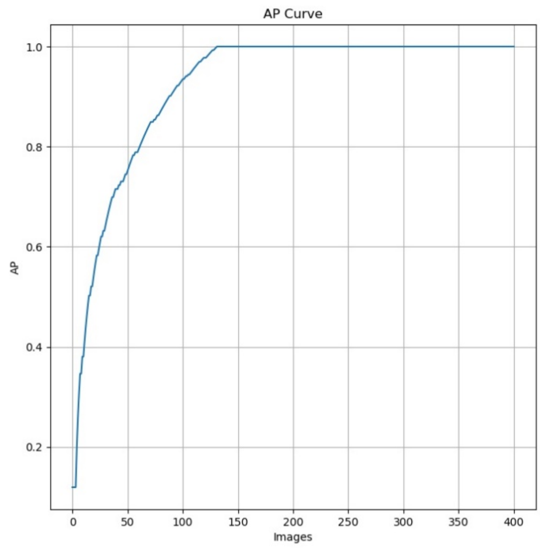


Fig. 21: Top 100 ranks.

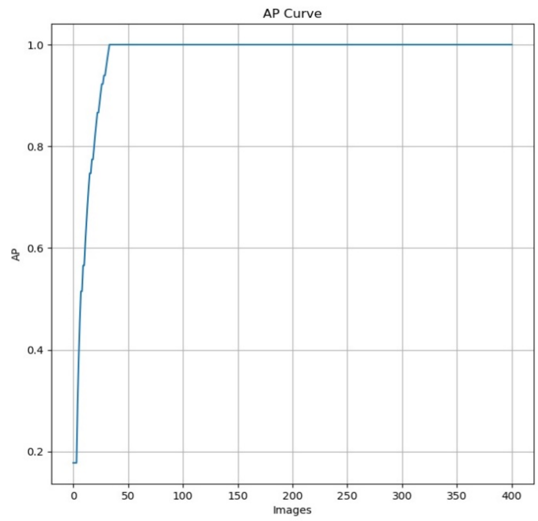


Fig. 22: Top 25 ranks.

5.5 Comparison: Results Summary

Summarize results for Comparing with other studies.

All ranks	100 ranks	25 ranks	Dataset
98.6	98.96	99.62	Market
79.96	93.66	97.93	DukeMTMC

Table.1. Summarize results for Comparing with other studies.

Comparing with state of art:

Study	Details	mAP	Dataset

TAUDL [12**Error! Reference source not found.**]	Unsupervised Reid	41.2	Market
TAUDL [12]	Unsupervised Reid	43.33	DukeMTMC
SGGNN [13]		82.8	Market
SGGNN [13]		68.2	DukeMTMC
Lin et al in [15],	Baseline 1 Baseline 2	43.50 24.17	DukeMTMC
Zheng et al in [16]	ResNet-50	70.33	Market
Ours	NCNN	79.96	DukeMTMC
Ours	NCNN	98.6	Market

Table.2. Comparing with state of art.

5.6 Results discussion

The results showed that the developed approach is reliable and robust in Re-ID systems. Even if dataset images have small size, the developed system resulted in high identification accuracy. The main reason of this is the structure itself and specifically the batch normalization layers in the design.

6 Conclusions

In this project, we have implemented a project of person re-identification using deep learning model; we have developed our skills in many things, Main useful points:

- ❖ Understanding Person re identification.
- ❖ Implementation many methods on different datasets.
- ❖ Develop these algorithms to get better results.
- ❖ Compare among these methods and previous research.
- ❖ Understanding CMC Curve and implement it.
- ❖ Enhance results by implement Relevant Score.

For future work, this system might be validated on different datasets and implemented in real time. Optimization algorithm might be used to adapt some parameters of the deep learning model.

References

[1] Subramaniam, A., Nambiar, A., & Mittal, A, " Co-segmentation inspired attention networks for video-based person re-identification," in Proceedings of the IEEE/CVF international conference on computer vision, 2019, pp. 562-572.

[2] Yin, J., Wu, A., & Zheng, W. S., “Fine-grained person re-identification”, International journal of computer vision, vol. 128, no.6, pp. 1654-1672, 2020.

[3] Huang, Y., Wu, Q., Xu, J., Zhong, Y., & Zhang, Z., “Clothing status awareness for long-term person re-identification”, in Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 11895-11904, 2021.

[4] Tian, Y., Wei, L., Lu, S., & Huang, T., "Free-view gait recognition", PloS one, vol.14, no. 4, e0214389, 2019.

[5] Saleh, A. M., & Hamoud, T., "Analysis and best parameters selection for person recognition based on gait model using CNN algorithm and image augmentation", Journal of Big Data, vol. 8, no.1, pp. 1-20, 2021.

[6] Li, W., Zhao, R., Xiao, T., & Wang, X., "Deepreid: Deep filter pairing neural network for person re-identification", in Proceedings of the IEEE conference on computer vision and pattern recognition, pp.152-159, 2014.

[7] Liu, J., et al., "Multi-scale triplet cnn for person re-identification", in Proceedings of the 24th ACM international conference on Multimedia, pp. 192-196, 2016, October.

[8] Chong, Y., et al., "Learning domain invariant and specific representation for cross-domain person re-identification", Applied Intelligence, vol. 51, no. 8, pp. 5219-5232, 2021.

[9] Mortezaie, Z., Hassanpour, H., & Beghdadi, A., "People re-identification under occlusion and crowded background", Multimedia Tools and Applications, pp. 1-21, 2022.

[10] Wei, W., Yang, W., Zuo, E., Qian, Y., & Wang, L., "Person re-identification based on deep learning-An overview", Journal of Visual Communication and Image Representation, 103418, 2021.

[11] Yu, H. X., Wu, A., & Zheng, W. S., "Unsupervised person re-identification by deep asymmetric metric embedding", IEEE transactions on pattern analysis and machine intelligence, vol. 42, no. 4, pp. 956-973, 2018.

[12] Li, M., Zhu, X., & Gong, S., "Unsupervised person re-identification by deep learning tracklet association", in Proceedings of the European conference on computer vision (ECCV), pp. 737-753, 2018.

[13] Dai, P., et al., "Cross-modality person re-identification with generative adversarial training", in IJCAI, vol. 1, no. 3, p. 6, 2018, July.

[14] Shen, Y., et al., "Person re-identification with deep similarity-guided graph neural network", in Proceedings of the European conference on computer vision (ECCV), pp. 486-504, 2018.

[15] Lin, Y., et al., "Improving person re-identification by attribute and identity learning", Pattern Recognition, vol. 95, pp.151-161, 2019.

[16] Zheng, Z., Zheng, L. and Yang, Y., "A discriminatively learned cnn embedding for person reidentification", ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 14, no. 1, pp.1-20, 2017.

[17] Ahmed, E., Jones, M., & Marks, T. K. (2015). An improved deep learning architecture for person re-identification. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3908-3916).

[18] Zheng, Liang and Shen, Liyue and Tian, Lu and Wang, Shengjin and Wang, Jingdong and Tian, Q. (2015). Scalable Person Re-identification: A Benchmark. Computer Vision, IEEE International Conference on

[19] Li, Wei and Zhao, Rui and Wang, Xiaogang. (2012). Human Reidentification with Transferred Metric Learning. ACCV.

[20] Putzu, L., Piras, L. and Giacinto, G., "Ten Years of Relevance Score for Content Based Image Retrieval", in International Conference on Machine Learning and Data Mining in Pattern Recognition, pp. 117-131. Springer, Cham, 2018, July.

[21] Luo, H., et al., "Convolutional neural networks: Computer vision-based workforce activity assessment in construction", Automation in Construction, vol. 94, pp.282-289, 2018.

[22] Yamashita, R., Nishio, M., Do, R.K.G. and Togashi, K., "Convolutional neural networks: an overview and application in radiology", Insights into imaging, vol. 9, no. 4 , pp.611-629, 2018.

Notes on contributor

Nisreen Abd Alhadi Jabr received the B.Sc. in Computer Science from Mustansiriyah University- Iraq (2009), and M.Sc. in Computer Science from Mustansiriyah University-Iraq (2016). She has nearly 12 years of experience in teaching in Computer science department. She’s currently a lecturer with the Branch of Theoretical Sciences, College of Physical Education and Sports Science, Mustansiriyah University. She has many published papers in the Computer Science field. Her research interests are in Artificial Intelligence, pattern recognition, and image processing.