

사람 재인식을 위한 개선된 PersonNet

Advanced PersonNet for Person Re-Identification

박 성 현*, 강 석 훈*

Seong-Hyeon Park*, Seok-Hoon Kang*

Abstract

This paper propose and experiment advanced PersonNet, a human identification model, with advanced performance. We apply the inception layer to extract feature points, and increase the existing 32 feature points to 154. Also, we modify the CND method used by PersonNet to mitigate asymmetry, and apply weights to the feature map of pedestrian images in three parts, thereby making the features more distinct. Three databases were used for performance evaluation : CUHK01, CUHK03 and Market-1501. The experiment results showed 27-31% improvement in performance.

요 약

이 논문에서는 사람 재식별 모델인 PersonNet의 성능을 개선하는 방법을 제안하고 실험한다. 특징점 추출을 위해 인셉션 레이어를 접목하여, 기존 32개의 특징점을 154개로 증가시켜 강화하였다. 또한, PersonNet에서 사용하는 CND 방식을 수정하여 비대칭성을 완화하였고, 보행자 이미지의 특징점을 3부분으로 나누어 가중치를 적용한 방법을 적용하여 특징을 더 뚜렷하게 파악하도록 하였다. 성능 평가를 위해 CUHK01, CUHK03 그리고 Market-1501 3가지의 데이터베이스를 사용하였고 실험 결과 27~31% 성능이 개선되었다.

Key words : PersonNet, CNN, Inception Layer, Cross neighborhood difference, Person re-Identification

1. 서론

사람을 인식하는 주목받는 학습 모델 중에, Siamese Network[1]를 보행자 이미지 처리에 특화시켜 발전시킨 PersonNet[2]이 있다. 사람 재인식의 주목적은 여러 카메라에서 관측된 다양한 특징을 가진 보행자들을 서로 동일인인지 판단하는 것이다. 한 쌍의 보행자 이미지를 비교, 유사성을 판단해, 높은 유사 점수를 받은 쌍은 동일인으로 판단하고 그렇지 않으면 서로 다른 사람으로 판단한다. PersonNet은 CNN 계층을 이용해 두 이미지에

서 특징점을 뽑아내어 이를 비교하는 방식을 사용한다. CNN 계층을 이용한 VGG[3]모델의 경우 3x3 컨볼루션 레이어를 16~19개를 쌓아 구성하였고 객체 분류에서 7.32%의 오류율을 보였다.[3] 하지만 이처럼 망이 깊어지게 되면 파라미터 수가 늘어나게 되고 학습시간과 메모리 사용량에 많은 영향을 미치게 되므로, 모델의 망 내 연결을 줄이고 세부 행렬 연산은 최대한 밀집하게 구성할 필요가 있다. 본 논문에서는 객체 분류에서 6.67%의 오류율을 보인 Inception 모델[4]에 기반을 두어, 개선된 PersonNet을 제안한다. 본 논문에서는 기존의

* Dept. of Embedded Systems Engineering, Incheon National University

★ Corresponding author

E-mail : hana@inu.ac.kr, Tel : +82-32-835-8760

Manuscript received Nov. 8, 2019; revised Dec. 11, 2019; accepted Dec. 16, 2019.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

CNN 계층 대신, 인셉션 레이어에 기반하여 특징점의 개수를 32개에서 154개로 늘리도록 조정하고, 이를 위해 최적화된 파라미터를 적용하여 효율과 성능이 개선되도록 한다. 그리고, 실험을 통하여 최적의 가중치를 부여하여 이미지 특징을 더 뚜렷하게 나타낼 수 있도록 하였다. 또한, PersonNet의 CND(Cross Neighborhood Difference)[2]가 가지는 모호성을 줄이도록 하여, 종합적으로 성능이 개선되도록 하였고, 그 실험결과를 보이도록 한다.

II. 본론

1. 관련 연구

사람 재인식에 관한 연구는 여러 가지가 있다. [5]는 Siamese Network 구조를 토대로 PersonNet과 유사한 구조로 되어있다. 입력 이미지 쌍이 들어오게 되면 공유 가중치를 가진 CNN 계층에서 특징들을 뽑아낸 후 CIND(Cross-Input Neighborhood Differences)[5]를 거쳐 특징들을 비교한다. 이후 완전연결 계층에 연결시켜 두 이미지 쌍이 동일인인지, 아닌지를 판별하는 구조이다. PersonNet과 다른점은 컨볼루션 필터가 5x5 크기이고 레이어 층도 더 적게 구성된 것이다.

[6]은 2가지의 구조로 이루어진 방법이다. 첫 번째는 기존의 전통적인 컨볼루션, 풀링 계층으로 이루어진 CNN 계층, 두 번째는 동일한 입력 이미지의 다양한 종류의 히스토그램을 통해 특징을 추출하는 ELF16 계층이다. 이러한 두 계층을 통과하여 퓨전 레이어로 합쳐지게 되고 입력 이미지가 어떤 ID 인지 판별하는 완전연결 계층을 통해 결과를 내게 된다. 퓨전 레이어의 입력은 아래와 같다.

$$x = [ELF16, CNNFeatures] \quad (1)$$

활성화 함수 $h()$ 를 이용한 출력은 다음과 같다.

$$Z_{Fusion}(x) = h(W_{Fusion}^T x + b_{Fusion}) \quad (2)$$

[7]은 이미지의 ID를 판단할 때 이미지마다 사람의 자세가 다르고, 이로 인해 오차가 생길 확률이 높으므로 이를 줄이기 위해 사람의 자세를 파악하는 네트워크를 이용하여 사람의 팔, 다리와 같은 부분을 추출하는 전처리를 수행하여 인식하는 모

델이다. Pose를 파악하는 네트워크와 이를 이용해 ID를 판단하는 네트워크 총 2개의 네트워크로 구성되어 있다.

PersonNet[2]은 사람의 이미지들을 입력으로 넣게 되면 두 이미지의 사람이 동일인인지(Same) 다른 사람인지(different) 판단하는 신경망 구조이다. 2개의 입력 이미지를 3개의 CNN 계층을 통과시켜 특징점을 뽑아내는데 컨볼루션 필터의 크기는 3x3 크기로 설정을 한다. 이 컨볼루션 필터를 이용하여 모델의 구조가 7~10개의 가중치 층으로 늘어나게 된다. 이후 CND를 통해, 각 이미지의 특징점들을 비교한다. 다음은 CND 계층의 수식이다.

$$K_i(x,y) = f_A(x,y)I(3,3) - N[h_f(x,y)] \quad (3)$$

여기서 $f_f, h_f \in R^{32 \times 7}$ 을 만족할 때 각 특징 위치에 이웃하는 픽셀값의 차이는 $K_i \in R^{32 \times 7 \times 3 \times 3}$ ($i = 1, \dots, 32$) 이고 각 특징 위치를 중심으로 3x3크기의 윈도우를 가지고 있다. 이는 각각의 3x3블록으로 구성된 32x7크기의 격자모양 이므로 $K_i(x,y) \in R^{32 \times 7}$ ($1 \leq x \leq 7, 1 \leq y \leq 32$) 로 다시 나타낼 수 있다. $I \in R^{3 \times 3}$ 에서 I 는 각 요소가 1로 구성된 3x3 행렬이고, $N[h_f(x,y)]$ 는 h_f 의 x,y 좌표를 기준으로 주변의 3x3크기의 픽셀값이다. 이 방법을 이용하면, 단순히 정확히 위치가 같은 특징점을 비교하는게 아닌 주변의 특징점과 비교를 하므로, 이미지들을 비교할 때 신뢰성 있는 비교가 가능하다. CND 계층을 통과하게 되면 특징 맵의 크기가 3배가 되는데, 다시 크기를 복구시키기 위해 CNN 계층의 스트라이드를 3으로 설정하여 원래 크기로 변환시킨다. 이후 완전연결 계층 3개를 통과 후, 두 이미지 쌍이 동일인인지(same) 아닌지(different) 판단하게 된다.

2. 개선된 PersonNet

본 논문에서는 세 단계에 걸쳐 PersonNet의 인식을 개선한다.

첫 번째는 인셉션 구조의 적용 및 관련 파라미터의 조정이다. 기존 PersonNet의 CNN 계층은 3x3 convolution 필터를 사용해 특징점들을 추출하였다. 이를 이용해 추출한 특징점들의 개수는 32개로 고정되어 있고, 모델의 판별능력에 제한을 줄 수 있다. 이를 해결하고자 단순히 CNN의 층과 출력

특징점을 늘리게 되면 파라미터 수가 늘어나게 되고, 학습시간과 메모리 사용량에 많은 영향을 미치게 된다. 따라서 이 논문에서는 모델의 망 내 연결은 줄이고 세부적인 행렬 연산은 최대한 밀집하게 구성하여 자원의 손실이 적어 이러한 문제점들을 해결할 수 있는 인셉션 계층에 기반을 두어, 여러 크기의 필터를 통해 최종적으로 154개의 특징점을 추출하였다. 그림 1은 본 논문에서 사용한 인셉션 레이어로 특징점을 추출하기 위한 구조를 나타내었다.

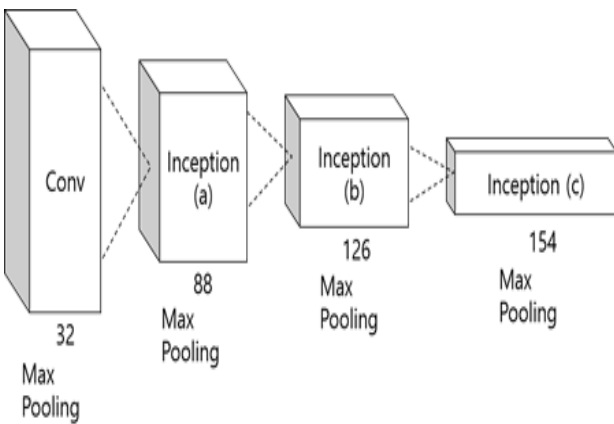


Fig. 1. Structure of Inception Layer.
그림 1. 인셉션 계층의 구조

인셉션 레이어의 파라미터 수를 설정하기 위해 기존 Inception-V1(gooGLeNet)[4]의 인셉션 레이어의 파라미터를 바탕으로 한다. 본 논문은 삼 네트워크 구조를 기반으로 한 대칭적 구조를 가진 모델이므로 파라미터를 동일하게 적용한다면 계산량이 2배가 되므로 GPU메모리 부족 현상을 유발시키게 된다. 또한, Inception-V1의 입력 이미지의 크기는 224x224x3으로 이 논문에서 사용하는 입력 이미지의 크기(160x60x3)와 비교하면 약 5배의 크기 차이가 있다. 이로 인해 Inception-V1의 인셉션 레이어 파라미터의 수를 그대로 적용하게 되면 효율이 낮아지게 된다. 이러한 이유로 인해 각 인셉션 레이어 파라미터의 수를 3~4배 낮추어 구성하였고 이를 통해 GPU 메모리 부족 현상과 학습 속도 저하 현상을 제한하였다. 또한, 인셉션 레이어 중간에 Max Pool 레이어를 추가하였는데 이를 통해 출력값이 너무 방대해져 GPU 메모리가 부족해지는 현상을 막을 수 있다. 본 논문에서 최적화한 파라미터 값은 표 1에 나타내었다.

Table 1. Layer Parameters of Modified Structure.

표 1. 수정된 구조의 레이어 파라미터

Name	Type	Output Dim	FS	Stride
Conv0	C	160x60x32	3x3	1
Pool0	MP	80x30x32	2x2	2
inception(a)	C	80x30x88		
Pool1	MP	40x15x88	2x2	2
inception(b)	C	40x15x126		
Pool2	MP	20x8x126	2x2	2
inception(c)	C	20x8x154		
Pool3	MP	10x4x154	2x2	2
CND	-	10x4x154	3x3	3
Conv4	C	10x4x154	3x3	1
Pool4	MP	5x2x154	2x2	2
FC1	FC	-	4096	-
FC2	FC	-	4096	-
FC3	FC	-	512	-

Table 2. Detailed setup of the Inception Layer parameters.

표 2. 인셉션 레이어 파라미터의 세부 설정

Name	#1x1	reduce / #3x3	reduce / #5x5	#pool
inception(a)	10	46 / 70	2 / 4	4
inception(b)	20	60 / 90	4 / 8	8
inception(c)	30	66 / 100	6 / 12	12

표 2는 각 인셉션 레이어를 구성하는 필터의 개수를 나타낸다. #NxN은 크기가 N인 컨볼루션 필터의 출력 채널의 개수, reduce는 #NxN 컨볼루션 필터를 적용하기 전 크기가 1인 컨볼루션 레이어로 감소시킬 채널의 개수, #pool은 풀링 과정을 적용한 후 크기가 1인 컨볼루션 레이어를 적용하여 출력할 채널의 개수이다.

두 번째는 특징점 분할 및 가중치 추가를 통한 특징점의 확대이다.

그림 2(논문의 끝에 위치)와 같이 사람을 이미지에서 인식할 때 위치에 따른 정보량은 차이가 있으며, 그 목적에 따른 위치 정보량의 차이도 클 수 있다. 따라서 본 논문에서는 본 논문에 맞는 최적의 정보량을 실험을 통해 구하고, 이것을 전체 모델에 적용하여 최적의 인식률이 구해질 수 있도록 하였



Fig. 2. Market-1501 dataset image.
그림 2. Market-1501 데이터 세트 이미지

다. 이를 위해 CND 방법 전에 뽑아낸 특징점들을 3등분 하고, 특징 위치의 특징점에 가중치를 더 부여하여 비교하는 방법을 사용하였다.

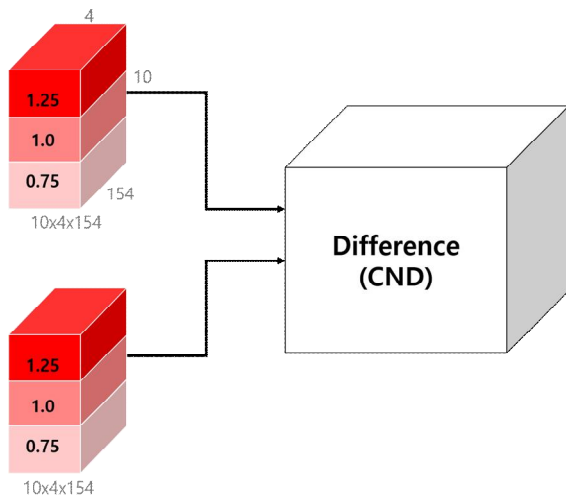


Fig. 3. An example of High weight on top area.
그림 3. 상단 부분에 높은 가중치의 적용한 CND 입력

예를들어, 그림 3은 CND를 통과하기 전 3등분한 입력들의 상단 부분에 더 높은 가중치를 부여한 방법을 나타낸다. 이와 같은 방식으로 각 특징점을 위치에 따라 나누고 여기에서 얻어지는 정보량의 비대칭성에 의해 생기는 인식률을 실험하였다. 본 논문에서는 여러 개로 분할한 특징점을 비교했으며 또 각 위치의 가중치도 변경하여 실험하였다. 실험방법은 재인식 분야에서 평가를 위해 쓰이는 Rank 평가방법을 이용했으며, 각 데이터 세트에서 100명의 ID를 뽑아 평가하였다. 평가방법에 관한 자세한 기술은 ‘3.실험 결과’ 부분에 기술하였으며, 여기에서는 가중치 분류에 따른 결과판으로 최적의 가중치 부분을 결정한다.

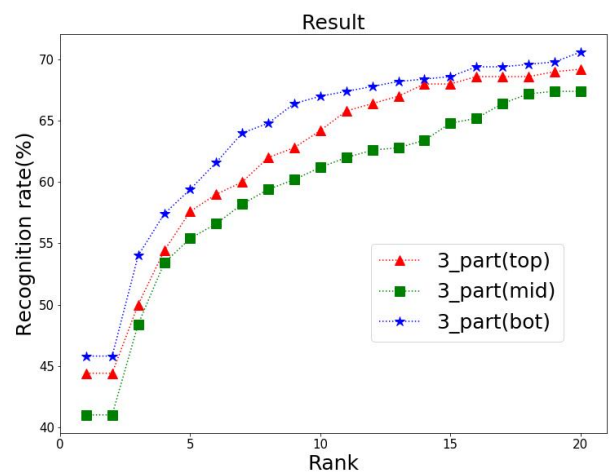


Fig. 4. Results according to the maximum weighting position of the three-division method.
그림 4. 3등분 분할 방법의 최대 가중치 위치에 따른 결과

Table 3. Weights setting for the three-division method.
표 3. 3등분 분할 방법의 가중치 설정

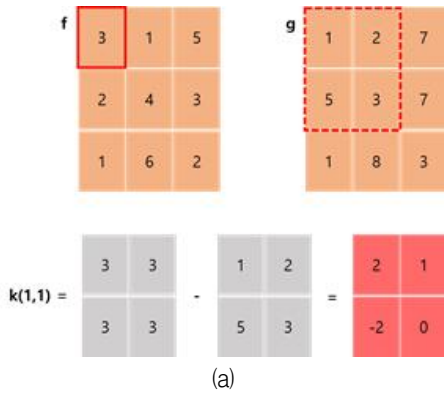
	Top	Middle	Bottom
top	1.25	1.0	0.75
mid	1.0	1.25	0.75
bot	0.75	1.0	1.25

먼저 그림 4는, 특징점을 3부분으로 나누고 각각 다른 부분에 강한 가중치를 적용하여 실험한 결과이다. 상단 가중치를 강하게 설정한 ‘top’, 중단 가중치를 강하게 설정한 ‘mid’ 마지막으로 하단 가중치를 강하게 설정한 ‘bot’으로 구분하여 실험하였다. 이때 설정한 가중치 값은 표 3에서 확인할 수 있다. 이 실험에 사용된 데이터 세트는 CUHK01[8], CUHK03[9], Market-1501[10], PKU-Reid[11], 3DPes[12] 총 5종류로 CUHK03 데이터 세트로 학습하였고 그 외의 데이터 세트로 테스트하였다. 결

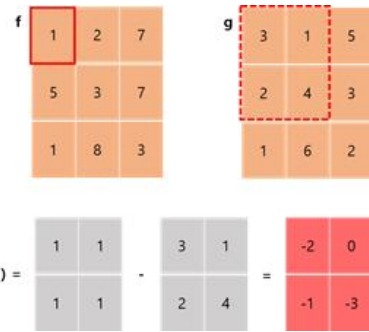
과값은 4가지 데이터 세트 결과의 평균값이다. 그림 4에 따르면 상단에 큰 가중치를 적용한 방법보다 하단에 큰 가중치를 적용한 방법이 모든 Rank에서 다른 방법에 비해 높은 인식률을 보인다.

그림 5의 경우(논문의 끝에 위치), 3등분 한 방법을 적용하여 특징점의 상위 부분에 강한 가중치를 주고 비교한 결과이다. 그림 5와 같이, 일부 입력 데이터의 경우, 하단보다, 상단에 더 많은 특징이 분포하는 경우도 있다. 그러나, 그림 4에서 보이듯이, 대부분의 데이터 실험에서 하단에 가중치를 주는 것이 평균적으로 더 향상된 성능을 보이는 것을 알 수 있었다.

세 번째는 CND 방식이 가지는 모호성의 감쇄이다



(a)



(b)

Fig. 6. Difference layer output of the same input.

그림 6. 동일 입력값의 Difference 계층 출력

기존의 CND 방식은 입력1 이미지에서 입력2 이미지의 특징점의 차이를 특정 방법으로 계산하는 방식이다. 여기서 차이를 구할 때 각 특징점의 값들을 뺄셈 연산을 통해 구하게 되는데 입력1과 입력2의 이미지 입력 순서가 바뀌게 되면 결과가 달라질 수 있게 된다. 예를 들어 그림 6에서 (a)와 (b)의 입력값은 동일하지만 입력의 순서가 다르므로 다른 결과가 출력되는 것을 알 수 있다. 이는 모델의 학습 시 모호한 기준으로 반영될 수 있고 이로 인해 신뢰성 없는 결과가 도출될 수 있다. 따라서 이 논문에서는 뺄셈 연산 이후에 절댓값을 취하여 입력이 서로 바뀌게 되더라도 상이한 결과값이 나오게 되는 것을 어느 정도 방지하였다.



Fig. 5. Results of the three-division(Top) method.

그림 5. 3등분 방법의 결과(Top)

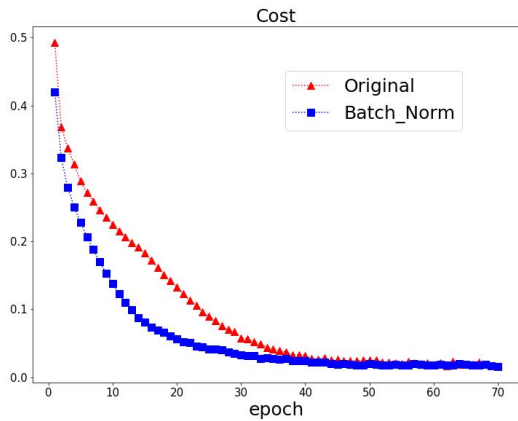


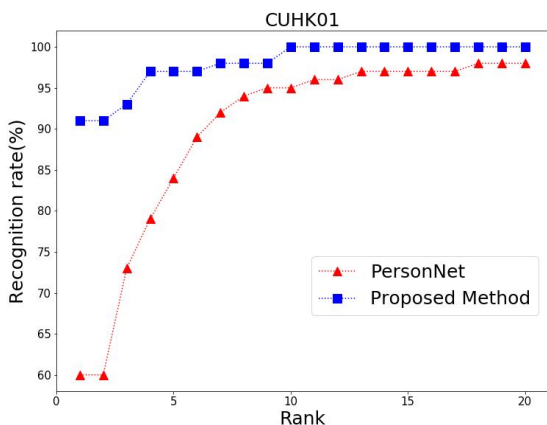
Fig. 7. Error function convergence graph after applying Batch Normalization.

그림 7. Batch Normalization 적용 후 오차 함수 수렴 그래프

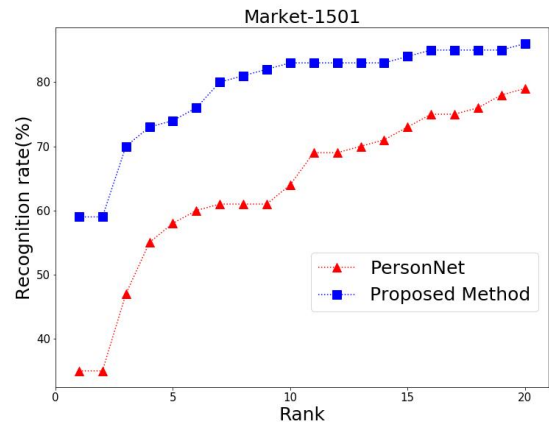
또한, Batch Normalization[13] 방식을 적용하여 학습 과정에서 오차 함수가 빠르게 수렴할 수 있도록 하였다.

3. 실험 결과

본 논문에서는, CUHK01, CUHK03 그리고 Market-1501 3가지 데이터베이스를 이용하였고 학습은 CUHK03 데이터, 테스트는 CUHK01, Market-1501 2가지 데이터베이스를 이용해 수행하였다. 학습에 사용한 CUHK03 데이터는 1360명의 보행자의 13,164장의 이미지로 구성되어 있다. 학습 데이터는 한 사람당 10쌍의 긍정 이미지와 50쌍의 부정 이미지로 구성하였다. 테스트 시 데이터베이스마다 100명의 사람을 뽑아 그 사람들의 전체 이미지에서 기준 이미지와의 유사성을 계산하여 1위부터 n 위까지의 유사성 Rank를 측정하여 판단하였다.



(a) CUHK01 Dataset



(b) Market-1501 Dataset

Fig. 8. Performance comparison of PersonNet with the proposed method.

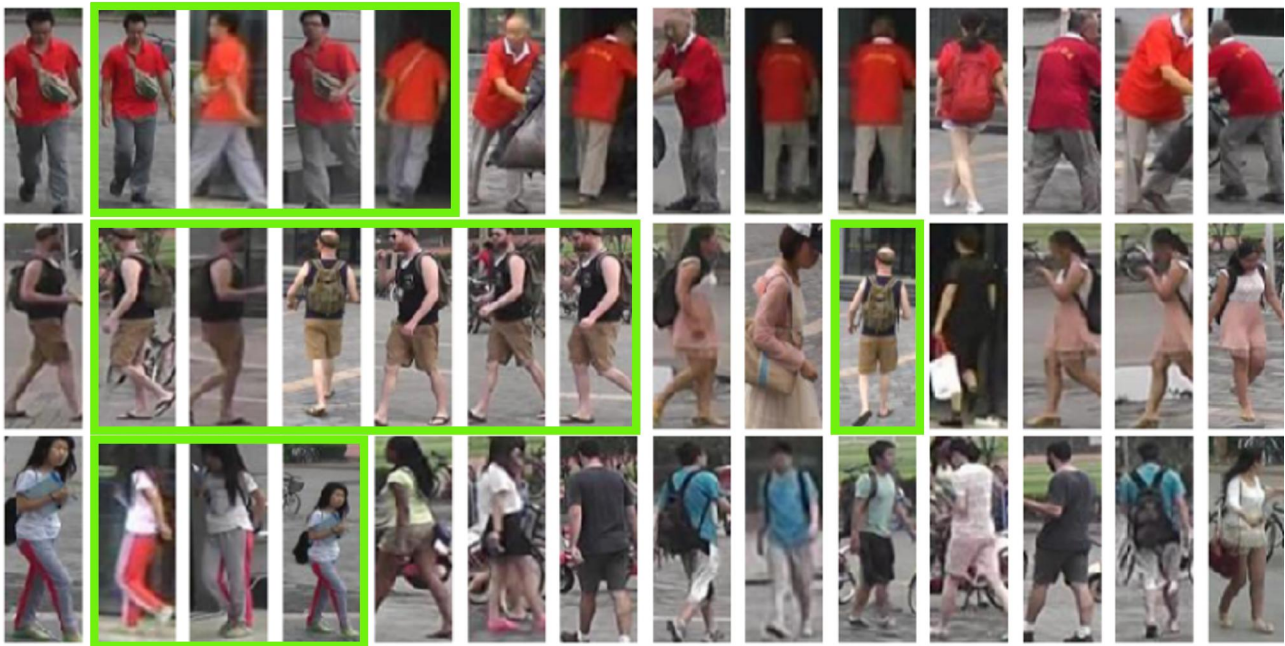
그림 8. PersonNet과 제안하는 방법의 성능 비교

그림 8은 본 논문의 실험결과를 보인다. CUHK01 데이터 세트에서는 상위 랭크(Rank-1)에서의 인식률이 90% 이상으로 다른 방법들에 비해 50%가량 좋아진 것을 관찰할 수 있다. 또한 Market-1501 데이터 세트에서도 상위 랭크에서의 인식률이 24% 정도 향상된 것을 알 수 있다.

그림 9는(논문의 끝에 위치) PersonNet과 본 논문의 방법의 Rank 결과 사진들이다. 첫 번째 사진들은 기준이 되는 사진들이고, 이 사진들과 가까이 있을수록 Rank가 높다는 것을 의미한다. 초록색 상자로 표시된 사진은 기준 그림에 대한 긍정 응답이다. (a)보다 (b)방법의 인식률이 더 높은 것을 볼 수 있고, (a)의 첫 번째 사진의 경우에는 단순히 빨간색 옷이라는 특징만 파악하고 있지만 (b)의 경우 빨간 옷뿐만 아니라 가방의 특징을 파악하여 잘 판단해주는 모습을 볼 수 있다. (a)의 두 번째 사진의 Rank-1, 2, 3의 이미지는 사람이 판단하기에도 확실히 다른 사람이다. 하지만 상위 Rank로 판단이 되어있다. 이와 반대로 (b)는 상위 Rank의 결과가 정확한 것을 볼 수 있다. (a)의 세 번째 사진의 경우 동일인을 찾지 못하였지만 (b)의 경우 동일인을 상위 Rank로 구분하였다. 본 논문의 전체 성능은 표 4에 나타내었다.



(a) PersonNet



(b) proposed method

Fig. 9. PersonNet and the proposed method results,
그림 9. PersonNet과 제안하는 방법의 결과

Table 4. Rank-1, Rank-5, Rank-10, Rank-20 recognition rate of each method.

표 4. 각 방법의 Rank-1, Rank-5, Rank-10, Rank-20 인식률

(a) CUHK01 Dataset

Method	r=1	r=5	r=10	r=20
PersonNet	60	89	96	98
Proposed Method	91	97	100	100

(b) Market-1501 Dataset

Method	r=1	r=5	r=10	r=20
PersonNet	35	60	69	79
Proposed Method	59	76	83	86

이 실험의 결과를 통해 기존의 방법보다 27~31% 개선되었음을 보여준다.

III 결론

본 논문에서는 PersonNet의 CNN계층이 추출하는 제한된 특징점의 개수로 인해 발생하는 이미지 판단의 불확실성을 해결하기 위해, 인셉션 계층을 적용하였고 기존 32개의 특징점보다 더 증가된 154개의 특징점을 추출하였다. CNN의 출력 특징점과 층의 개수를 증가시키는 방법은 학습 파라미터 수의 증가, 교육시간과 메모리 사용량의 증가를 발생시키므로, 이로 인해 느려지는 모델의 학습 속도와 방대해진 출력 데이터의 크기로 인한 제한사항의 개선을 위해, 파라미터의 개수 설정을 최적화하고 Max Pool 레이어를 적용하여 해결하였고, 두 이미지의 특징점을 CND로 비교하기 전 가중치를 조정하여 특징값을 극대화하였다. 또한, 기존 CND 방법을 수정하여 최종적으로 27~31%의 성능 개선을 보였다. 기존의 PersonNet이 가진 특징점 분석 능력은 한계가 있지만 인셉션 레이어를 적용하여 더 많은 특징점을 뽑아내어 비교하였을 때 특징이 뚜렷한 옷의 색깔 정보뿐만 아니라 더 고차원적인 특징을 파악한다는 것을 결과 사진들을 통해 알 수 있었고, 이러한 방법으로 기존 PersonNet보다 성능을 향상할 수 있었다. 이미지의 특징점을 3등분으로 나누어 가중치를 적용하여 이미지의 특징이 하단에 더 많이 분포한 경우 인식률이 높게 나오게 하였다. 본 논문의 실험에서는, 하단 부분에 특징점이 많고, 사람 재인식의 경우 하단에 더 많은 특징이 있다는 것을 확인하였지만, 가중치를 데이터, 또는 데이터 환경에 따라 적응적으로 설정하게 적용한다면 인식률을 더 높일 수 있을 것이다. 향후 연구 방향으로는 단순히 2개의 이미지를 비교하는 것이 아닌 각 이미지의 비슷한 영역을 추출하여 비교하는 방식을 적용하여 더욱 신뢰성 있는 Re-ID 방식을 구축할 수 있을 것이다.

References

[1] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," A thesis for the degree of Master of Science, University of Toronto, 2015.
 [2] L. Wu, C. Shen, and A. v. d. Hengel,

"Personnet: Person reidentification with deep convolutional neural networks.," *arXiv preprint arXiv : 1601.07255*, [Online]. Available: <https://arxiv.org/abs/1601.07255>, 2016
 [3] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", *arXiv preprint arXiv:1409.1556*, [Online]. Available: <https://arxiv.org/abs/1409.1556>, 2014.
 [4] C. Szegedy, W. Liu, Y. Jia, P. sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.1-9, 2015.
 [5] E. Ahmed, M. Jones, and T. K. Marks, "An Improved Deep Learning Architecture for Person Re-Identification," *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp.3908-3916, 2015.
 DOI: 10.1109/CVPR.2015.7299016
 [6] S. Wu, Y. C. Chen, X. Li, A. C. Wu, J. J. You and W. S. Zheng, "An Enhanced Deep Feature Representation for Person Re-identification," *2016 IEEE winter conference on applications of computer vision (WACV)*. IEEE, pp.1-8, 2016.
 [7] L. Zheng, Y. Huang, H. Lu and Y. Yang, "Pose Invariant Embedding for Deep Person Re-Identification," *IEEE Transactions on Image Processing*, pp.4500-4509, 2019.
 DOI: 10.1109/TIP.2019.2910414
 [8] W. Li, R. Zhao, and X. Wang, "Human reidentification with transferred metric learning," *Asian conference on computer vision. Springer, Berlin, Heidelberg*, pp.31-44, 2012.
 DOI: 10.1007/978-3-642-37331-2_3
 [9] W. Li, R. Zaho, T. Xiao, and X. Wang, "Deepreid: Deep filter pairing neural network for person re-identification," *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp.152-159, 2014. DOI: 10.1109/CVPR.2014.27
 [10] L. Zheng, L. Shen, L. Tian, S. wang, J. Wang, and Q. Tian, "Scalable person re-identification: A benchmark," *Proceedings of the IEEE international*

conference on computer vision. pp.1116–1124, 2015. DOI: 10.1109/ICCV.2015.133

[11] L. Ma, H. Liu, L. Hu, C. Wang, Q. Sun, “Orientation driven bag of appearances for person re-identification,” *arXiv preprint arXiv:1605.02464*, [Online]. Available: <https://arxiv.org/abs/1605.02464>, 2016.

[12] D. Baltieri, R. Vezzani, R. Cucchiara, “3DPes: 3D People Dataset for Surveillance and Forensics,” *Proceedings of the 2011 joint ACM workshop on Human gesture and behavior understanding*. ACM, pp.59–64, 2011. DOI: 10.1145/2072572.2072590

[13] S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” *arXiv preprint arXiv:1502.03167*, [Online]. Available: <https://arxiv.org/abs/1502.03167>, 2015.

BIOGRAPHY

Seong-Hyeon Park (Member)



2014~ : BS degree in Embedded Systems Engineering, Incheon National University.
A.I. Deep Learning, Embedded Systems Engineering

Seok-Hoon Kang (Member)



1989 : BS degree in Electronic Communications Engineering, Hanyang University.
1995 : PhD degree in Electronic Communications Engineering, Hanyang University.

2004~ : Professor, Embedded Systems Engineering, Incheon National University.
A.I., Deep Learning, Mobile/Embedded System, Wearable System, Natural Language Processing