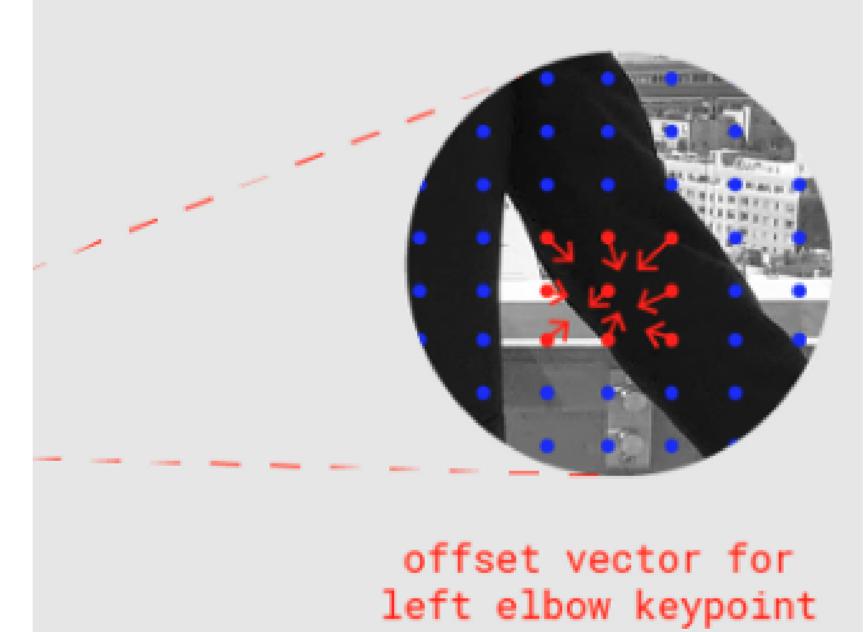


각각의 히트맵은 해상도 * 해상도 * 24+1의 3D 텐서

24는 디텍하고자 하는 키포인트의 개수

+1 는 객체의 중앙 Keypoints

히트맵의 각각의 포지션은 신뢰 점수를 가지고, 이것은 해당 히트맵 포인트에서 벡터를 따라 이동할 때 키포인트의 정확한 위치를 예측하는데 사



offset vector는 해상도 해상도 48의 3D Tensor이다. 48는 키포인트 * 2(x,y)를 의미한다.

히트맵은 키포인트가 어디 위치하는지를 나타내는 근사치이고 offset vectors는 히트맵 포인트에 해당하며 해당 히트 맵 포인트에서 벡터를 따라 이동할 때 키포인트의 정확한 위치를 예측하는데 사용된다.

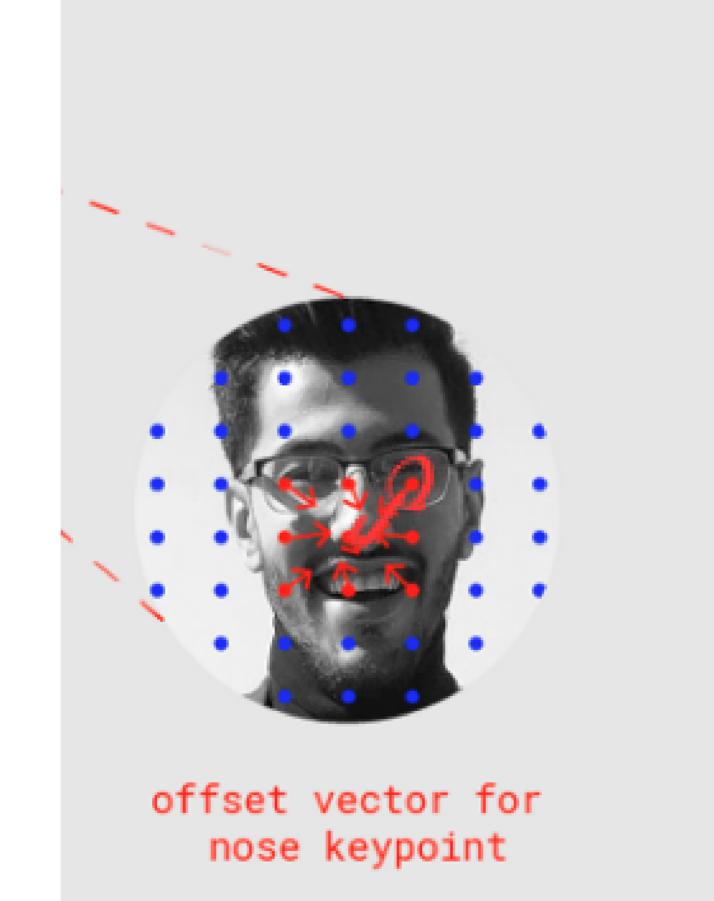
Offset, $F_k(x_i) = l_k - x_i$

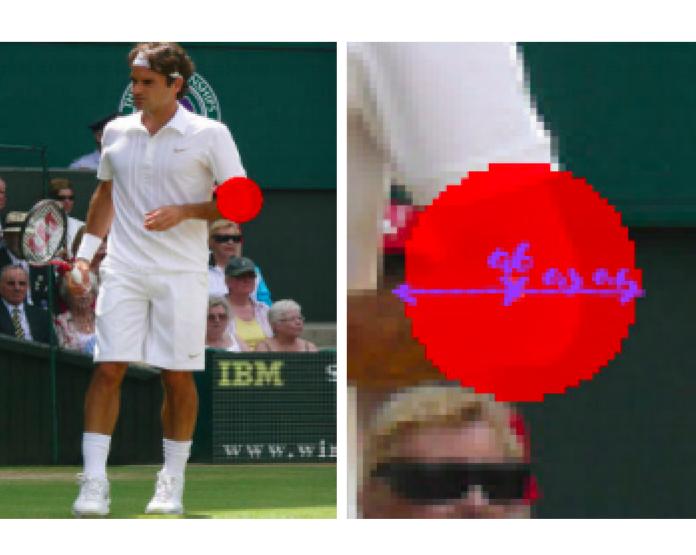
 x_i 는 픽셀의 location 을 의미한다.

 $_k$ 는 임의의 key point (여기서는 17개중 하나)를 의미한다.

 l_k 는 key point (ground-truth로 이해) 위치

즉, offset이란 keypoint로부터 거리가 얼마나 떨어져 있냐이고, 그림 07은 이것을 시각화 한 것





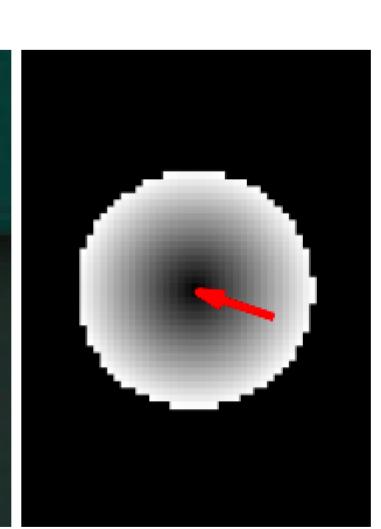


Figure 2: Network target outputs. Left & Middle: Heatmap target for the left-elbow keypoint (red indicates heatmap of 1). Right: Offset field L2 magnitude (shown in grayscale) and 2-D offset vector shown in red).

- * Pose 추출 순서
- 1. 추출한 Center Heatmap에서 Maxpooling 연산을 진행(이때 Heatmap의 사이즈는 불변)
- 2. Maxpooling 된 Center Heatmap에 존재하는 Confidence score 중 미리 설정한 top k 개의 x,y 좌표 추출
- --> 이 과정에서 한장의 이미지에서 몇개의 object가 어느 영역(x,y) 에 있는지 근사화(Coarse Inference)
- 3. Offset Vector 는 위에서 구한 heatmap의 coarse하게 추출한 x,y 좌표에서의 offset vector의 값을 추출 --> offset vector : [offsets(y,x,k) --> x vector , offsets(y,x,k+1) --> y vector]

