

객체 분리 및 인코딩을 이용한 애완동물 영상 세부 분류 인식

김지혜

LIG넥스원

Fine grained recognition of breed of animal from image using object segmentation and image encoding

Ji-hae Kim

LIG Nex1

E-mail : jihae.kim@lignex1.com

요 약

본 논문은 개와 고양이에 해당하는 애완동물 영상에서 세부 분류인 동물의 종을 인식하는 것을 목표로 한다. 영상의 세부 분류 인식에 대한 연구는 계속적으로 발전하고 있지만, 다형성의 성질을 갖는 동물에 대한 객체인식 연구는 더디게 진행되고 있다. 본 논문에서는 객체 분리를 위해 Grab-cut 알고리즘을 이용하고, 영상 인코딩을 위해 Fisher Vector를 이용하여 더 높은 동물 객체인식 방법을 제안한다.

ABSTRACT

A goal of this paper is doing fine grained recognition of breed of animal from pet images. Research about fine grained recognition from images is continuously developing, but it is not for animal object recognition because they have polymorphism. This paper proposes method of higher animal object recognition using Grab-cut algorithm for object segmentation and Fisher Vector for image encoding.

키워드

fine grained recognition, object recognition, object segmentation, image encoding

I. 서 론

본 논문은 개와 고양이에 해당하는 애완동물 영상에서 세부 분류인 동물의 종을 인식하는 것을 목표로 한다. 영상의 세부 분류 인식에 대한 연구는 국제적으로도 많은 연구가 진행되고 있지만 [1,2], 다형성의 성질을 갖는 동물에 대한 객체인식 연구는 더디게 진행되고 있다. 본 논문에서는 세부 분류 인식 단계 중 객체 분리와 영상 인코딩 단계에서 개선할 수 있는 방법을 설명한다. 객체 분리를 위해 Grab-cut 알고리즘[3]을 이용하고, 영상 인코딩을 위해 Fisher Vector[4]를 이용하여 더 높은 동물 객체인식 방법을 제안한다.

본 논문은 다음과 같이 구성되어있다. 먼저 2장에서는 본 논문에서 사용한 객체 분리 방법을 소개하고, 3장에서는 인코딩 방법을 소개한다. 4장에

서는 제안하는 방법을 이용한 실험결과 및 분석에 대하여 설명하며 마지막으로 5장에서는 결론을 내린다.

II. 객체 분리

객체 분리란 어떤 장면 혹은 패턴에서 주의와 시각의 대상이 되는 중요한 대상에 해당하는 객체와 그 나머지인 배경으로 분리되는 것이다. 효과적인 객체 인식을 위해 객체 분리 단계는 중요한 단계이다. 본 논문은 객체 분리를 위해 Grab-cut 알고리즘을 이용하였다.

Grab-cut 알고리즘[3]은 Graph cut을 기반으로 하는 영상 분할 방법으로, 객체 주위의 사용자 지정 경계 박스에서 전경 객체에 해당할 가능성이 있는

화소를 마크하여 결과 영상을 생성한다. 주로 정지 영상에서 객체를 추출할 때 적합하다.

III. 인코딩

영상의 세부 분류 인식을 위해서는 학습 영상으로부터 추출한 다량의 특징점을 사용된다. 추출한 모든 특징점을 인식을 위해 사용한다면 엄청난 복잡도에 빠지게 될 것이다. 이를 해결하기 위해서는 특징점의 변별력은 유지하면서, 특징점들의 차원은 줄이는 방법을 사용하게 되는데 이 과정이 영상 인코딩에 해당된다. 본 논문에서는 효과적인 인코딩을 위해 Fisher-Vector[4]를 사용하였다.

Fisher-Vector(FV)는 Fisher-kernel을 기반으로 하는 영상 인코딩 방법으로 학습영상으로부터 GMM(Gaussian Mixture Model)을 생성하고 입력된 특징벡터를 가우시안 모델에 적용하는 방법이다. 이 방법은 근사화 오류를 최소화한다는 장점이 있다.

IV. 실험 결과 및 분석

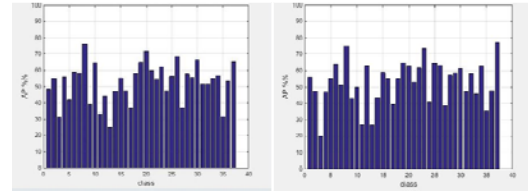
본 논문에서는 제안하는 방법을 검증하기 위해 37개의 종에 해당하는 총 7393장의 개와 고양이 입력영상에 대하여 실험을 진행하였다.

먼저 영상 인코딩 방법에 따른 정확도 차이를 알기위한 실험을 진행하였다. 정확도 평가를 위해 기본적인 영상분류 방법인 Bag of Words(BoW)[5] 방법을 사용하였을 때의 정확도 값과 제안하는 방법을 이용하였을 때의 정확도 값을 비교하였다. 표 1에서는 Fisher-Vector(FV) 사용에 따른 실험결과를 보여주고 있다. BoW 방법만 사용하였을 때는 35.46%, FV를 함께 사용하였을 때는 51.24%의 정확도로, 15.78% 상승된 정확도를 확인할 수 있다.

다음으로는 객체 분리 방법에 따른 정확도 차이를 알기위한 실험을 진행하였다. 그림 1에서는 본 논문에서 제안하는 방법인 Grab-cut 알고리즘을 적용여부에 따른 정확도를 그래프로 보여주고 있다. (a)와 같이 Grab-cut 알고리즘을 적용하지 않았을 때는 51.24%의 정확도를, (b)와 같이 적용하였을 때는 53.98%의 정확도를 확인할 수 있다. 제안하는 방법 적용을 통해 2.74% 상승된 정확도를 확인할 수 있다.

표 1. Fisher-Vector(FV) 사용에 따른 실험결과

	정확도
BoW	35.46 %
BoW + FV	51.24 %



(a) w/o Grab-cut (b) w Grab-cut

그림 1. Grab-cut 알고리즘 적용에 따른 실험결과

V. 결 론

본 논문에서는 개와 고양이에 해당하는 애완동물 입력영상에서 동물의 세부 분류에 해당하는 종을 인식하는 방법을 제안하고 실험을 통하여 그 효과를 검증하였다. 더 높은 객체 인식률을 위해 객체 분리와 영상 인코딩 단계에서의 개선방법을 제안한다. 추후에는 입력영상 학습 및 인식을 위해 딥러닝 방법을 사용하여 정확도에 상승에 기여하는 추가적인 연구를 진행하고자 한다.

References

- [1] Lingxi Xie, Jingdong Wang, Bo Zhang, and Qi Tian, "Fine-Grained Image Search," *IEEE Transactions on multimedia*, Vol. 17, No. 5, pp. 636-647, May. 2015.
- [2] C. Spampinato, S. Palazzo, P.H. Joalland, "Fine-Grained object recognition in underwater visual data," in *Proceeding of the Springer Science+Business Media*, New York: NY, pp. 1-20, 2015.
- [3] C Rother, V Kolmogorov, A Blacke, "Grabcut: Interactive foreground extraction using iterated graph cuts" *ACM transactions on Graphics(TOG)*, Vol. 23, No. 3, pp. 309-314, August. 2004.
- [4] C. Dance and F. Perronnin, "Fisher Kernels on Visual Vocabularies for Image Categorization", in *proceeding of the IEEE Conference on Computer Vision*, Minneapolis, MN, pp. 1-8, 2007.
- [5] Gabriella Csurka, Christopher R.Dance, Lixin Fan, Jutta Willamowski, Cedric Bray, "Visual categorization with bags of keypoints", in *proceeding of the workshop on statistical learning in Computer Vision, ECCV*, pp. 1-16, 2004.