

Segmentation Propagation in ImageNet

H.Tuğrul Erdoğan

Neden Bu Çalışma?

- Segmentasyon; ön/arkaplan ayırıştırma, poz tahmini için ön işleç
- Çok geniş veri kümesi üzerinde çalışmak: elle sınıflandırmak
 - arama-silme zor sınıf sayısının sınırlı elle sınırlı sayıda sınıf

Yönteme genel bakış

- Yankılama ile test görüntüsüne en benzer görüntüleri içeren alt küme ile çalışılabilir
- T0 anında temel doğru (ground truth) değerler kullanılıyor.
- T1 anında, artık sadece ana nesneye göre etiketlenmiş olan görüntüler
 - Alt kümede, T0'daki etiketlerle aynı etiketteki görüntüler.

Yönteme genel bakış

- Sonraki Tx anlarında sadece daha önceden bulunan sınıflar ve yeni sınıflarla bağlantılı eski sınıflar alt kümesinde çalışma
- En nihayetinde hep TD çevresinde yankılanan alt kümeler ile çalışma
 - İyi yanı; gittikçe genişliyor ve hep en bağlantılı görüntüler var
 - Kötü yanı; yapılan hata da etkisini aktarıyor

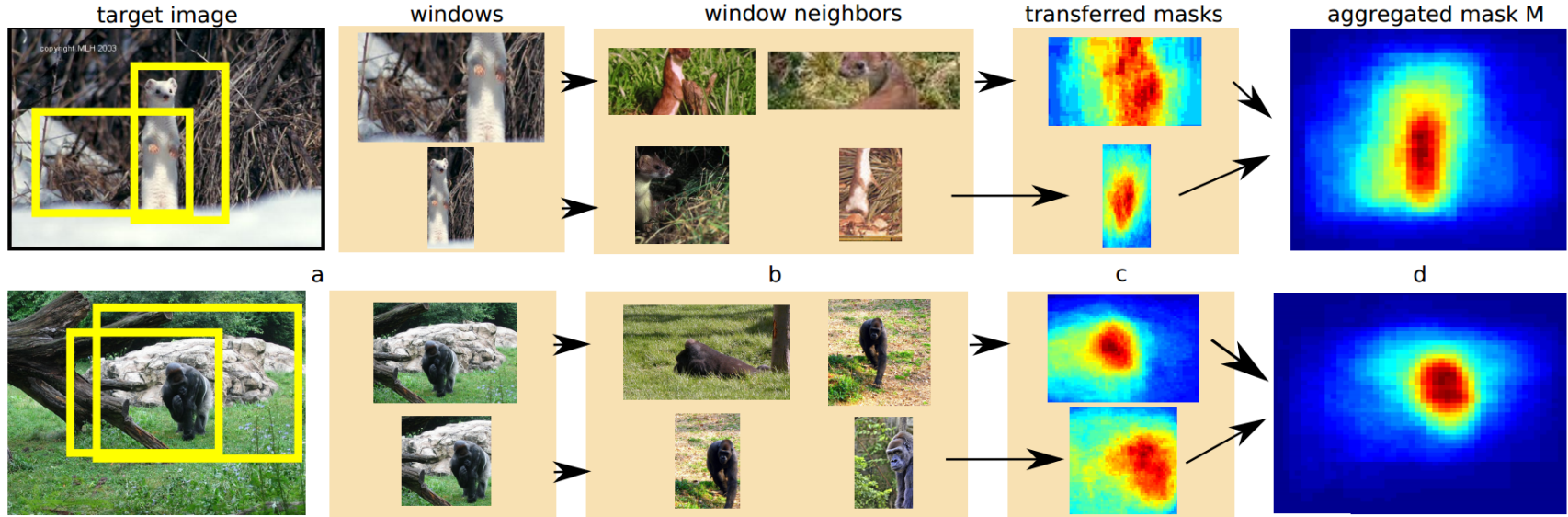
Detaylı tasarım: Pencere yankılanması

- Her test görüntüsü ile, o anki örnekler kümesinden global önitelikler bakımından en benzerler alt-kümesi alınıyor
- Benzer tanımında “nesneli (objectness)” parçlar kullanılıyor
 - Böylece bir nesne varsa ortalananmış oluyor
 - hem de bütün görüntüye göre o nesneyi daha karakterize edebiliyor

Detaylı tasarım: Pencere yankılanması

- Gelen görüntülerden bir maske elde et
- maske, arkaplan mı yada önplan bilgisi taşıyor olma olasılığını içeriyor $[0,1]$
 - Böylece önceki deneyimlerden elde edilmiş sonuçlar yankılanabiliyor
 - Güncel bölütlemenin, geçmişteki sonuçlar ile bütünvari olması sağlanıyor

Detaylı tasarım: Pencere yankılanması



Detaylı tasarım: Segmentasyon yankılanması

- Yöntemin başarımı alt küme çıkarımı sırasında olabildiğince test görüntüsüne yakın görüntüler ile çalışabilinmesinde
- Gist yerine HOG. 100k görüntü için 10^6 uzaklık hesabı var. Sonuçlar saklansa 310Gb
 - Bunun için “sıralı niceleme (terative quantization)” ile Hamming uzaklığı ile aynı değere sahip yüksek boyutlu kısa vektörler ile L2 normu kullanıldı.

Detaylı tasarım: Maskelerin oluşması

- En yakın k tane görüntüden ağırlıklı toplam ile olasılık haritası çıkarılıyor
 - Daha akıllı olarak, her görüntü için farklı katsayı hesabı

$$\min_{\{\lambda_k\}} \sum_w \left\| m_w - \sum_{k=1}^K \lambda_k m_{s_k} \right\|_F^2 \quad \text{s.t.} \quad \forall k, \lambda_k \geq 0, \quad \text{and} \quad \sum_{k=1}^K \lambda_k = 1. \quad (1)$$

$$\lambda_k = \exp(\hat{\lambda}_k) / \sum_{k=1}^K \exp(\hat{\lambda}_k). \quad (2)$$

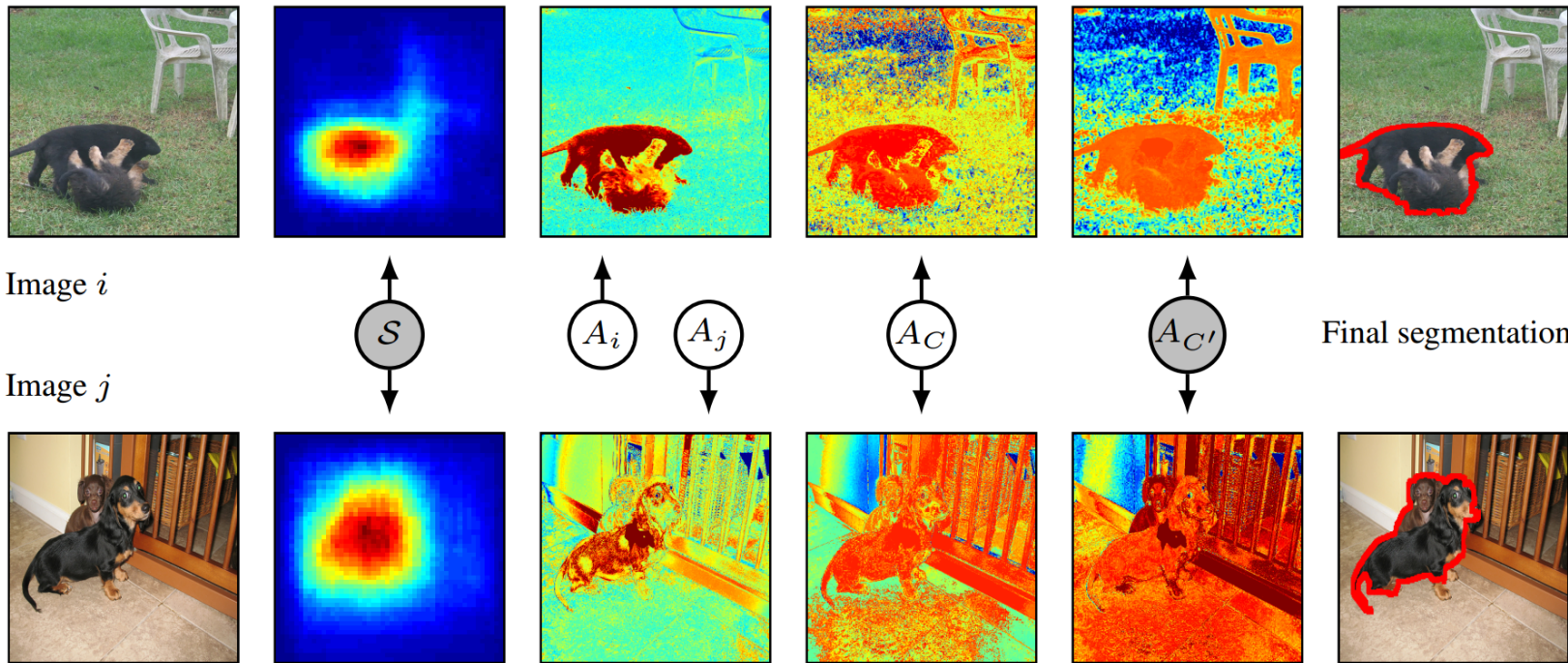
Detaylı tasarım: Bölütleme

- Enerji minimizasyon problemi

$$E(\mathbf{x}; \mathcal{A}, \mathcal{S}) = \sum_i \left(\sum_p E_{ip}(x_{ip}; \mathcal{A}, \mathcal{S}) + \sum_{p,q} E_{ipq}(x_{ip}, x_{iq}) \right) \quad (3)$$

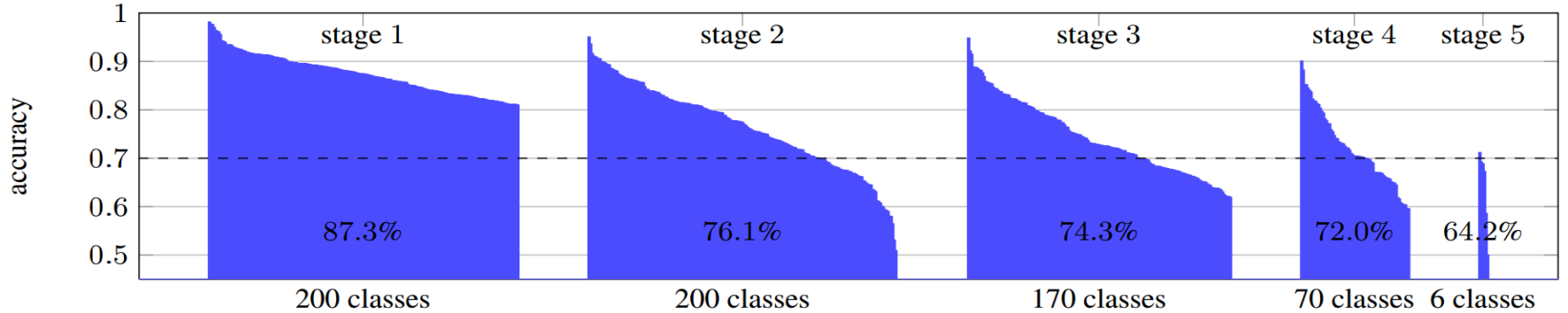
- Benzer görüntülerin sınıflarının nesne modellerinin de eklenmesi
 - Benzer çalışmalardan farklı nokta tamamen otomatik

Detaylı tasarım: Bölütleme



Sonuçlar

	[16]	[15]	image only \approx GrabCut [1]	class only \approx Batra [36]	image+class	image+transfer \approx Kuettel [2]	image+transfer +class
Accuracy	78.9	85.4	82.4	83.6	88.2	87.6	91.4



Stage 2-5 tam segmentasyon yankılama yerine hep T0 kümesi kullanımını tartışmaya açıyor.

Teşekkürler