

# Marka Erişimi İçin Derin Özniteliklerin İncelenmesi

## Analyzing Deep Features for Trademark Retrieval

Cemal Aker, Osman Tursun, Sinan Kalkan

Kovan Araştırma Laboratuvarı, Bilgisayar Mühendisliği Bölümü, Orta Doğu Teknik Üniversitesi, Ankara, Türkiye

Email: {cemal, wusiman.tuerxun, skalkan}@ceng.metu.edu.tr

**Özetçe** —Marka erişimi (ME), marka başvurularında ve ihlallerinde artan eğilimlerden dolayı, önemli ve zorlayıcı bir problem haline gelmiştir. Şimdiye kadarki çalışmalar elle tasarlanmış öznitelik tabanlı olup, yeterli başarı gösterememişlerdir. Derin öğrenme yöntemlerine olan artan eğilimi göz önünde bulundurarak, bu çalışmada birkaç yaygın Evrişimsel Sinir Ağı (ESA) modeli ME problemine uygulanmıştır. Modeller, daha önce önerilen çözümlerin aksine, geniş çaplı bir veri kümesi kullanılarak değerlendirilmiş ve ESA'ların yetersiz olduğu noktalar tartışılmıştır. Karşılaşılabilecek problemler için özel ince ayar, uzaklık ölçütü öğrenme, özniteliklerin yerleştirilmesi ve görüntü boyutundan bağımsızlaştırılması gibi çözümler önerilmiştir.

**Anahtar Kelimeler**—marka erişimi, derin öğrenme.

**Abstract**—The rapid rise in the amount of trademark applications and trademark infringements has led the trademark retrieval (TR) to become an important and formidable task to solve. Existing studies based on hand-crafted features show unsatisfying performance. Taking the popularization and increasing success of the deep learning methods into consideration, in this work, many well-known Convolutional Neural Network (CNN) models are applied to the TR problem. Models are tested using a large scale trademark dataset in contrast with the previously proposed solutions, and their failure points are discussed in this study. For the problems that can be encountered, solutions such as fine-tuning, distance metric learning, using CNN features locally, and making them invariant to aspect ratio of the trademark are suggested.

**Keywords**—trademark retrieval, deep learning.

### I. GİRİŞ

Marka, bir şahsa, kuruluşa veya tüzel kişiliğe ait bir ürün veya servisi tanımlayıp diğerlerinden ayıran işaret veya ilişkili yazıdır. Kayıtlı markalar yasal mal olduğundan, korsanlık ve marka ihlaline karşı korunmalıdır. Bu yüzden, patent ofislerinde yeni marka kaydı yapılırken, daha önce kayıtlı olanlarla görsel ve anlamsal benzerlik taşıyıp taşımadığı araştırılmalıdır. Bu araştırma, bir insanın, ihlal ihtimali bulunan markaları gözle karşılaştırması ile veya marka erişimi (ME - ing. Trademark Retrieval) sistemleri kullanılarak yapılmaktadır.

İçerik Tabanlı Görüntü Erişimi (İTGE - ing. Content-Based Image Retrieval), bilgisayarlı görü alanında çok çalışılmış ve çalışılmakta olan, temel ve zorlayıcı problemlerden biridir. İçerik tabanlı marka erişimi (İTME) ise, sadece görsel eşleştirme yerine, farklı seviyelerde ve farklı yönlerden benzerliklerin tespitine odaklanarak İTGE'den ayrılır. Şekil 1'de bazı marka benzerliği çeşitleri görülebilir. Günümüzde, ME sistemleri geliştirilirken İTME problemi için önerilmiş çözümlerden yararlanılmaktadır.



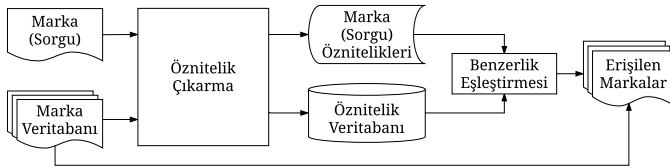
Şekil 1: Marka benzerliği çeşitleri.

İTME konusundaki önceki çalışmalar, geleneksel öznitelik çıkarma yöntemlerini kullanmışlardır. Bunlara düşük/orta seviye küresel öznitelikleri, basit ve düşük maliyetli şekil özniteliklerini ve renk tabanlı yaklaşımları örnek gösterebiliriz. Sonraki çalışmalarda ise gradyan tabanlı öznitelikler ve kelime çantası yöntemi kullanılarak başarı artırılmıştır. Ancak, bahsedilen geleneksel yöntemler yeterli başarıya ulaşamamıştır – örn. bknz. [1].

Görüntü sınıflandırma alanındaki çok büyük başarılarını [2] takiben, derin öğrenme ve özellikle derin evrişimsel sinir ağı (ESA - ing. Convolutional Neural Network) modelleri, nesne tespiti [3], anlamlı bölütleme [4] ve süper çözünürlük [5] gibi bir çok farklı bilgisayarlı görü probleminde en gelişmiş yöntem haline geldi. Görüntü tanıma problemi için danışmanlı eğitim ile öğrenilmiş genel öznitelikleri, diğer tanıma ve erişme problemlerine taşımada [6], [7], [8] erişilen hayret verici başarılar, ESA'ların aşırı karmaşık ve ince ayarlanmış, en başarılı yöntemlere çok sıkı rakip olabileceğini göstermektedir. İTGE probleminin son zamanlarda derin öğrenme camiasının [9], [6], [7], [10], [11] ilgisini çekmeye başladığı da göz önünde bulundurulduğunda, İTME probleminde de yüksek başarılar gösterebileceği düşünülebilir. Derin öğrenme yöntemlerinin İTME problemine uygulanmasını inceleyen ilk çalışmada [1], iyi bilinen ESA modellerinin doğrudan ve yalnızca öznitelik çıkarıcı olarak kullanılmasının, herhangi bir iyileştirme yöntemi kullanılmadan bile İTME probleminde, en başarılı geleneksel yöntemlerden üstün geldiği kanıtlanmıştır. Ancak, bu çalışmada derin öğrenmenin gerçekten bu problem için uygun olup-olmadığı irdelenmemiştir. Bu çalışmada, derin öğrenmeyi İTME problemine uygulamanın ötesinde, detaylı bir irdeleme sunulmaktadır.

### II. DENEYLER

Bu çalışmada, AlexNet [2], VGGNet [12], GoogLeNet [13], and ResNet [14] gibi yaygın olarak kullanılan önceden eğitilmiş yapay sinir ağları, METU Trademark Dataset [1] isimli geniş çaplı bir veri kümesi kullanılarak değerlendirilmiştir. Bu veri kümesinin seçilmesinin sebebi çok geniş çaplı bir test kümesinden (923,343 marka) ve gerçekte karşılaşılan marka ihlal durumları gözetilerek oluşturulmuş bir sorgu kümesinden oluşmasıdır.



Şekil 2: Deneylerin yapıldığı marka erişim sisteminin genel yapısı

TABLO I: ESA modelleri ve elle tasarlanmış öznitelik çıkarma yöntemlerinin sonuçlarının karşılaştırması.

Tür	Yöntem	Ortalama sıra	Normalleştirilmiş ortalama sıra
Elle tasarlanmış	HOG	242,166.1 ± 118,686.6	0.262 ± 0.129
	SIFT	<b>164,837.7 ± 133,932.5</b>	<b>0.179 ± 0.145</b>
	SURF	191,304.1 ± 139,696.4	0.207 ± 0.151
	Color	369,598.3 ± 161,895.1	0.400 ± 0.175
	LBP	254,971.8 ± 131,399.5	0.276 ± 0.142
	GIST	234,087.1 ± 159,585.2	0.254 ± 0.173
	SHAPEMES	203,408.2 ± 171,317.4	0.220 ± 0.186
Derin öznitelikler	AlexNet(FC7)	103,549.2 ± 157,877.9	0.112 ± 0.171
	AlexNet(Pool5)	125,300.9 ± 157,739.5	0.136 ± 0.171
	GoogLeNet(77S1)	108,662.5 ± 127,619.1	0.118 ± 0.138
	VggNet16(Pool5)	88,829.1 ± 112,370.7	0.096 ± 0.122
	VggNet16(FC7)	<b>79,538.5 ± 98,961.3</b>	<b>0.086 ± 0.107</b>
	VggNet16(FC8)	98,716.9 ± 100,910.4	0.107 ± 0.109
	ResNet50(FC1000)	101,570.0 ± 123,086.6	0.110 ± 0.133
	ResNet50(Pool5)	87,785.2 ± 127,086.5	0.095 ± 0.138
	ResNet101(FC1000)	87,964.9 ± 107,574.3	0.095 ± 0.117
	ResNet101(Pool5)	80,623.1 ± 118,812.7	0.087 ± 0.129
	ResNet152(FC1000)	101,237.6 ± 119,987.6	0.110 ± 0.130
	ResNet152(Pool5)	89,558.4 ± 128,587.2	0.097 ± 0.139

Değerlendirme esnasında her bir marka sorgusu için, sorgu kümesinde o sorgu ile benzer olarak işaretlenmiş markalar test kümesine eklenerek Şekil 2’de ana hatları görülen ME sistemi kullanılmıştır. ESA özniteliklerini karşılaştırmak için ortalama sıra (Denklem 1) ve normalleştirilmiş ortalama sıra (Denklem 2) ölçütleri kullanılmıştır. Bu tanımlarda,  $N$  test kümesi boyutunu,  $N_{ben}$  test kümesine eklenmiş benzer resim sayısını ve  $R_i$  eklenmiş benzer resimlerin sırasını ifade etmektedir.

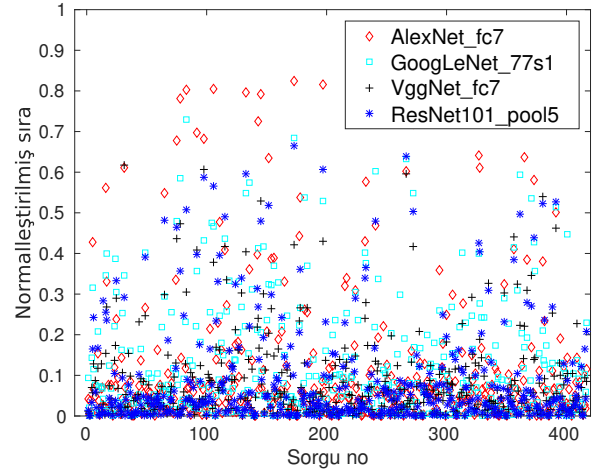
$$OS = \frac{1}{N_{ben}} \sum_{i=1}^{N_{ben}} R_i, \quad (1)$$

$$NOS = \frac{1}{N \times N_{ben}} \left( \sum_{i=1}^{N_{ben}} R_i - \frac{N_{ben}(N_{ben} + 1)}{2} \right). \quad (2)$$

Çalışma kapsamındaki deneyler, Intel i7-4770K işlemci, 32GB DDR3 hafıza ve GeForce GTX 760 ekran kartı kullanan bir bilgisayarda gerçekleştirilmiştir. Öznitelik çıkarma işlemleri için Caffe derin öğrenme altyapısı ve değerlendirme işlemleri için Matlab kullanılmıştır.

### III. SONUÇLAR VE TARTIŞMA

Marka kaydı esnasından yapılan sorgu sonucu ME sistemleri yardımıyla erişilen benzer markalar, uzmanlar tarafından incelenerek ihlal olup olmadığına karar verilmektedir. Uzmanların karşılaştırma yapabilmesi için sorgu sonuçları da belirli bir sayı ile sınırlandırılmalıdır. Bu sayı pratikte, uzmanlara ödenen kontrol ücretleri göz önünde bulundurularak, genellikle 1000 civarında seçilmektedir. Milyonlarca kayıtlı marka



Şekil 3: Tüm sorguların normalleştirilmiş ortalama sıra sonuçları (Noktalar yatay eksene ne kadar yakınsa, yöntem o kadar başarılıdır)

olduğu göz önüne alınırsa, bu rakamın oldukça küçük olduğu, dolayısıyla da ME sistemlerinin çok hassas olması gerektiği görülmektedir. Başarılı bir ME sisteminin ihlale sebep olabilecek tüm benzer markalara sonuç sınırı dahilinde erişim sağlaması beklenmektedir. Aksi takdirde benzer markalar incelemeye alınmayacak ve marka ihlali gerçekleşecektir.

Bölüm II’de açıklanan deneylerin sonuçları detaylı bir şekilde incelendiğinde, şu sonuca varılabilir: ESA’ların, geleneksel yöntemlere göre çok açık bir şekilde daha başarılı olduğu gösterilmiş [1] olsa da, sonuçlar beklentilerden halen çok uzakta kalmıştır. Tablo I ve Şekil 3’e bakılarak, şu yorum yapılabilir: ESA’lar bazı sorgular için benzer markaları dönme konusunda başarısız olmaktadır ve bu da ESA’ların kullanılmasına engel olabilir. ESA’ların, İTME problemini çözme yoluna taş koyan şey nedir? Sonraki kısımlarda bu konu ile ilgili bazı incelemelere ve muhtemel çözümlere yer verilmiştir.

#### A. ESA Özniteliklerinin Markaları Yeterince Betimleyememesi

ESA özniteliklerinin kimi sorgular için başarısız olmasının ilk sebebi, kullanılan modellerin farklı bir problem için eğitilmiş olmasıdır. İncelenen tüm modeller ImageNet [15] veri kümesi kullanılarak nesne sınıflandırma problemi için tasarlanmış ve eğitilmişlerdir. Bu problemin, İTME probleminden ayrıldığı noktalardan üçü bu noktada önemli rol oynamaktadır:

I Öncelikle, markalar bu veri kümesinde bulunan resimler gibi renk ve desen konularında fazla çeşitliliğe ve karmaşıklığa sahip değildir. Bu durum, Tablo I’de görülebilen, daha genel öznitelikler öğrenen katmanlar kullanılarak çıkarılan özniteliklerin daha özel öznitelikler öğrenenlere göre daha iyi sonuçlar göstermesi gerçeği ile desteklenebilir.

II İkinci nokta, İTME için sınıf içi farklılıklara odaklanan özniteliklere ihtiyaç duyulmasına rağmen, sınıflandırma probleminde sınıflar arası değişikliklere odaklanan özniteliklerin öğrenilmiş olmasıdır. ESA’ların katmanlı yapısı düşünüldüğünde, en son katmanlarda öğrenilen özniteliklerin sınıflandırma problemi için aşırı özelliğiği söylenebilir. Babenko vd. de en yüksek başarımın en yüksek kat-

manlarda değil, bunlardan iki seviye şağıdaki katmanlarda elde edildiğini göstermişlerdir [11]. Bununla beraber, çok altlarda kalan katmanların da problemler için yeterli ayırt edici öznelikler öğrenemediği söylenebilir.

III Son olarak, marka ihlalleri düşünülürken dönme, örtme, ölçek, zıtlık ve benzeri zorlukların İTME problemi için sınıflandırma problemine göre çok daha önemli olduğu söylenebilir. Bu zorlukların etkileri Tablo III'te görülebilir.

Bütün bu farklılıklardan kaynaklanan sorunları çözmek adına modellere ince ayar yapmak gerekir. İnce ayar esnasında veri çeşitlemesi (ing. data augmentation) kullanımı bahsedilen zorlukları yenmeye yardımcı olacaktır. Fakat kullanılan modellere markalardan İTME problemine yönelik öznelik çıkarmak için ince ayar yapmak zordur. Aşağıda, bu zorluğu aşmak için bazı yöntemler önerilmiştir:

- Girschik vd. [3] ve Long vd. [4] tarafından belirtildiği gibi, başarılı bir ince ayar için etiketli bir veri kümesine ihtiyaç vardır. Ancak, markalar için uygun etiketler bulmak veya oluşturmak çok güçtür. Vienna sınıflandırma kodu [16] markaları etiketlemede kullanılabilir. Bu yöntem her bir marka için üç seviyeden oluşan hiyerarşik bir etiket atamaktadır. Bütün olası kodlar hesaba katıldığında, problemi 1503'lü bir sınıflandırma problemine dönüştürerek ince ayar yapılabilir.
- Siyam mimariler (ing. siamese architectures) [17], Babenko ve diğerlerinin [11] İTGE problemi için de önerdiği gibi, İTME için de bir ince ayar yöntemi olarak kullanılabilir. İkiz veya üçüz siyam ağları ile bir marka grubu kullanılarak daha başarılı tanımlayıcılar elde edilebilir. İkiz olduğu durumda ağlar iki benzer marka ile beslenirken, üçüz olduğu durumda bunlara bir tane de benzer olmayan marka eklenir.
- Bir diğer olası yöntem ise yarı danışmanlı veya danışmansız eğitim ile ince ayar yapmaktır; örneğin derin modelleri simetrik şekilde bağlayıp oto-eşleyici (ing. autoencoder) eğiterek [18] veya görüntülerin içeriğini keşfederek [19].

#### B. ESA Özneliklerinin Küresel Tanımlayıcılar Olarak Kullanılması

İkinci sebep ise, yapılan deneylerde ESA'lar kullanılarak çıkarılan öznelikler ile oluşturulmuş tanımlayıcıların, küresel tanımlayıcılar olmasıdır. Bu yöntem, iki marka aynı seviyede tasarımı düzenine sahip olduğunda çok iyi çalışmaktadır. Tablo II'de görüldüğü üzere, tasarımı düzeni aynı olduğunda, sıra ve uzaklık (beklenildiği gibi) çok düşük çıkarken, düzeni değiştirmek, içerik aynı olsa da, sırada ve uzaklıkta yüksek bir artışa sebep olmaktadır. Buna ek olarak, örtme, kesme (ing. truncation), ve yazı veya çerçeveler gibi fazladan içerikler, İTME problemlerinde çok yaygın olarak görülmektedir ve başarıyı önemli ölçüde düşürmektedir. Tablo III'te bu zorlukların başarıyı ne kadar çok etkilediğine dair örnekler görülebilir. Tablo II ve III'teki örneklerde de görülebileceği gibi, marka içerisinde bulunan yazılar başarının düşmesinde en büyük rolü oynamaktadır. Tursun vd. [1], geleneksel öznelikler kullanarak karşılaştırma yaparken yazı bulunan kısımları yoksaymanın

başarıyı önemli derecede artırdığını göstermişlerdir. Buna dayanarak ESA öznelikleri çıkarılırken de aynı yöntemin izlenmesinin başarıyı artıracığı öngörülebilir. Küresel öznelikleri bir kenara bırakırsak, Tablo I'e göre, geleneksel yöntemler arasından yerel olanların, küresel olanlara göre çok daha yüksek başarı gösterdiği açık bir gerçektir. Bunun yanında, ESA özneliklerini yerel tanımlayıcılar olarak kullanmak İTGE problemlerinde en başarılı çözüm haline gelmiştir [20], [21]. Bütün bunlar göz önünde bulundurulduğunda, bahsedilen zorlukları aşmada ESA'ların bölgesel [6], [21] veya yerel [20] bir şekilde kullanılmasının başarıyı önemli oranda artıracığı söylenebilir.

#### C. Farklı Boyut ve En-Boy Oranındaki Markalar

ESA özneliklerinin başarısını olumsuz etkileyen bir diğer sebep de, "METU Trademark Dataset" in çok çeşitli boyutlarda ve en-boy oranlarında markalar içermesidir. Kullanılan ESA modelleri girdinin belir bir boyutta (örn.  $224 \times 224$ ) olmasını gerektirmektedir. Yüksek çözünürlüklü markalarda boyutlandırılmadan dolayı bazı bilgiler kaybedilirken, düşük çözünürlüklü markalar anlamlı bilgi sağlamada yetersiz kalmaktadır. Farklı boyut ve en-boy oranındaki markaları daha doğru bir şekilde tanımlayabilmek için, düşük çözünürlüklü markaların çözünürlüğü yeniledikten [5] sonra yerel ESA özneliklerini kümelemek [22] gibi yöntemler uygulanabilir.







#### D. Uzaklık Ölçütü

Önceki kısımlarda öznelikler ile ilgili sebepler tartışıldı. Fakat, başka bir sebep de kullanılan uzaklık ölçütü olabilir, çünkü İTGE ve İTME problemleri için uzaklık ölçütü hayati önem taşımaktadır. Bu çalışmada, markalar arasındaki benzerlik, ilintili ESA öznelik vektörleri arasındaki Öklid mesafesi kullanılarak hesaplanmıştır. Fakat, Öklid mesafe, deneysel olarak seçilmiş olup, en iyi uzaklık ölçütü olmayabilir. Bu noktada diğer uzunluk ölçütleri de değerlendirilerek daha iyisi olup olmadığı araştırılabilir. En iyi uzaklık ölçütü, benzer markalar arasındaki mesafeyi düşürürken benzemeyenler arasındakiini artıracaktır. Uzaklık ölçütü öğrenme, diğer görüntü erişim problemleri için çok çalışılmış bir alandır [23]. Fakat, bilgimiz dahilinde, İTME problemi için henüz böyle bir girişimde bulunulmamıştır. İTGE problemi için geliştirilmiş ikili kısıtlamalarla küresel uzaklık ölçütü öğrenme yöntemi [24], bu probleme kolayca uyarlanabilir. Bu yöntem ile aynı kategorideki veri noktalarını olabildiğince yakın tutmaya çalışırken, farklı kategorilerden olan veri noktalarını da olabildiğince uzak bir şekilde ayıran ölçütler öğrenmek hedeflenmektedir. Birbirine benzeyen markaları aynı kategoride, benzemeyen markaları da farklı kategoride gibi düşünerek belirlenecek kısıtlamalar ile yöntem İTME problemine uyarlanabilir. Bu sayede ME sistemleri için en uygun uzaklık ölçütü elde edilebileceği için başarı da önemli ölçüde artacaktır.

## IV. SONUÇ

ESA modellerini küresel tanımlayıcılar olarak doğrudan kullanmak, geleneksel metodlara göre daha başarılı olmuştur. Bu başarı, İTME problemi için özelleştirilmiş ESA'lar yardımıyla beklentilere ulaşılacağına dair umut verirken, yeni çalışmalara da yol göstermektedir. Bu çalışmada, ESA özneliklerini basit bir küresel yöntem ile kullanmanın doğasında var olan zayıflıklar incelenip özel ince ayar, probleme özgü uzaklık

TABLO II: Küresel ESA özniteliklerinin başarı ve başarısızlık örnekleri











Sorgu	Beklenen	Sıra	Uzaklık
	 CITY OF MELBOURNE	9	0.2933
	 CITY OF MELBOURNE	66,867	0.6847
		2	0.2153
	 MOTOROLA	496,591	0.8372

ölçütü öğrenme, ESA özniteliklerini yerel tanımlayıcılar olarak kullanma ve farklı boyutlara ve en-boy oranlarına sahip markaları işleme gibi çözümler önerilmiştir. Gelecekteki çalışmalarımızda, bahsedilen çözümlerin tamamamını denemeyi planlıyoruz olsak da siyam mimariler ile ince ayar ve uzaklık ölçütü öğrenme en umut verici çözümler olduğu için onlar ile devam etmeyi düşünmekteyiz.

#### BİLGİLENDİRME

Bu çalışma, SANTEZ-0029.STZ.2013-1 kodlu proje kapsamında T.C. Bilim, Sanayi ve Teknoloji Bakanlığı tarafından desteklenmiştir. Usta Bilgi Sistemleri A.Ş. ve Grup Ofis Marka Patent A.Ş.'ye veri kümesini oluşturmamızdaki katkılarından dolayı teşekkür ederiz.

TABLO III: Küresel ESA özniteliklerinin zorluklar karşısında başarısız olduğu durumlara örnekler

Sorgu	Beklenen	Sıra	Uzaklık	Sebebi
	 MOTOROLA	22,132	0.6494	Zıtlık
		64,039	0.6627	Çerçeve
		103	0.3776	Dönme
		5,349	0.5197	
		2,658	0.4685	Kesme
		35,620	0.6417	

#### KAYNAKLAR

- [1] O. Tursun, C. Aker, and S. Kalkan, "A large-scale dataset and benchmark for similar trademark retrieval," *arXiv preprint arXiv:1701.05766*, 2017.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *NIPS*, 2012.
- [3] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *CVPR*, 2014.
- [4] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *CVPR*, 2015.
- [5] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *European Conference on Computer Vision*. Springer, 2014, pp. 184–199.
- [6] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "Cnn features off-the-shelf: an astounding baseline for recognition," in *CVPR Workshops*, 2014.
- [7] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition." in *Icml*, vol. 32, 2014, pp. 647–655.
- [8] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *NIPS*, 2014.
- [9] A. Gordo, J. Almazán, J. Revaud, and D. Larlus, "Deep image retrieval: Learning global representations for image search," in *ECCV*, 2016.
- [10] K. Yan, Y. Wang, D. Liang, T. Huang, and Y. Tian, "Cnn vs. sift for image retrieval: Alternative or complementary?" in *ACM Int. Conf. on Multimedia*, 2016.
- [11] A. Babenko, A. Slesarev, A. Chigorin, and V. Lempitsky, "Neural codes for image retrieval," in *ECCV*. Springer, 2014, pp. 584–599.
- [12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014.
- [13] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," *CoRR*, vol. abs/1409.4842, 2014. [Online]. Available: <http://arxiv.org/abs/1409.4842>
- [14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *arXiv preprint arXiv:1512.03385*, 2015.
- [15] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *CVPR*, 2009.
- [16] WIPO, "O.h.i.m. figurative classification manual," 2012. [Online]. Available: <http://euipo.europa.eu/pdf/mark/vienna-en.pdf>
- [17] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *CVPR*, 2005.
- [18] J. Dong, X.-J. Mao, C. Shen, and Y.-B. Yang, "Unsupervised feature learning with symmetrically connected convolutional denoising auto-encoders," *arXiv preprint arXiv:1611.09119*, 2016.
- [19] C. Doersch, A. Gupta, and A. A. Efros, "Unsupervised visual representation learning by context prediction," in *ICCV*, 2015.
- [20] J. Yue-Hei Ng, F. Yang, and L. S. Davis, "Exploiting local features from deep networks for image retrieval," in *CVPR Workshops*, 2015.
- [21] A. S. Razavian, J. Sullivan, S. Carlsson, and A. Maki, "Visual instance retrieval with deep convolutional networks," *arXiv preprint arXiv:1412.6574*, 2014.
- [22] A. Babenko and V. Lempitsky, "Aggregating local deep features for image retrieval," in *ICCV*, 2015.
- [23] L. Yang and R. Jin, "Distance metric learning: A comprehensive survey," *Michigan State University*, vol. 2, no. 2, 2006.
- [24] E. P. Xing, A. Y. Ng, M. I. Jordan, and S. Russell, "Distance metric learning with application to clustering with side-information," in *NIPS*, 2002.