Market Ürünlerinin Görüntü İşleme ile Planogram Uygunluğunun Kontrolü Planogram Matching Control in Grocery Products by Image Processing

Ünsal Gökdağ¹, Mehmet Yasin Akpınar¹, Yusuf Sinan Akgül²

¹İdea Teknoloji Çözümleri, İstanbul, Türkiye
{unsal.gokdag,mehmet.akpinar}@ideateknoloji.com.tr

²Gebze Yüksek Teknoloji Enstitüsü, İstanbul, Türkiye
akgul@gyte.edu.tr

Özetçe —Market raflarındaki ürünlerin dizilimi bir takım pazarlama stratejilerine göre tasarlanmaktadır. Bu tasarımlar planogram adı verilen şemalarla ifade edilir. Marketlerdeki ürünlerin planogramlara uygunluğu ise günümüzde bir görevli tarafından manüel olarak kontrol edilmekte olup zaman kaybına neden olmaktadır. Bu bildiride bu soruna çözüm bulma amacıyla bir sistem önerilmektedir. Tasarlanan sistem market raf görüntüsünü girdi olarak alır ve planogram uygunluğunu çıktı olarak verir. Önerilen yöntem raf çizgisi tespitinde %99.03, ürün tespitinde %97.31 kesinlik değerine ulaşırken, marka tanımada ise %99.21 doğruluk göstermiştir.

Anahtar Kelimeler—Kademeli Nesne Tanıma, Gauss Karışım Modeli, Fisher Vektör, SIFT, Ekstrem Öğrenme Makineleri.

Abstract—The placement of products in grocery shelves is designed according to several marketing strategies. These designs are represented by schemas named as planograms. Currently, the planogram matching control of grocery products are counducted by a personel manually and causes time loss. In this paper, a system that aims to solve this problem is proposed. The sytem takes a grocery shelf image as the input and gives the planogram compatibility as the output. Proposed system performs with %99.03 precision in shelf line detection, %97.31 precision product in detection and %99.21 accuracy in brand identification.

Keywords—Cascaded Object Detection, Gaussian Mixture Model, Fisher Vector, SIFT, Extreme Learning Machines.

I. GİRİŞ

Market raflarındaki ürünlerin dizilimleri planogram adı verilen şemalara uygun olarak yapılmakta ve bu şemalar birtakım pazarlama stratejilerine ve kurallarına göre tasarlanmaktadır. Raflara dizilen ürünlerin bu şemalara uygunluğunun kontrolü ise bir görevli tarafından büyük ölçüde manüel olarak yapılmakta olup zaman kaybına yol açmaktadır. Bu işlemin otomatik olarak gerçekleştirilebilmesi bu kaybı önleyerek envanter yönetimi acısından büyük önem tasımaktadır.

Bu çalışmada yukarıda bahsedilen probleme odaklanan bir sistem önerilmektedir. Sistem 2 aşamada çalışmaktadır. İlk aşamada mobil bir cihaz aracılığıyla alınan raf görüntüsü

üzerinde ürün kutuları tespit edilirken, ikinci aşamada ise tespit edilen bu kutular üzerinden marka tanıma gerçekleştirilip planograma uygunluğu kontrol edilmektedir.

Bildirinin akışı ise şu şekildedir. II. bölümde tekniğin bilinen durumu kaynaklar ile birlikte özetlenmektedir. Ardından III. bölümde önerilen yöntem detaylarıyla beraber verilmektedir. IV. bölümde ise çalışmadan elde edilen sonuçlar tekniğin bilinen durumundaki sonuçlar ile kıyaslanarak sunulmaktadır. V. bölüm olan kapanış bölümünde ise bildirinin özeti yapılıp gelecek çalışmalardan bahsedilmiştir.

II. İLGİLİ ÇALIŞMALAR

Bu konuda yapılan çalışmalar arasında planogram çıkartma amacıyla SIFT (Scale-Invariant Feature Transform) [1] ve OCR (Optical Character Recognition) yöntemlerini birlikte kullanan bir patent çalışması yer almaktadır [2]. 2008 yılında tamamlanan bir çalışmada ise gerçek hayat görüntüleri üzerinde uzamsal piramit madenciliği ve SIFT tanımlayıcıları ile logo tanıma yapılmaktadır [3]. Bir başka çalışmada ise görme engelliler için market raflarında ürün tanıma yapan bir sistem önerilmiştir [4]. Bu çalışmada hız bakımından daha yüksek performansa sahip olan SURF (Speeded Up Robust Features) tanımlayıcıları kullanılmıştır. [5]'te SIFT tanımlayıcıları ile bir görüntüde veritabanındaki imgelerden hangisinin bulunduğunu tespit eden bir çalışmaya yer verilmiştir. Bir başka çalışmada ise bir kütüphanedeki kitapları tanıma üzerine bir yöntem önerilmiştir [6].

2014 yılında önerilen bir yöntemde market raflarından planogram çıkartma işlemi 3 aşamalı bir şekilde gerçekleştirilip, ümit verici sonuçlar elde edilmiştir [7]. Çalışmada ilk adım olarak raf çizgileri tespit edilmeye çalışılmış, sonrasında tespit edilen raflar içerisinde ürün kutularının tespiti yönüne gidilmiştir. Son aşamada ise ürün kutularının içerisinde kalan alanlardan marka tanıma gerçekleştirilmiştir ve 1 sınıf üzerinde test edilmiştir. Aynı yıl tamamlanan diğer bir çalışmada ise marka tanıma modülü güçlendirilip 10 sınıf üzerinde test edilmiştir [8].

III. YÖNTEM

Giriş bölümünde basedildiği gibi bu çalışmada önerilen yöntemde market ürünlerinin tespiti için 2 aşamalı bir işlem uygulanmaktadır. Sistemin 2 aşamalı olmasındaki asıl amaç marka tanıma işlemini bütün görüntü içerisinde yapmak yerine, önce ürün kutularını mümkün olduğunca doğru bir şekilde tespit edip, sonrasında görüntüden kırpılan bu kutular içerisinden marka tanıma işleminin gerçekleştirilmesidir. Böylece ürün tespitinden daha maliyetli olan marka tespiti işlemi kolaylaştırılıp toplam sürenin kısaltılması hedeflenmektedir. [7]'deki çalışmadan farklı olarak raf tespiti ve ürün tespiti ayrı ayrı ele alınmayıp doğrudan ürün tespiti yönüne gidilmiştir.

A. Ürün Tespiti

Ürün tespiti aşamasının çıktısı doğrudan marka tanıma aşamasında kullanıldığı için bu adım oldukça önemlidir. Sisteme gelen görüntü üzerinde gerçekleştirilen ilk işlem Viola-Jones algoritmasını [9] baz alan bir Kademeli Nesne Tanıyıcı (Cascaded Object Detector) ile ürün kutularının tespit edilmesidir. Bu algoritma en boy oranı sabit ya da çok az değişen nesnelerde oldukça başarılı çalışmasıyla bilinmektedir. Ancak tespit edilecek ürünlerin raf görüntüsü içerisindeki boyutları önceden bilinmediğinden bu tanıyıcı en boy oranı sabit olacak şekilde irili ufaklı kutular tespit etmektedir.

Tespit edilen kutuların doğruluğundan emin olmak için bir dizi işlemden geçirilmeleri gerekmektedir. Bunlardan ilkinde tespit edilen kutuların ortalama ve medyan yükseklikleri hesaplandıktan sonra, ortalama yüksekliğin 0,5 ve 1,5 katı ile medyan yüksekliğin 0,8 ve 1,4 katı arasında olmayan kutular elenmektedir. Bu işlemin amacı çok büyük ve çok küçük olan kutuların yok edilmesini sağlamaktır.

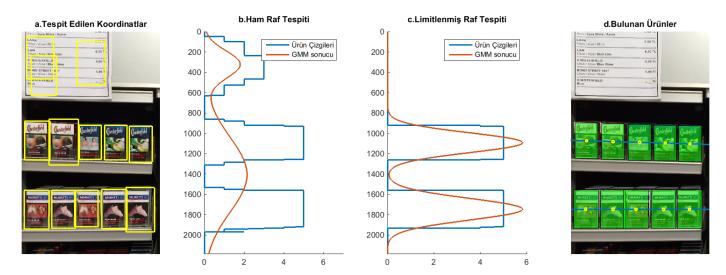
İkinci işlemde ise kutuların dikey eksendeki koordinatları üzerinden eğitilen bir Gauss Karışım Modeli (Gaussian Mixture Model) ile kutuların kümelenmesi sağlanır. Bu işlem gerçekleştirilirken görüntü dikey eksende taranarak dikey eksenindeki her bir koordinat üzerinde kaç adet sigara paketi izdüşümü olduğu çıkartılır. Bu bilgi daha sonra Gauss Karışım Modeli eğitimi için kullanılır. Böylece kutuların satırlara göre ayrıştırılması sağlanmış olur.

Ürün koordinatlarının dikey eksen üzerine izdüşümü işlemi sonucunda eğitilen Gauss Karışım Modeli'nde sistematik hatalar meydana gelmektedir. Yanlış pozitif ürün kutularının dikey eksen üzerindeki izdüşümü ürün koordinatları histogramını etkilemekte ve elde edilen veriler üzerine oturtulan Gauss Karışım Modeli yanlış sonuçlar çıkartmaktadır (Şekil 1.b). Bu hataları engellemek için dikey eksen üzerindeki her ürün sayısı eğer raf üzerindeki toplam ürün sayısını %60'ından daha yüksek ise alınmakta, aksi takdirde 0 olarak değerlendirilmektedir (Şekil 1.c). Bu sayede yanlış pozitif ürün koordinatları raf koordinatlarının yanlış girilmesini etkileyememektedir.

Bir sonraki adımda satırlarına göre ayrıştırılan kutuların merkez koordinatları bulunup bu noktalardan geçen doğrular saptanmaktadır. Ürün kutuları, yatay eksendeki koordinatları sabit tutulacak şekilde merkezleri saptanan bu çizgilere oturtularak, medyan kutu boyutlarında yeniden oluşturulmaktadır. Bu aşamada resmin çekildiği perspektifin etkilerinden kurtulmasının yanında dikey eksendeki ufak kaymaların da düzeltilmesi sağlanmaktadır.

Bu işlemlerin ardından tespit edilemeyen ürün kutuları planograma uyacak şekilde oluşturulmaktadır. Bunun için bütün görüntü içerisinde tespit edilen kutulardan yatay eksen koordinatı en küçük ve en büyük olanı bulunmaktadır. Her raf içindeki eksik olan kutular yine bir önceki adımda tespit edilen çizgilere merkezleri yerleşecek şekilde ve medyan kutu boyutlarında oluşturulmaktadır. Takip edilen sıra ise aşağıdaki şekilde belirlenmiştir:

 Satır ortası: bir satırın ilk ve son ürünleri arasında büyük boşluklar bulunuyorsa (örneğin, medyan kutu genişliğinden daha fazla) bu boşluk giderilinceye



Şekil 1: a) Kademeli Nesne Tanıyıcı algoritmasının bulduğu kutulardan ortalama ve medyan elemesi sonucu elde edilen koordinatlar. b) Gauss Karışım Modeliyle hesaplanan ham raf tespiti (rafların dışında kalan üç kutu eğitim sırasında modeli yanıltmaktadır). c) Limitleme işleminden sonra eğitilen Gauss Karışım Modeliyle elde edilen yeni raf tespiti. d) Ürün tespiti aşamasının diğer adımları da tamamlandıktan sonra bulunan ürün kutuları.

- kadar kutu eklenir.
- Satır başı: bir satırın başındaki kutu bütün görüntüde tespit edilen yatay koordinatı en küçük kutuya medyan kutu genişliğinden daha uzaksa, bu uzaklık giderilinceye kadar kutu eklenir.
- 3) Satır sonu: bir satırın sonundaki kutu bütün görüntüde tespit edilen yatay koordinatı en büyük kutuya medyan kutu genişliğinden daha uzaksa, bu uzaklık giderilinceye veya satırda olması gereken kutu sayısı elde edilene kadar kutu eklenir.

Bu aşamada dikkat edilmesi gereken nokta ise yatay eksende tespit edilen en küçük ve en büyük koordinatların aşılmamasıdır.

Bütün bu işlemler tamamlandığında ürün tespiti aşaması sona ermektedir. Sonrasında tespit edilen bu kutular içindeki alanlar kırpılarak marka tanıma algoritmasına verilmektedir.

B. Marka Tanıma

Verilen bir görüntü üzerindeki ürün koordinatlarının tespitinden sonra belirlenen çerçeveler içerisinde kalan kısımlarından marka tanıma işlemi yapılmaktadır. Yapılan çalışmalarda sigara paketleri üzerindeki ayırt edici bilgilerin paket görüntülerinin üst %40'lık kesiminde bulunduğu tespit edilmiş ve marka tanıma işleminde kullanılmıştır.

Marka tanıma kısmı birkaç aşamadan oluşmaktadır. İlk aşama olarak üst %40'lık kısımdaki görüntülerin (ürün görüntüsü) hepsi 100*160 piksel boyutuna getirilmektedir. Bu işlem sırasında piksel değerleri küçültülürken lineer interpolasyon kullanılmıştır. Bütün ürün görüntüleri aynı boyuta getirildiği zaman bu görüntüler üzerinden yoğunSIFT (denseSIFT) özellikleri çıkartılmaktadır. Özellik çıkartma aşamasında yoğunSIFT özelliklerinin kullanılmasının sebebi veri boyutunu eşitleme imkanı sağlamasıdır. Yapılan deneyler sonucu yoğunSIFT algoritmasındaki parça boyutu P parametresi 15, adım mesafesi G parametresi 5 olarak belirlenmiştir.

Marka tanıma aşamasında daha önce kullanılan Kelime Torbalama işleminden (Bag of Words), Peronnin ve Dance [10] tarafından kullanılan Fisher Vektörlerine geçilmiştir. Fisher Vektörlerin Kelime Torbalama işleminden farkı kümeleme aşaması sırasında Gauss Karışım Modellerinin kullanılarak özelliklerin küme merkezlerine göre üyelik bilgilerinden (norm 0) ziyade kümenin merkezi ve varyansına oranla uzaklığının kullanılmasıdır (norm 1 ve norm 2). Elde edilen işlem sonucunda çıkan özellik boyutu 2DK boyutunda olup, D yoğunSIFT algoritmasında kullanılan özellik boyutunu ifade ederken, K Gauss Karışım Modeli'nin küme sayısıdır. Fisher Vektörlerin farklı ışık koşullarından ve poz açılarından kaynaklanacak olan hatalara karşı daha dayanıklı olması için Fisher Vektör değerleri norm 1 ve norm 2 değerlerine göre -1 ile 1 arasında olacak şekilde normalize edilmiştir.

YoğunSIFT özellikleri şekil tabanlı çalıştığından dolayı Fisher Vektörlerinin resmin şekli ile ilgili hangi marka içerisinde bulunduğu bilgisinin histogramını barındırır. Ürün görüntüleri içerisindeki renk bilgilerini kullanmak için ürün görüntüsünden çıkartılan öznitelik vektörlerine renk histogramı da eklenmektedir. Birleştirilecek olan özelliklerin bir diğerini domine etmemesi için histogram değerleri 0 ile 1 arasında normalize edilmiştir.

Şekil ve renk tabanlı yöntemlerin birlikte çalışma başarısı oldukça yüksek olmasına rağmen farklı ışık koşulları ve flaşlı görüntülerin oluşturduğu bölgesel ışık yoğunluğu değişimlerine karşı dayanıklı olmadığı tespit edilmiştir. Bu sorunu aşmak için Yerel İkili Örüntüler (Local Binary Patterns) kullanılmıştır. 100*160 boyutundaki ürün görüntülerinden 98*158 boyutunda elde edilen Yerel İkili Örüntü bilgisi çıkartılmış ve bir önceki aşamada yapıldığı gibi 0 ile 1 arasına normalize edilmiştir.

Tüm vektörlerin birleştirilmesi ile özellik vektörü oluşturulmuş olur:

$$f(I) = \begin{bmatrix} f_s(I) \\ f_c(I) \\ f_l(I) \end{bmatrix} \tag{1}$$

Burada ürün görüntüsünden çıkartılan özellik f üç adet özellik vektörünün birleştirilmesinden oluşur, f_s yoğunSIFT özelliklerinden elde edilen şekil tabanlı Fisher Vektörünü, f_c renk histogramını ve f_l Yerel İkili Örüntülerin bilgisini ifade eder

Özellik vektörünün oluşturulmasından sonra sınıflandırma işlemi ELM (Extreme Learning Machines) [11] ile gerçekleştirilir. Bu aşamada dikkat edilmesi gereken bir detay özellik vektörlerinin Mercer Kuralı'na [12] uygunluğundan dolayı ELM eğitiminde kernel kullanılabilmesidir. Bu çalışmada kullanılan kernel fonksiyonu Denklem 2'de görülmektedir:

$$K(f(I), X) = f(I)X^{\top} \tag{2}$$

Denklem 2'deki f(I) ürün görüntüsünden çıkartılan özellik vektörünü, X ise marka tanıma aşamasında kullanılan eğitim kümesinin Fisher Vektörlerinden oluşan matrisi ifade eder. Böylece, kernel fonksiyonu eğitim matrisini kovaryans matrisine dönüştürür ve bu matris kullanılarak ELM eğitimi tamamlanır. Bu aşamada marka görüntüleri öznitelik vektörü uzayından marka uzayına aktarılmış ve bu sayede ham özellik vektörü değerleri yerine diğer özellik vektörlerinin toplamları cinsinden (marka uzayı) ifade edilmiş olur.

Test aşamasında ise sınıflandırılacak olan özellik vektörü eğitim matrisinin transpozu ile çarpılarak marka uzayına projekte edilir ve sınıflandırma işlemi gerçekleştirilir. Elde edilen sonuç tanımlanmak istenilen özellik vektörünün her bir sınıf için sahip olduğu güven katsayısını verir ve sistem bu katsayılardan en yüksek olan değerin indisi üzerinden sınıflandırma işlemini gerçekleştirir.

Ürün tespiti ve marka tanıma işlemleri tamamlandıktan sonra veritabanında bulunan planogram bilgisiyle, tespit edilen ürünlerin dizilimlerinin birebir eşleştirilerek hangilerinin doğru ve hangilerinin yanlış olduklarının çıkarılması işlemi gerçekleştirilir. Böylece market rafındaki ürünlerin, tasarlanan planograma uygunluğunun kontrolü tamamen otomatik bir şekilde yapılmaktadır.

IV. SONUÇLAR VE KARŞILAŞTIRMA

Bu bölümde önerilen sistemin çeşitli bölümlerinde yürütülen testlerin sonuçları daha önceki çalışmalarla karşılaştırarak sunulmuştur.

İlk test raf çizgileri doğruluğunu ölçmek için gerçekleştirilmiştir. Bunun için sistemin saptadığı raf çizgilerinin yatay

eksene göre orta noktalarının, rafların olması gereken orta noktalarına olan öklid uzaklıkları hesaplanmıştır. Bu uzaklık medyan kutu boyutunun %5-%15-%25'inden daha küçük ise raf çizgisi doğru olarak kabul edilecek şekilde 3 defa tekrarlanmıştır. Sonuçlar [7]'deki çalışma sonucuyla birlikte Tablo I'de verilmiştir.

Tablo I: Raf Çizgisi Tespit Başarımı

	İzin Verilen Hata Payı	Tespit Başarımı
[7]	-	%91.27
Önerilen Yöntem	%5	%67.43
	%15	%98.42
	%25	%99.03

Tablo I'de de görüldüğü gibi raf çizgisi tespit başarımı izin verilen hata payı %15 olarak kabul edildiğinde bile eski çalışmadan daha üstün performans sergilemektedir.

Ölçüm yapılan bir diğer nokta ise tespiti yapılan ürün kutularının doğruluğuyla ilgilidir. Bu ölçüm için [8]'de kullanılan 354 raf görüntüsünden 4 tanesi görüntülerin çok fazla bulanık olması veya bazı ürünlerin yarısına yakın kısmının görüntü dışında kalması nedeniyle elenerek geri kalan 350 raf görüntüsü üzerinden ürün tespiti algoritması çalıştırılmıştır. Elde edilen kutuların merkez koordinatları, kutuların olması gereken merkez koordinatlarına olan öklid uzaklıkları hesaplanmıştır. Yine bir önceki testteki gibi bu uzaklık medyan kutu boyutunun %5-%15-%25'i olacak şekilde tekrarlanmıştır. Elde edilen bulgular karşılaştırmalı olarak Tablo II'de sunulmuştur.

Tablo II: Ürün Tespit Başarımı

	Izin Verilen Hata Payı	Kesinlik	Geri Getirme	F1-Skoru
[7]	-	%89	%88	%88.50
[8]	-	%81	%94	%87.02
Önerilen Yöntem	%5	%43.95	%45.48	%44.70
	%15	%91.99	%95.19	%93.56
	%25	%94.05	%97.31	%95.65

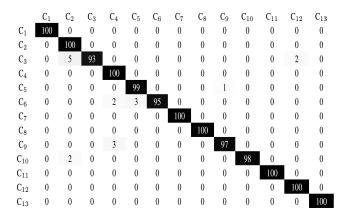
En yüksek başarım yüzdesi izin verilen hata payı %25 iken tespit edilmiş olup önerilen yöntemin başarımı eski çalışmaların ikisinden de yüksektir.

Son olarak güçlendirilen marka tanıma algoritmasını test etmek için manüel kırpılmış olan ürün görselleri üzerinden test gerçekleştirilmiştir. 13 sınıf ile yapılan testlerde elde edilen hata matrisi Şekil 2'de paylaşılmıştır. [8]'de %92,3 olan doğruluk oranı bu çalışmada önerilen yöntem ile %99,21 olarak kaydedilmiştir.

V. KAPANIŞ

Bu bildiride market raflarındaki ürünlerin planogram adı verilen şemalara uygunluğu otomatik olarak kontrol edebilecek bir sistem önerilmiştir. Giriş bölümünde bildirinin genel amacı açıklanıp, akıştan bahsedilmiştir. Ardından İlgili Çalışmalar bölümünde tekniğin bilinen durumu kaynaklarla anlatılmıştır. 3. bölüm olan Yöntem bölümünde ise önerilen metot detaylarıyla açıklanmış, eski yöntemlerle olan benzer ve farklı yönleri ele alınmıştır. Sonuçlar ve Karşılaştırma bölümünde elde edilen bulgular karşılaştırılmalı olarak sunulmuştur.

Mevcut projeye ileriki zamanlarda yapılması planlanan geliştirmeler sabit özellik çıkartma algoritmalarından uzaklaşıp



Şekil 2: Marka Tanıma Hata Matrisi

görüntü özelliklerinin otomatik olarak Derin Öğrenme (Deep Learning) aracılığı ile çıkartılması ve ürün koordinatlarının bulunması için Tekrarlı Sinir Ağlarının (Recurrent Neural Networks) kullanılması planlanmaktadır.

VI. TEŞEKKÜR

Bu çalışmamız TÜBİTAK TEYDEB tarafından 3130322 no'lu VisionFetch Projesi kapsamında desteklenmiştir.

KAYNAKÇA

- A. Opalach, A. Fano, F. Linaker, and R. Groenevelt, "Planogram extraction based on image processing," Patent US 8 189 855, May 29, 2012.
- [2] D. G. Lowe, "Distinctive image features from scaleinvariant keypoints," International journal of computer vision, vol. 60, no. 2, pp. 91–110, 2004.
- [3] J. Kleban, X. Xie, and W. Y. Ma, "Spatial pyramid mining for logo detection in natural scenes," in Proc. IEEE International Conference on Multimedia and Expo (ICME), Jun 2008.
- [4] T. Winlock, E. Christiansen, and S. Belongie, "Toward real-time grocery detection for the visually impaired," in Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on, June 2010, pp. 49 –56.
- [5] A. Auclair, L. D. Cohen, and N. Vincent, "How to use SIFT vectors to analyze an image with database templates." in Adaptive Multimedia Retrieval, ser. Lecture Notes in Computer Science, N. Boujemaa, M. Detyniecki, and A. Nürnberger, Eds., vol. 4918. Springer, 2007, pp. 224–236.
- [6] Y. Akiyama and M. Ito, "Book recognition from color images of book shelves." in MVA, 1998, pp. 106–110.
- [7] G. Varol, R. S. Kuzu, Y. S. Akgül, "Rafta ürün diziliminin görüntü işleme ile tespiti (Product placement detection based on image processing)," 22nd IEEE Conference on Signal Processing and Communications Applications (SİU), Trabzon, April 2014.
- [8] G. Varol, R. S. Kuzu, "Toward retail product recognition on grocery shelves." International Conference on Image, Vision and Computing, Paris, September 2014.
- [9] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features." 2001, pp. 511–518.
- [10] F. Perronnin, C. Dance, "Fisher kernels on visual vocabularies for image categorization." In: Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, June 2007, pp. 1–8
- [11] G.B. Huang, Q.Y. Zhu, C.K. Siew, "Extreme learning machine: Theory and applications. Neurocomputing 70(13), 489 – 501 (2006), neural Networks Selected Papers from the 7th Brazilian Symposium on Neural Networks (SBRN '04) 7th Brazilian Symposium on Neural Networks
- [12] J. Mercer, "Functions of positive and negative type and their connection with the theory of integral equations", Philosophical Transactions of the Royal Society A 209, 1909, (441–458): 415–446