

Inception-Unet ve Klasik Yöntemler ile Bina Bölütleme

Building segmentation with Inception-Unet and classical methods

İbrahim Delibaşoğlu

Fen Bilimleri Enstitüsü^a, Yazılım Mühendisliği^b

Yalova Üniversitesi^a, Sakarya Üniversitesi^b

Yalova^a, Sakarya^b, Türkiye

ibrahimdelibasoglu@sakarya.edu.tr

Müfit Çetin

Bilgisayar Mühendisliği

Yalova Üniversitesi

Yalova, Türkiye

mufit.cetin@yalova.edu.tr

Özetçe—Uydu görüntülerinde uzamsal çözünürlüğün artması ile birlikte kentsel alanlardaki binaların otomatik tespitine yönelik uzaktan algılama çalışmaları giderek önem kazanmıştır. Bu konuda bir çok klasik görüntü işleme temelli yöntem önerilmiş, son zamanlarda ise derin öğrenme yöntemleri popülerlik kazanmıştır. Bu çalışmada Unet mimarisi ek bölütleme modülü ile bir arada kullanılmış ve Unet performansını arttırmak için Inception-Unet mimarisi önerilmiştir. Elde edilen sonuçlar klasik yöntemlerle karşılaştırılmıştır.

Anahtar Kelimeler — unet; inception-unet; bina bölütleme; derin öğrenme.

Abstract—Remote sensing studies for automatic detection of buildings in urban areas have become important with the increase in spatial resolution of satellite images. Many classical image processing based methods have been proposed in this regard, and deep learning methods have gained popularity recently. In this study, Unet architecture is used with proposed additional segmentation module, and Inception-Unet architecture is proposed to improve Unet performance. The results are compared with classical image processing methods.

Keywords— unet; inception-unet; building segmentation; deep learning.

I. GİRİŞ

Uydu görüntülerinde uzamsal çözünürlüğün artması, metre (hatta cm) mesafelerine kadar düşmesi nispeten küçük nesnelerin de algılanabilmesine olanak sağlamıştır. Uzamsal çözünürlüğü yüksek olan bu görüntüler kullanılarak kent alanlarındaki değişikliklerin analizi, otomatik gözlemlenmesi gibi uygulamalar gerçekleştirilebilmektedir. Bunun için uzaktan algılama görüntülerinde binaların tespit edilmesi üzerine birçok çalışma yapılmıştır. Binaların tespit edilmesinde yararlanılan önemli bir özellik binaların oluşturduğu gölgelerdir[1,2]. Bazı çalışmalarda bina olması muhtemel yerler tespit edildikten sonra karar aşamasında hipotezin doğrulanması ve bina yüksekliğinin tahmin edilmesi probleminde gölge bilgisi kullanılmıştır[3,4].

978-1-7281-7206-4/20/\$31.00 ©2020 IEEE

Liow ve Pavlidis tarafından gerçekleştirilen çalışmada ilk olarak kenar bilgileri çıkarılmış daha sonra ise kenarların oluşturduğu bölgelerin hemen yanındaki gölge alanlar kontrol edilerek binalar tespit edilmeye çalışılmıştır [5]. Bu yöntemde kenar bilgisinden tam ve doğru bir şekilde kapalı alanların bulunması büyük bir problemdir. Gerçekleştirilen bir başka çalışmada [6] ise renk bilgisi kullanılarak gölge alanlar, daha sonra ise güneş ışığının yönüne bakılarak binaların konumu belirlenmeye çalışılmıştır. Önerilen yöntemde canny kenar bulma algoritması ile bulunan kenarlar kullanılarak dikkörtgen alanlar belirlenmiştir. Yöntem kenar bilgisine fazla bağımlı/hassas ve sadece dikkörtgen şeklindeki binalar için iyi sonuçlar vermektedir.

Akçay ve Aksoy [7] tarafından geliştirilen yöntemde ise ilk olarak bina olmaya aday yerler gölge bilgisi ile belirlenir. Bunun için bulunan tüm alanlar ile gölge alanları arasındaki ilişkiler için kısıtlamalar belirlenir ve segmentlere ayrılmış bu alanlar arasında kümeleme işlemi yapılır. Ali özgün Ok [8] tarafından gerçekleştirilen çalışmada tüm bu evrelere ek olarak CRF optimizasyon algoritması kullanılarak en son karar aşamasında piksel seviyesinde sınıflama yapılmıştır. Bu çalışmada ilk olarak gölge alanlar üzerinden bina olması muhtemel alanlar bulunmuştur. Bunun için de renk bilgisi kullanılarak ilk olarak gölge alanları genişletilerek bulunan küçük parça halindeki gölgeler birleştirilmiştir. Daha sonra ise her bir bölge için CRF algoritması çalıştırılmış, pikseller bina ya da bina değil olarak etiketlenmiştir. Algoritmanın temelinde başlangıçtaki gölge alanların bulunması yer almaktadır ve bu aşamada bulunan bölgelerin birleştirilmesi son derece önem arz etmektedir.

Literatürde geçen birçok çalışmanın, gölge ve önışleme adımlarının üzerine kurulduğu, eşik değerlere bağımlı ve çok aşamalı sınıflama işlemlerinden sonra belirli bir başarıya ulaşılabildiği düşünüldüğünde, tüm bu izlenen yöntemler yerine görüntüden özellik çıkarımı için derin öğrenme tekniklerinin daha genel ve iyi sonuçlar verebileceği düşünülebilir. Son dönemde derin öğrenmenin görüntü işleme problemlerinin çözümünde başarılı sonuçlar vermesiyle birlikte, uydu görüntülerinden bina tespiti problemi için de

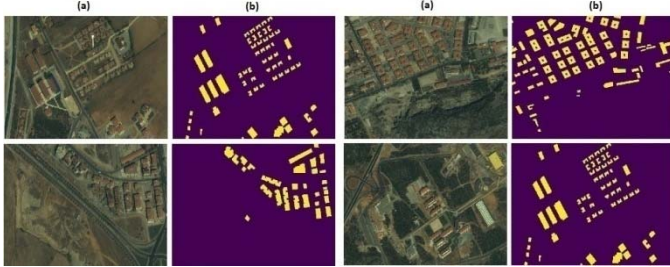
derin öğrenme tercih edilmeye başlanmıştır. Klasik görüntü işleme temelli yöntemlerde, kullanıcı tanımlı veya veri kümesine göre değişebilecek bazı parametreler ile hassas yaklaşımlar mevcuttur. Derin öğrenme ise bir çeşit makine öğrenmesi yöntemidir ve bu tip yöntemlerde uzman kişilerin özellik çıkarımı için spesifik ve detaylı algoritmalar geliştirmesi yerine, çözülecek problem için örnek veri setlerinin hazırlanıp etiketlenmesi suretiyle makine öğrenmesi algoritmalarına girdi olacak veriler hazırlanır. Etiketlenmiş veriler modele verilerek, modelin eğitilmesi sağlanır. Görüntü işleme problemlerinde kullanılan derin öğrenme modellerinde ise, görüntü işleme için özel hazırlanmış katmanlar ile eğitim işlemi esnasında görüntüden özellik çıkarım işlemi de modele dahil edilmiş olur.

Bu çalışmada biyomedikal görüntü işleme ve bina tespitinde yaygın kullanılan *Unet*[9] derin öğrenme mimarisi kullanılmıştır. *Unet* performansı değerlendirilmiş ve gölge bilgisi ile elde edilen sonuçların iyileştirilmesi üzerine bir yöntem geliştirilmiştir. Böylece az sayıda eğitim verisi kullanılarak bile tatmin edici sonuçlar elde edilebildiği gösterilmiştir. Ayrıca Inception blokları ile *Unet* mimarisine genişlik kazandırılarak yeni bir mimari önerilmiştir. Önerilen yöntem, literatürdeki MPP [10], GrabCut [11], MLP [8] gibi klasik görüntü işleme temelli yöntemler ile karşılaştırılmıştır.

II. VERİ SETİ VE GÖLGE TESPİTİ

A. Veri kümesi

Bu çalışmada kullanılan görüntüler IKONOS (1 m) ve Quickbird (60 cm) olmak üzere iki farklı uydudan elde edilmiştir. Tüm görüntüler B-G-R ve NIR olmak üzere 4 band içermektedir ve 11 bit radyometrik çözünürlüğe sahiptir. Veri kümesi farklı çözünürlüklerdeki toplam 14 görüntüden oluşmaktadır[12].



Şekil. 1. Veri kümesinden örnek görüntüler

Şekil 1'de veri kümesinden yer alan 4 örnek görüntünün RGB görünümüleri (a) sütunlarında, her bir görüntüdeki binaların etiketlendiği görüntüler ise (b) sütunundagösterilmiştir.

B. Gölge tespiti

Gölge bilgisi bina gibi insan yapımı nesnelere tespit edilmesinde yaygın olarak kullanılmaktadır. Örneğin; uydu görüntülerinden binaların tespit edilmesi amaçlanıyorsa sadece şekil özelliklerine bakılarak otomatik bir tespit yapılması çok da kolay değildir. Fakat görüntülerin aydınlık ortamda elde edildiğini düşünülürken, her binanın güneşin gelme açısına göre bir gölgesi oluşacaktır. Yani şöyle ki, her binanın bir gölgesi olacaktır fakat her gölgesi olan şey de bina değildir. Uydu görüntülerinin çok bantlı görüntüler olduğunu da düşünürsek gölge alanların bulunması nispeten daha kolay

olabilir, bu sebeple gölgeler nesne tespitinde kuvvetli bir yardımcı bilgi olarak kullanılabilir. Bu çalışmada gölge tespitinde Teke ve arkadaşları [13] tarafından önerilen yöntem kullanılmıştır.

III. BİNA TESPİTİ

A. *Unet* mimarisi

Uzaktan algılama uygulamalarında ve hiperspektral verilerin sınıflandırılmasında derin öğrenme algoritmaları son zamanlarda tercih edilmekte ve literatürdeki diğer yöntemlere nazaran düşük hata oranları ile tatmin edici performans göstermektedir [14,15]. Geleneksel görüntü tanıma yöntemlerinde ham veriden konunun uzmanları tarafından kompleks ve ayırt edici özellikler çıkarılırken, derin öğrenme modellerinde ise probleme özgü özellik çıkarma işlemi otomatikleştirilmektedir. Böylece düşük seviyeden yüksek seviyeye kadar hiyerarşik özellikler çıkarılarak makine öğrenmesi süreci tamamlanmış olur [16,17]. Bu çalışmada biyomedikal görüntülerde iyi sonuçlar vermiş *unet*[9] modelinden esinlenilerek derin öğrenme tabanlı bina tespit algoritması önerilmiştir. Tasarlanan *unet* mimarisi 224*224 boyutlarında girdi görüntülerini kabul etmektedir.

Veri kümesi iki farklı uydudan elde edilen 14 görüntüden oluşmaktadır. Eğitim kümesi oluşturulması için IKONOS uydusuna ait 4 görüntüden ikisi, Quickbird uydusuna ait 10 görüntüden dördü seçilmiştir. Ağ eğitimi için veri kümesindeki 1,2,5,6,7,8 numaralı görüntüler kullanılarak her bir görüntüden 224*224 boyutlarında 10 farklı alan seçilmiş ve 60 görüntüden oluşan bir eğitim seti hazırlanmıştır. Eğitim işleminde hata fonksiyonu olarak *dice* katsayısı kullanılmış ve eğitim neticesinde 0.8134 doğruluk değerine ulaşılmıştır.

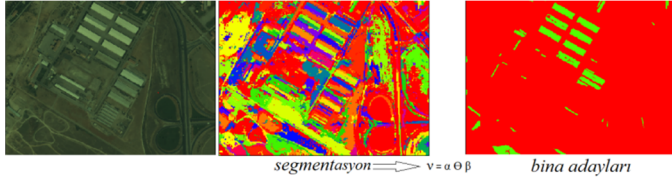
B. *Felzenszwalb* bölütleme gölge doğrulama

Derin öğrenme modeline ek olarak görüntüde fark edilmesi nispeten kolay ve keskin yapıdaki binaları tespit edebilmek için ek bir bölütleme modülü bina tespit sürecine dahil edilmiştir. Yardımcı bölütleme modülü geliştirilmesinde *felzenszwalb*[18] algoritması kullanılmaktadır. RGB görüntüler ilk olarak graf temelli *felzenszwalb* algoritması ile analiz edilir ve bölgeler çıkarılır. Bulunan her bir bölgenin şekil bakımından özellikleri incelenerek keskin hatlara (insan yapımı yapılar genelde keskin hatlara sahip olacağından) sahip bölgeler bina adayları olarak seçilir. *Felzenszwalb* ve şekil doğrulama ile elde edilen bina adaylarına son olarak gölge doğrulama işlemi uygulanır. Gölge doğrulama neticesinde elde edilen nihai sonuçlar, derin öğrenme modeliyle elde edilen sonuçlar ile birleştirilir.

Veri kümesindeki örnek bir görüntünün *felzenszwalb* algoritması ile elde edilen bölütleme sonucu şekil 2'de gösterilmiştir. Elde edilen her bir segment şekil bakımından incelenerek, keskin hatlara sahip olup olmamasına göre bina adayları belirlenir. Bir bölgenin bina adayları olarak seçilebilmesi için iki farklı kriter tanımlanır. Seçilen bölgeyi tamamen içerisine alacak bir konveks kapalı alan (H) tanımlanır ve seçilen bölgenin alanının (B_a) ilgili bölgeyi tamamen kapsayan konveks alanına (H_a) oranlanmasıyla 0-1 arasında bir değer elde edilir ($\alpha = B_a / H_a$). Bir başka kriter olarak ise, hedef bölgeyi içerisine alabilecek minimum büyüklükte bir dikdörtgen hesaplanarak, bulunan dikdörtgenin

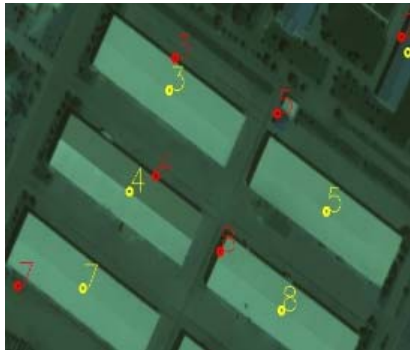
uzunluğu (b_l) ve genişliği (b_w) oranlanır ($\beta = b_l / b_w$). Böylece dairesel alanlar için 1 değerini verecek olan, 0-1 arasında bir katsayı elde edilir (1). Bu kriter ile uzun ve ince yapıdaki düzgün şekillerin (özellikle yolların) bina adayı olarak seçilebilmesinin önüne geçilmiştir. Her iki kriter tüm bölgelere uygulandıktan sonra bina adayları çıkarılmış olur. Denklem (1)'de formülde Θ operatörü her iki kriterin aynı anda yerine getirilmesini, yani kesişim işlemini ifade etmektedir.

$$\alpha = B_a / H_a, \beta = b_l / b_w, v = \alpha \Theta \beta \quad (1)$$



Şekil 2. Örnek bölütleme ve bina adaylarının bulunması

Bina adayları çıkarıldıktan sonra, her bir bina adayı için gölge doğrulama işlemi gerçekleştirilir. Gölge doğrulama süreci iki aşamadan oluşur. İlk olarak bina adayları kendisine en yakın gölgeler ile eşleştirilir. Bina adaylarını en yakın gölgeler ile eşleştirmek için $G_{yön}$ kriteri tanımlanır ve bu kriterle uyumlu alanlar eşleştirilir: $E(B, G) = d_{min} \cdot G_{yön}$. İşlem güneş açısına göre bina ve gölgelerin yön kontrolünü ifade eder. d_{min} ise her bir bina adayına en yakın uzaklıktaki gölgelerin bulunmasını ve aradaki minimum mesafeyi ifade eder. Şekil 3'de kırmızı renkli alanlar tespit edilen binaların merkez noktalarını, sarı renkler ise bina adaylarına en yakın gölge alanlarının merkez noktasını göstermektedir. Eşleştirilen bina ve gölge alanları ise aynı numaralarla ifade edilmektedir.



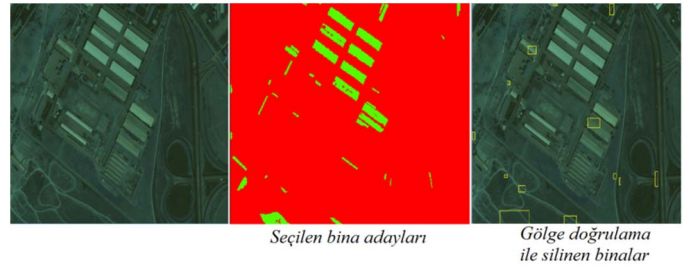
Şekil 3. Gölge yönü ($G_{yön}$) ve mesafeye (d_{min}) göre gölge-bina eşleştirilmesi

$E(B, G)$ ile gölge ve bina adayları eşleştirildikten sonra, seçilen bina adayının doğrulanması için iki kriter kullanılır. Gölge alanı (G_a) ile bina adayı alanları (B_a) karşılaştırılır ve eğer gölge alanı bina alanından büyükse ilgili bina adayı elenir. İkinci kriter olarak ise gölge ile bina adayı arasındaki mesafenin belirli bir eşik değerden küçük olup olmaması kontrol edilir. Böylece bina adayına yeterince yakın gölge bulunmaması durumunda ilgili bina adayı elenir. Burada kullanılacak eşik değerin belirlenmesi için bir yöntem önerilmiştir. Buna göre bina adayının içerisine alan minimum dikdörtgenin kısa kenar uzunluğu hesaplanarak gölge ile bina bölgesi uzaklığı arasında olması gereken eşik değeri belirlenir.

Şekil 4'de bulunan bazı hatalı bina adayları ve eşleştirildikleri gölgeler aynı sayı numaraları ile gösterilmektedir. Uygulanan gölge doğrulama kriterleri ile hatalı olarak seçilmiş 4, 5 ve 6 numaralı sarı renkli bina adayları elenmiştir.



Şekil 4. Hatalı bina adayları ve eşleştirildikleri gölge alanları

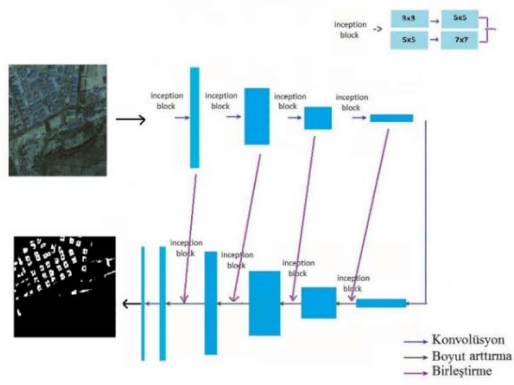


Şekil 5. Gölge doğrulama ile hatalı bina adaylarının elenmesi

Böylece ilk bölütleme sonucunda elde edilen tüm alanlar içerisinde, şekil bazı kontroller yapılarak ilk bina adaylarının belirlenmesi ve gölge doğrulama ile hatalı bina adaylarının elenmesi süreçlerinden oluşan bina tespit algoritması önerilmiştir. Uygulanan ek bölütleme yöntemi ile sadece şekil bakımından keskin hatlara sahip bina adayları gölge bilgisinin de yardımıyla tespit edilebilmektedir. Bu yöntem derin öğrenme ile birlikte kullanılarak önerilen bina tespit algoritmasına dahil edilmiş, ve her iki yöntem ile elde edilen sonuçlar birleştirilerek algoritma performansı artırılmıştır.

C. Inception-UNet mimarisi

Nesne tanıma amaçlı önerilen derin öğrenme mimarilerinden ResNet ile daha derin yapıda mimariler elde edilebilmesine olanak sağlanmıştır. Modelin derinliğinin yanısıra genişliğinin de artırılmasının hedeflendiği ağ mimarisine ise Inception diye adlandırılan blokların kullanıldığı mimari ile geçilmiştir. UNet mimarisine genişlik kazandırılması için (3x3-5x5) ve (5x5-7x7) boyutlarında paralel bloklar uygulanarak, Inception-UNet mimarisi önerilmiştir (Şekil-6). Mimaride görüntüyü sıkıştırma aşamalarında boyutların düşürülmesi için Inception bloklarının ardından ortaklama işlemi uygulanmıştır.



Şekil. 6. Inception-Unet mimarisi

IV. SONUÇLAR

İlk olarak Unet ile elde edilen sonuçlar, sırasıyla gölge maskesi tespiti, felzenswalb bölütleme, şekil doğrulama, gölge eşleştirme ve gölge doğrulama sonucu elde edilen sonuçlar ile birleştirilmiştir. Unet ile ek bölütleme modülünün bir arada kullanılması suretiyle elde edilmiş bina tespit sonuçlarına dair F_1 skorları Tablo 1’de Unet* ile gösterilmiştir. Sonuçlar incelendiğinde ek bölütleme ve gölge doğrulama işlemi ile Unet performansının 0.666 dan 0.70’e ulaştığı görülmektedir. Uygulanan felzenswalb ve gölge doğrulama temelli klasik yöntemde sadece keskin hatlardaki binalar tespit edilebilmektedir. Veri kümesi az sayıda görüntü içerdiğinden, Unet’in yetersiz eğitimden dolayı zayıf kaldığı noktalarda ek segmentasyon modülünün en azından keskin yapıdaki binaları tespit etmede yardımcı olduğu söylenebilir. Karmaşık yapıdaki şehir görüntülerinde ise çok aşamalı görüntü işleme teknikleri uygulanmak zorundadır. Unet ile klasik görüntü işleme temelli MPP ve Grabcut yöntemlerinden daha iyi sonuçlar elde edildiği gözlemlenmiştir.

TABLO I. KLASİK YÖNTEMLER İLE ÖNERİLEN YÖNTEMİN KİYASLANMASI

No	MPP	Grabcut	MLP	Unet	Unet*	Inception Unet
Ortalama F_1 skor	0,531	0,592	0,628	0,666	0.703	0.7420

Klasik yöntemler ile derin öğrenme yöntemleri kıyaslandığında, özellikle karmaşık yapıdaki şehir görüntülerinden özellik çıkarma konusunda derin öğrenmenin başarısı göze çarpmaktadır. Bu çalışmada önerilen Inception-Unet mimarisi ile Unet mimarisine genişlik kazandırılmış ve performans artışı sağlanmıştır. Inception-Unet ile ek segmentasyon modülü kullanmadan bile Unet’e nazaran çok daha iyi sonuçlar elde edilmesi önemlidir. Az sayıda görüntü kullanılarak eğitilen modellerin başarısı, uzmanlık ve detaylı analiz gerektiren, farklı görüntüler için değişkenlik gösterebilecek özellik çıkarma ve sınıflama süreçlerinin yerini alması bakımından önemlidir. Dolayısıyla bu çalışmada uygulanan klasik yöntemlerin karşılaştığı zorluklar da göz önüne alındığında, makine öğrenmesi temelli yöntemlerin bina bölütleme problemindeki başarısı açıkça görülmektedir. Bina

bölütleme konusunda derin öğrenme mimarilerinin geliştirilmesine yönelik çalışmalarımız devam etmektedir.

KAYNAKLAR

- [1] Lin, Chungan, and Ramakant Nevatia. "Building detection and description from a single intensity image." *Computer vision and image understanding* 72.2 : 101-121, 1998.
- [2] Izadi, Mohammad, and Parvaneh Saeedi. "Three-dimensional polygonal building model estimation from single satellite images." *IEEE Transactions on Geoscience and Remote Sensing* 50.6: 2254-2272, 2012.
- [3] Irvin, R. Bruce, and David M. McKeown. "Methods for exploiting the relationship between buildings and their shadows in aerial imagery." *OE/LASE'89, 15-20 Jan., Los Angeles, CA*. International Society for Optics and Photonics, 1989.
- [4] Kim, Taejung, and Jan-Peter Muller. "Development of a graph-based approach for building detection." *Image and Vision Computing* 17.1: 3-14, 1999.
- [5] Liow, Yuh-Tay, and Theo Pavlidis. "Use of shadows for extracting buildings in aerial images." *Computer Vision, Graphics, and Image Processing* 49.2: 242-277, 1990.
- [6] Sirmacek, Beril, and Cem Unsalan. "Building detection from aerial images using invariant color features and shadow information." *Computer and Information Sciences, 2008. ISICIS'08. 23rd International Symposium on*. IEEE, 2008.
- [7] Akçay, H. Gökhan, and Selim Aksoy. "Building detection using directional spatial constraints." *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*. IEEE, 2010.
- [8] Ok, Ali Ozgun. Automated detection of buildings from single VHR multispectral images using shadow information and graph cuts. *ISPRS journal of photogrammetry and remote sensing*, 86: 21-40, 2013.
- [9] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.
- [10] C. Benedek, X. Descombes, and J. Zerubia, "Building detection in a single remotely sensed image with a point process of rectangles," in *Proc. ICPR*, pp. 1417-1420, 2010.
- [11] Ok, Ali Ozgun, Caglar Senaras, and Baris Yuksel. "Automated detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery." *IEEE Transactions on Geoscience and Remote Sensing* 51.3: 1701-1717, 2013.
- [12] Kovacs, A.; Ok, A.O., "Building Detection From Monocular VHR Images by Integrated Urban Area Knowledge," *IEEE Geoscience and Remote Sensing Letters*, vol.12, no.10, pp.2140-2144, Oct. 2015.
- [13] Teke, M., Başeski, E., Ok, A. Ö., Yüksel, B., & Şenaras, Ç., Multi-spectral false color shadow detection. In *Photogrammetric Image Analysis* (pp. 109-119). Springer Berlin Heidelberg, 2011.
- [14] Chen, Y., Lin, Z., Zhao, X., Wang, G., Gu, Y.: Deep learning-based classification of hyperspectral data. *Selected Topics in Applied Earth Observations and Remote Sensing*, IEEE Journal of 7, 2094-2107, 2014.
- [15] Makantasis, K., Karantzas, K., Doulamis, A., Doulamis, N.: Deep Supervised Learning for Hyperspectral Data Classification through Convolutional Neural Networks. In: *IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2015)*, 2015.
- [16] Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H.: Greedy layer-wise training of deep networks. In: *Advances in NIPS 19*, 153-160, 2007.
- [17] Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. *Science* 313, 504-507, 2006.
- [18] P. F. Felzenswalb, D. P. Huttenlocher. *Efficient graph-based image segmentation*. *International Journal of Computer Vision* 59 (2) (2004) 167-181.