

# Data Augmentation Investigating in Improving the Performance of Deep Learning Model for Building Segmentation Using Aerial Images

Somaye Karimpour<sup>1\*</sup>, Javad Sadidi<sup>2</sup>  and Seyed Mohammad Tavakoli Sabour<sup>3</sup> 

1. M.Sc. Student, Department of Remote Sensing and Geographic Information Systems, Kharazmi University, Tehran, Iran
2. Associate Professor, Department of Remote Sensing and Geographic Information Systems, Kharazmi University, Tehran, Iran
3. Assistant Professor, Department of Remote Sensing and Geographic Information Systems, Kharazmi University, Tehran, Iran

\*Corresponding Author's E-mail: [sm.karimpour@yahoo.com](mailto:sm.karimpour@yahoo.com)

## Abstract

*Deep learning is a modern method of image processing and data analysis that has entered the field of urban management with promising results and high potential. The purpose of this study is to investigate data augmentation techniques in improving the results of segmentation of buildings using aerial images with high spatial resolution and deep learning method. For this purpose, MSB building data set and MapNet model were used. The model was trained and evaluated in three stages without data augmentation, with data augmentation of geometric transformations and with data augmentation of geometric and photometric transformations. The results of model evaluation showed that using geometric transformations as data enhancement techniques, F-1 and IoU score evaluation criteria have increased by 0.5 and 0.55%, respectively, and using data techniques Incremental geometric and photometric transformations increased by 1.41 and 1.57 percent. This increase was visually observed in the improvement of the segmentation of dense areas of the building and the discontinuity of large-scale buildings.*

**Keywords:** Photometric transformations, Geometric transformations, Aerial images, Data augmentation, Segmentation, Deep learning



## COPYRIGHTS

© 2022 by the authors. Published by Aerospace Research Institute. This article is an open access article distributed under the terms and conditions of [the Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

## How to cite this article:

S. Karimpour, J. Sadidi and S. M. Tavakoli Sabour, "Data Augmentation Investigating in Improving the Performance of Deep Learning Model for Building Segmentation Using Aerial Images," *Journal of Space Science and Technology*, Vol. 15, No. 3, pp. 23-32, 2022 (in Persian), <https://doi.org/10.30699/jsst.2022.1364>.

# بررسی داده‌افزایی در بهبود عملکرد مدل یادگیری عمیق جهت قطعه‌بندی ساختمان‌ها با استفاده از تصاویر هوایی

سمیه کریم پور<sup>۱\*</sup>، جواد سدیدی<sup>۲</sup> و سیدمحمد توکلی صبور<sup>۳</sup>

سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده علوم جغرافیایی، دانشگاه خوارزمی، تهران، ایران

\* ایمیل نویسنده مخاطب: [sm.karimpour@yahoo.com](mailto:sm.karimpour@yahoo.com)

## چکیده

یادگیری عمیق یک روش مدرن پردازش تصویر و تجزیه و تحلیل داده‌هاست که با داشتن نتایج امیدوارکننده و پتانسیل بالا، وارد حوزه مدیریت شهری شده است. هدف از تحقیق حاضر، بررسی تکنیک‌های داده‌افزایی در بهبود نتایج قطعه‌بندی ساختمان‌ها با استفاده از تصاویر هوایی با قدرت تفکیک مکانی بالا و روش یادگیری عمیق می‌باشد. برای این منظور از مجموعه داده ساختمان MSB و مدل MapNet/استفاده شد. مدل در سه مرحله بدون داده‌افزایی، با داده‌افزایی تبدیلات هندسی و با داده‌افزایی تبدیلات هندسی و فتومتریک مورد آموزش و ارزیابی قرار گرفت. نتایج ارزیابی مدل نشان داد که با استفاده از تکنیک‌های داده‌افزایی تبدیلات هندسی معیارهای ارزیابی امتیاز اف-یک و IoU به ترتیب به میزان ۰/۵ و ۰/۵۵ درصد و با استفاده از تکنیک‌های داده‌افزایی تبدیلات هندسی و فتومتریک به میزان ۱/۴۱ و ۱/۵۷ درصد افزایش پیدا کرد. این افزایش به صورت بصری در بهبود قطعه‌بندی مناطق متراکم ساختمان و ناپوستگی ساختمان‌های بزرگ مقیاس مشاهده شد.

واژه‌های کلیدی: تبدیلات فتومتریک، تبدیلات هندسی، تصاویر هوایی، داده‌افزایی، قطعه‌بندی، یادگیری عمیق

## مقدمه

یادگیری عمیق به دسته‌ای از شبکه‌های عصبی مصنوعی اشاره دارد که از بسیاری از لایه‌های پردازشی تشکیل شده است. نوآوری‌های الگوریتمی، افزایش توانایی‌های محاسباتی با استفاده از GPU و مجموعه‌های داده بزرگتر، عواملی هستند که به پیشرفت اخیر یادگیری عمیق کمک کردند. یادگیری عمیق یک روش جدید و مدرن برای پردازش تصویر و تجزیه و تحلیل داده‌هاست که با داشتن نتایج امیدوارکننده و پتانسیل بالا، وارد حوزه مدیریت شهری هم شده است و در بسیاری از زمینه‌ها مانند آشکارسازی تغییرات، مدیریت بحران، کاربری اراضی و پوشش زمین، برنامه‌ریزی شهری کاربرد دارد. اطلاعات ساختمان بخشی اساسی از برنامه‌های سیستم اطلاعات جغرافیایی (GIS) در برنامه‌ریزی و مدیریت شهری است که این اطلاعات با رشد اقتصادی به سرعت در حال تغییر است.

## علائم و اختصارات

Geographic Information System (GIS)	سیستم اطلاعات جغرافیایی
Unmanned Aerial Vehicle (UAV)	وسیله نقلیه هوایی بدون سرنشین
High Spatial Resolution (HSR)	تصاویر هوایی با قدرت تفکیک مکانی بالا
Digital Surface Model (DSM)	مدل‌های رقومی سطح
Light Detection and Ranging (LiDAR)	ردیابی نور و فاصله‌یابی
Interferometric Synthetic Aperture Radar (InSAR)	رادار دیافراگم مصنوعی
Deep Convolutional Neural Networks (DCNN)	شبکه‌های عصبی کانولوشن عمیق
Multi Scale Building Dataset (MSB)	داده چند مقیاس ساختمان
Intersection Over Union (IoU)	معیار استاندارد عملکرد سگمنتیشن

داده‌های لیدار و مدل‌های رقومی زمین و ... وجود دارد، توسعه روش‌های استخراج خودکار ویژگی‌های ساختمان از تصاویر RGB به سرعت در حال افزایش است [18].

با توجه به اینکه یکی از چالش‌های موجود در زمینه قطعه‌بندی ساختمان‌ها از تصاویر هوایی در مدل‌های یادگیری عمیق، عدم وجود داده‌های برچسب‌دار در مقیاس وسیع و با تنوع بالاست و با توجه به اینکه بهبود توانایی تعمیم این مدل‌ها یکی از دشوارترین چالش‌ها در این زمینه است [19]. لذا هدف این تحقیق، استفاده از تکنیک‌های داده‌افزایی<sup>۷</sup> به منظور افزایش دقت و همچنین افزایش قدرت تعمیم‌پذیری مدل با بهبود بیش‌برازش است که این بررسی، بر تکنیک‌های داده‌افزایی به عنوان یک راه حل در فضای داده برای حل مشکل داده‌های محدود، متمرکز است.

داده‌افزایی شامل مجموعه‌ای از تکنیک‌هاست که اندازه و کیفیت مجموعه داده‌های آموزشی را افزایش می‌دهد، به طوری که می‌توان با استفاده از آن مدل‌های یادگیری عمیق بهتری را ایجاد کرد [19]. به منظور کاهش بیش‌برازش و افزایش قدرت تعمیم‌پذیری مدل، تکنیک‌های مختلفی برای داده‌افزایی ایجاد شده است. به طور کلی می‌توان تکنیک‌های داده‌افزایی استفاده شده در قطعه‌بندی ساختمان را به دو دسته تبدیلات هندسی<sup>۸</sup> و تبدیلات فوتومتریک<sup>۹</sup> تقسیم‌بندی کرد. متداول‌ترین تکنیک‌های تبدیل هندسی استفاده شده شامل چرخش<sup>۱۰</sup> [4, 20, 18, 21, 22, 24] و وارونگی (افقی/عمودی) [22, 23, 25]، [20] است. سایر تکنیک‌های تبدیل هندسی شامل لرزش<sup>۱۲</sup> [25]، مقیاس<sup>۱۳</sup> و تغییرات مقیاس<sup>۱۴</sup> [4, 18]، برش [21, 4]، تغییرات سایز<sup>۱۶</sup> و انتقال معنا<sup>۱۷</sup> [21]، ترجمه<sup>۱۸</sup> [22]، بزرگنمایی<sup>۱۹</sup> [23] و تبدیلات انعکاسی<sup>۲۰</sup> [24] را نیز می‌توان در مطالعات پیشین مشاهده کرد.

علاوه بر تبدیلات هندسی به منظور استخراج ساختمان‌ها از تصاویر با قدرت تفکیک مکانی بالا، لیو و همکاران (۲۰۱۹) از تکنیک بهبود رنگ<sup>۲۱</sup> استفاده کردند و خوشبرش ماسوله و همکاران (۲۰۲۰) نیز از تکنیک تباین استفاده کردند. خوشبرش ماسوله و همکاران

فتوگرامتری مبتنی بر وسایل نقلیه هوایی بدون سرنشین (UAV) به دلیل اینکه قابلیت بروزرسانی سریع دارند و همچنین دارای قدرت تفکیک مکانی بالا می‌باشند برای این امر مناسب هستند [1]. با در دسترس بودن تصاویر هوایی با قدرت تفکیک مکانی بالا (HSR)، تقاضا برای روش‌های استخراج عوارضی مانند ساختمان‌ها، جاده‌ها و سایر ویژگی‌های ساخته شده توسط انسان از تصاویر هوایی روز به روز در حال افزایش است. از طرف دیگر روش دستی برای استخراج این عوارض بسیار پرهزینه و وقت‌گیر است. طی دو دهه گذشته، الگوریتم‌های زیادی برای استخراج ساختمان ارائه شده است [2].

برای طراحی عملگرهای ویژگی در روش‌های سنتی استخراج ساختمان، از ویژگی‌های طیف، بافت، هندسه و سایه به منظور قطعه‌بندی ساختمان‌ها از تصاویر نوری استفاده شده است [3-7]. این ویژگی‌ها تحت شرایط مختلف نور، انواع حسگر و معماری ساختمان متفاوت است و برای انواع مختلف ساختمان‌ها تعمیم خوبی ندارند. روش‌های سنتی فقط می‌توانند مسائل خاص مربوط به داده‌های خاص را حل کنند [8-11].

تصاویر نوری ترکیبی با GIS [12] مدل‌های رقومی سطح (DSM) حاصل از ردیابی نور و فاصله‌یابی (LiDAR) یا تداخل‌سنجی رادار دیافراگم مصنوعی (InSAR) [13] برای تشخیص مناطق غیر ساختمانی که شباهت زیادی به ساختمان‌ها دارند، دقت قطعه‌بندی ساختمان را افزایش می‌دهند، ولی در عین حال، این طیف وسیع داده‌ها که از چند منبع کسب می‌شوند همیشه پرهزینه هستند [2]. از آنجا که ساختمان‌ها در تصاویر سنجنش از دور از نظر ساختار، شکل ظاهری و مقیاس متنوع هستند، الگوریتم‌های قطعه‌بندی ساختمان از روش‌های مبتنی بر ویژگی‌های دستی به سمت یادگیری روش‌های خودکار استخراج ویژگی، مانند شبکه‌های عصبی کانولوشن عمیق (dCNN) توسعه یافتند [2]. روش خودکار برای قطعه‌بندی ساختمان از تصاویر هوایی به دلایل مختلفی مانند ساختار، اندازه و رنگ ساختمان‌ها، وضوح تصاویر و تغییر در تصاویر مناطق مختلف، یک کار چالش برانگیز بوده است. همچنین، تباین<sup>۴</sup> بین پشت‌بام‌های ساختمان و پس زمینه می‌تواند گاهی مشکلاتی ایجاد کند. با در نظر گرفتن همه این موانع، باز هم روش یادگیری عمیق توانسته است به نتایج خوبی در این زمینه دست یابد [13].

به‌رغم پیشرفت‌های سریع در علم اطلاعات جغرافیایی، قطعه بندی ساختمان در مقیاس بزرگ یک چالش اساسی در نقشه‌برداری این حوزه است [14-17]. با توجه به اینکه معمولاً نیاز به پوشش سطح وسیعی می‌باشد و در عین حال تمایل به صرف هزینه و بکارگیری نیروی انسانی کمتر در مقایسه با تصاویر استریو<sup>۵</sup> و

6. Red, Green, Blue  
7. Data augmentation  
8. Geometric Transformations  
9. Photometric Transformations  
10. Rotation  
11. Flip (right left/up down)  
12. Jittere  
13. Scaling  
14. Rescaling  
15. Cropping or Slicing  
16. Resizing  
17. Shifting  
18. Translating  
19. Zooming in  
20. Reflective Transformations  
21. Color Enhancement

4. Contrast  
5. Stereo

این مجموعه داده به سه قسمت داده‌های آموزش، اعتبارسنجی و آزمون به ترتیب به نسبت ۸۰: ۱۰: ۱۰ به صورت تصادفی تقسیم شدند. جزییات این مجموعه داده‌ها در جدول (۱) آمده است.

جدول ۱- تقسیم‌بندی مجموعه داده

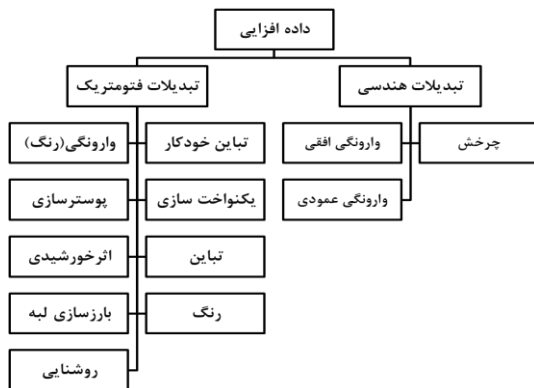
آموزش	اعتبارسنجی	آزمون
۲۳۲۰	۲۹۰	۲۹۰

## پیش پردازش داده‌ها

مجموعه داده قبل از عبور از مدل شبکه عصبی کانولوشنی، برداری‌سازی و سپس نرمال‌سازی شدند تا مقادیر پیکسل تمام تصاویر، در یک بازه مشخص قرار بگیرند. با توجه به اینکه تصاویر مورد استفاده ۸ بیتی هستند، نرمال‌سازی به صورت تقسیم هر تصویر بر ۲۵۵ انجام شد. علاوه بر این، از نرمال‌سازی دسته‌ای به صورتی که میانگین خروجی نزدیک به صفر و انحراف استاندارد خروجی نزدیک به یک باشد، استفاده شد.

## داده افزایی

به منظور بررسی تاثیرات داده‌افزایی بر نتایج، از تکنیک‌های داده‌افزایی به عنوان یک مرحله پیش‌پردازش استفاده شد تا شبکه در دامنه تغییرات این انتقالات نیز آموزش دیده و نتایج را بهبود ببخشد. تکنیک‌های داده‌افزایی مورد استفاده در شکل (۲) نشان داده شده است.



شکل ۲- تکنیک‌های داده‌افزایی مورد استفاده در تحقیق.

## معماری شبکه عصبی عمیق

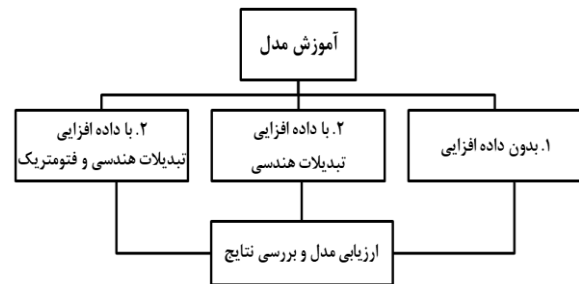
در این تحقیق از معماری MapNet [2] برای قطعه‌بندی ساختمان‌ها استفاده شد. این شبکه به طور کلی از سه مولفه به شرح زیر تشکیل شده است:

(۲۰۲۰) بیان کردند که طبق نتایج تجربی اشباع تصاویر با نرخ ۸۵ درصد به بهبود نتایج کمک می‌کند.

با توجه به نرخ رشد فضای شهری فعلی و تأثیرات آشکار ساخت و ساز ساختمان، از جمله در برنامه‌ریزی شهری، ارزیابی خسارت ساختمان و برآورد جمعیت، نظارت و اندازه‌گیری مداوم و همچنین اهمیت قطعه‌بندی ساختمان‌ها، هدف از انجام این تحقیق قطعه‌بندی ساختمان‌ها با استفاده از تصاویر هوایی با قدرت تفکیک مکانی بالا و روش یادگیری عمیق است که توانایی استخراج ویژگی‌ها را به صورت خودکار دارد و با حذف مرحله استخراج ویژگی به صورت دستی، باعث صرفه‌جویی در هزینه‌های مالی و زمانی شده و در عین حال نتایج قابل قبولی را نیز ارائه می‌دهد.

## مواد و روش‌ها

در این تحقیق مدل در سه مرحله مورد آموزش و ارزیابی قرار گرفت که روش کلی انجام تحقیق در شکل (۱) نشان داده شده است.



شکل ۱- فلوچارت کلی نحوه انجام تحقیق.

## مجموعه داده

در این تحقیق از مجموعه داده چند مقیاسه ساختمان MSB جهت آموزش و ارزیابی مدل استفاده شده است. این مجموعه داده توسط خوشبرش ماسوله و همکاران (۲۰۲۰) تهیه شده است که شامل تصاویر RGB از چندین مجموعه داده عمومی مختلف، با اندازه ۵۱۲ × ۵۱۲ پیکسل و قدرت تفکیک مکانی از ۰/۰۵ تا ۲/۵ متر تشکیل شده است. این مجموعه داده ۱۶ شهر مختلف (مانند ایندیاناپولیس، ماریون، تپتون، کلینتون، تایوان، اتاوا، ووهان، لس‌آنجلس، قاهره، میلان، سانتیاگو، ونیز، نیویورک، پوتسدام و بلینگهام) را پوشش می‌دهد. علاوه بر این، مجموعه داده تولید شده شامل تصاویر از حسگرهای مختلف ماهواره‌ای، از جمله WorldView-2، WorldView-3، QuickBird، ZiYun-3، IKONOS و تصاویر به دست آمده توسط حسگرهای هوایی است. داده‌های آموزش شامل تصاویر RGB چند مقیاس و برچسب‌های مربوط به ساختمان است که در آن‌ها پیکسل‌های ساختمان به صورت یک و پیکسل‌های غیر ساختمان به صورت صفر آورده شدند.

## اعتبارسنجی

از ماتریس ابهام (جدول خطا) برای ارزیابی مدل استفاده شد [26]. با استفاده از ماتریس ابهام، اعتبارسنجی با شمارش تعداد پیکسل‌های مثبت صحیح، مثبت ناصحیح، منفی صحیح و منفی ناصحیح محاسبه شد. بر اساس ترکیب چهار مورد فوق که در ماتریس ابهام محاسبه شد، دو معیار موثر در طبقه‌بندی باینری یعنی امتیاز اف - یک و IoU محاسبه شد. معادله‌های مربوط به هر معیار در جدول (۲) نشان داده شده است.

جدول ۲- معادله های ارزیابی مدل

معیار	معادله
اف-یک	$\frac{\text{صحيح مثبت} \times 2}{\text{ناصحيح منفي} + \text{ناصحيح مثبت} + (\text{صحيح مثبت} \times 2)}$
IoU	$\frac{\text{صحيح مثبت}}{\text{ناصحيح منفي} + \text{ناصحيح مثبت} + \text{صحيح مثبت}}$

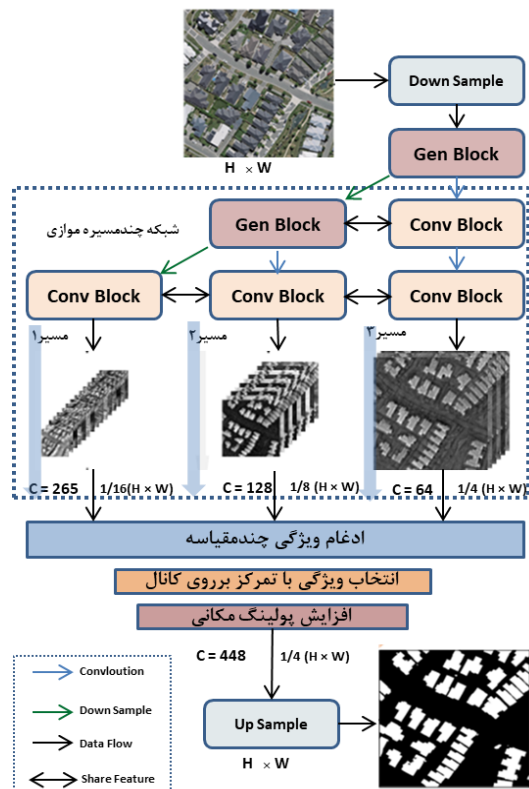
## نتایج و بحث

آموزش و ارزیابی مدل در سه مرحله مورد بررسی انجام شد که نتایج آن از نظر کمی و کیفی به شرح زیر است.

### نمودار تغییرات میزان خطا و میزان IoU دو مجموعه آموزش و اعتبارسنجی

برای ساختن مدل‌های مفید یادگیری عمیق، خطای اعتبارسنجی باید با خطای آموزش همچنان کاهش یابد. در شکل (۴) نمودارهای تغییرات میزان خطا (چپ) و میزان IOU (راست) دو مجموعه آموزش و اعتبارسنجی در حین آموزش مدل برای هر سه مرحله آموزش مورد بررسی یعنی مدل بدون داده‌افزایی (مرحله یک)، با داده‌افزایی تبدیلات هندسی (مرحله دو) و با داده‌افزایی تبدیلات هندسی و فتومتریک (مرحله سه) نشان داده شده است. شیب کاهش در نمودارهای تغییرات خطا بیان کننده این می‌باشد که میزان نرخ یادگیری (۰/۰۰۰۸) در آموزش مدل مناسب بوده است. همچنین شیب انتهایی به سمت صفر میل می‌کند که بیان می‌کند تعداد اپیاک (۲۰۰) و جهت آموزش مدل مناسب بوده است. کم بودن فاصله بین نمودارهای مجموعه آموزشی و اعتبارسنجی بیان کننده این موضوع می‌باشد که بیش‌برازش کمی در آموزش مدل رخ داده است. همانطور که در شکل (۴) شاهد هستیم در نمودار پرش‌هایی مشاهده می‌شود که یکی از دلایل وجود این پرش‌ها، عدم توازن بین مجموعه داده ورودی می‌باشد چرا که مجموعه داده مورد آموزش از ترکیب مجموعه داده‌های مختلف با تعداد و مقیاس‌های مختلف بوده است و این عدم یکنواختی باعث بروز پرش در نمودار شده است. از طرف دیگر معماری مدل نیز در بروز این پرش‌ها بی‌تأثیر نیست.

۱. یک شبکه چند مسیره موازی برای استخراج ویژگی‌های معنایی سطح بالای چند مقیاسه  
 ۲. یک فشرده‌ساز انطباقی ویژگی‌های چندمقیاسه چندمنظوره و ماژول پولینگ مکانی  
 ۳. یک ماژول افزایش نمونه‌گیری به سمت بالا بر پایه درونیایی به منظور استخراج ردپای ساختمان  
 بلوک Conv از مجموعه‌ای از ماژول‌های Residual برای استخراج ویژگی‌ها تشکیل شده است و بلوک Gen یک مسیر موازی جدید برای استخراج ویژگی‌های معنایی سطح بالاتر بر اساس بلوک Conv ایجاد می‌کند. شکل (۳) معماری این شبکه را نشان می‌دهد. این معماری دارای سه مسیر موازی به منظور استخراج ویژگی‌هاست و در نهایت مجموعه ویژگی‌های خروجی این سه مسیر باهم ادغام می‌شوند و وارد مرحله انتخاب ویژگی براساس کانال می‌شوند و پس از مرحله پولینگ مکانی با روش نمونه‌برداری به سمت بالا به سائز تصویر اولیه بازگردانده می‌شوند. این مدل به دلیل ارتباط متقابل بلوک‌های Conv و Gen و همچنین وجود مسیرهای موازی و مرحله انتخاب ویژگی‌ها بر اساس کانال، توانایی بسیار بالایی در استخراج ویژگی‌های سطح بالا دارد.



شکل ۳- ساختار شبکه MAP-Net [2]، H و W به ترتیب نشان دهنده تعداد کانال، ارتفاع و عرض تصویر ورودی هستند. Conv Block مجموعه‌ای از ماژول‌های Residual و Gen مسیر موازی جدید برای استخراج ویژگی‌های سطح بالاتر بر اساس Conv Block می‌باشد. Down sample و Up sample به ترتیب نمونه‌گیری رو به پایین و نمونه‌گیری رو به بالاست.

مرحله اول و سوم آموزش نشان می‌دهد که تعداد مثبت‌های صحیح ۴۴۴۴ عدد و منفی‌های صحیح ۱۳۱۷۱۵ عدد افزایش داشته است. همچنین در خصوص تغییرات مثبت‌ها و منفی‌های ناصحیح، در مرحله دوم، تعداد مثبت‌ها و منفی‌های ناصحیح به ترتیب به میزان ۱۰۲۱۶ و ۲۴۸۲۹ کاهش یافته است و در مرحله سوم آموزش، این مثبت‌های ناصحیح ۵۷۷۲ عدد افزایش و منفی‌های ناصحیح ۱۰۶۸۸۶ عدد کاهش یافته است.

### ارزیابی مدل

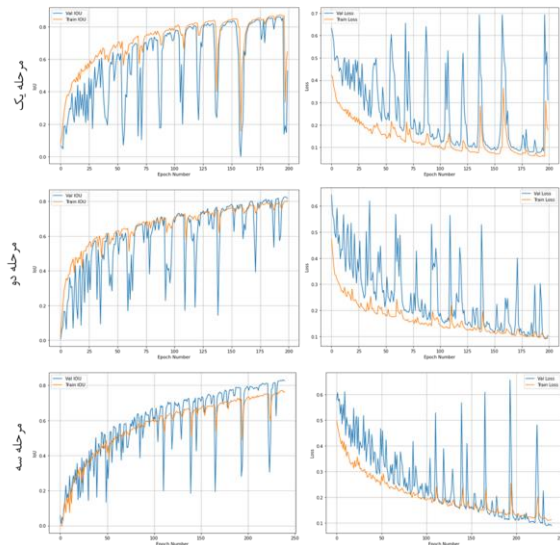
نتایج حاصل از قطعه‌بندی ساختمان در سه مرحله مورد بررسی برای مدل MapNet از نظر معیارهای امتیاز اف-یک و IoU در جدول (۴) نشان داده شده است.

جدول ۴- نتایج معیارهای ارزیابی مدل MapNet در سه مرحله

مرحله	مدل MapNet	امتیاز اف-یک %	تغییرات	IoU %	تغییرات
اول	بدون داده‌افزایی	۹۱/۳۹	-	۹۲/۴۶	-
دوم	با داده‌افزایی (تبدیلات هندسی)	۹۱/۹۴	+۰/۵	۹۲/۹۶	+۰/۵۵
سوم	با داده‌افزایی (تبدیلات هندسی و فتومتریک)	۹۳/۵۱	+۱/۴۱	۹۴/۳۷	+۱/۵۷

میزان معیار امتیاز اف-یک در مرحله اول، مدل MapNet بدون داده‌افزایی (۹۱/۳۹ درصد)، در مرحله دوم، مدل MapNet با داده‌افزایی تبدیلات هندسی (۹۱/۹۴ درصد) و در مرحله سوم، مدل MapNet با داده‌افزایی تبدیلات هندسی و فتومتریک (۹۳/۵۱ درصد) به ترتیب بالاترین تا کم‌ترین میزان را نشان داد.

بالاترین میزان معیار IoU در مدل MapNet مربوط به مرحله سوم آموزش مدل MapNet، با داده‌افزایی تبدیلات هندسی و فتومتریک (۹۴/۳۷ درصد) و مربوط به مرحله دوم آموزش مدل MapNet، با داده‌افزایی تبدیلات هندسی (۹۲/۹۶ درصد) و پایین‌ترین میزان مربوط به مرحله اول آموزش مدل MapNet، بدون داده‌افزایی (۹۲/۴۶ درصد) مشاهده شد. در مرحله دوم آموزش یعنی آموزش مدل با استفاده از داده‌افزایی تبدیلات هندسی نسبت به آموزش مدل بدون داده‌افزایی، میزان امتیاز اف-یک و IoU به ترتیب به میزان ۰/۵ و ۰/۵۵ درصد افزایش داشته است. همچنین در مقایسه مرحله دوم و سوم آموزش یعنی آموزش مدل با داده‌افزایی تبدیلات هندسی و آموزش مدل با داده‌افزایی تبدیلات هندسی و فتومتریک میزان امتیاز اف-یک و IoU به ترتیب ۱/۴۱ و ۱/۵۷ درصد افزایش را نشان می‌دهد. در مقایسه نتایج ارزیابی صحت مدل در مرحله اول و سوم یعنی آموزش مدل بدون داده‌افزایی و آموزش مدل با داده‌افزایی تبدیلات هندسی و فتومتریک میزان امتیاز اف-یک و IoU به ترتیب ۱/۹۱ و ۲/۱۲ درصد افزایش را نشان داد.



شکل ۴- نمودار تغییرات میزان خطا و IOU مجموعه آموزش و اعتبارسنجی بر اساس آپیک در سه مرحله مورد بررسی.

### ماتریس ابهام

ماتریس ابهام (جدول خطا)، ردیف‌ها و ستون‌هایی است که نمایانگر پیش‌بینی‌ها و نتایج واقعی (برچسب‌ها) برای یک طبقه‌بندی است که برای درک بهتر عملکرد مدل یا طبقه‌بندی بر اساس ارائه جواب صحیح استفاده می‌شود [26]. ماتریس ابهام در سه مرحله آموزش مورد بررسی در جدول (۳) آمده است.

جدول ۳- ماتریس ابهام مدل MapNet در سه مرحله مورد بررسی

مرحله سه		مرحله دو		مرحله یک		
N'	P'	N'	P'	N'	P'	
۲۷۲۹۶۰	۳۱۱۳۷۱۲	۲۶۷۱۸۸	۳۱۱۹۴۸۴	۲۷۷۴۰۴	۳۱۰۹۲۶۸	P
۱۵۵۲۰۵۱۰	۹۸۲۵۸	۱۵۴۱۳۶۲۴	۲۰۵۱۴۴	۱۵۳۸۸۷۹۵	۲۲۹۹۷۳	N
PP' : مثبت صحیح، NN' : منفی صحیح، PN' : مثبت ناصحیح، NP' : منفی ناصحیح.						

با توجه به جدول (۳)، مشاهده می‌شود که در مرحله دو آموزش، یعنی آموزش مدل با داده‌افزایی تبدیلات هندسی، میزان مثبت‌های صحیح و منفی‌های صحیح به ترتیب ۱۰۲۱۶ و ۲۴۸۲۹ عدد افزایش داشته است که نشان می‌دهد با استفاده از تبدیلات هندسی توانایی مدل در شناسایی هر دو نوع پیکسل‌های ساختمان و غیرساختمان بهبود داشته است و همچنین توانایی مدل در شناسایی پیکسل‌های غیرساختمانی خیلی بیشتر بوده است. همچنین مشاهده نتایج مرحله سوم آموزش، یعنی آموزش مدل با داده‌افزایی تبدیلات هندسی و فتومتریک نشان می‌دهد که میزان مثبت‌های صحیح ۵۷۷۲ عدد نسبت به مرحله دو آموزش کاهش داشته است و تعداد منفی‌های صحیح به تعداد ۱۰۶۸۸۶ عدد افزایش داشته است. مقایسه نتایج



پوشیده شده با پوشش گیاهی کمکی نکرده است درحالیکه استفاده از تبدیلات هندسی به تنهایی باعث بهبود این نتایج شده است.

### - مناطق متراکم ساختمان

مناطق متراکم ساختمان در تصاویر کوچک مقیاس قابل مشاهده خواهد بود. در شکل (۷) نمونه‌هایی از مناطق متراکم ساختمانی و همچنین نتایج پیش‌بینی مدل در سه مرحله مورد بررسی نشان داده شده است. همانطور که در شکل ۷ مشاهده می‌شود در مرحله اول یعنی در مرحله بدون داده‌افزایی، مدل توانایی کمی در تشخیص ساختمان‌های کوچک مقیاس دارد که در مرحله دو یعنی با استفاده از داده‌افزایی تبدیلات هندسی، مدل در شناسایی این ساختمان‌ها عملکرد بهتری پیدا کرده است. در مرحله سه یعنی با داده‌افزایی تبدیلات هندسی و فتومتریک نیز مدل عملکرد بهتری نسبت به دو مرحله قبل نشان داده است. بنابراین بهبود عملکرد مدل در شناسایی ساختمان‌های کوچک مقیاس طی این سه مرحله روند روبه رشدی داشته است.

### - ناپیوستگی در ساختمان‌های بزرگ مقیاس

بسیاری از اشیای زمینی (به‌عنوان مثال، ساختمان‌ها) که می‌توانیم آن‌ها را با تصاویر سنجنش از دور مشاهده کنیم، دارای ویژگی چند مقیاس هستند. این ویژگی باعث می‌شود مدل در ساختمان‌های بزرگ و کوچک مقیاس رفتارهای متفاوتی داشته باشد. یکی از مشکلات در خصوص ساختمان‌های بزرگ مقیاس، عدم پیوستگی در قطعه‌بندی سقف ساختمان‌ها می‌باشد. با بررسی نتایج بصری مشاهده شد که در مرحله دو یعنی تبدیلات هندسی توانسته است نسبت به مرحله یک (بدون داده‌افزایی) به بهبود مدل کمک کند و در مرحله سه (تبدیلات هندسی و فتومتریک) روند بهبود نسبت به مرحله دو افزایش داشته است. بنابراین تشخیص پیوستگی ساختمان‌ها در سه مرحله مورد بررسی روند روبه افزایشی را نشان می‌دهد. در شکل ۸ نمونه‌های بصری مرتبط با این مسئله نشان داده شده‌اند.

## بررسی بصری نتایج

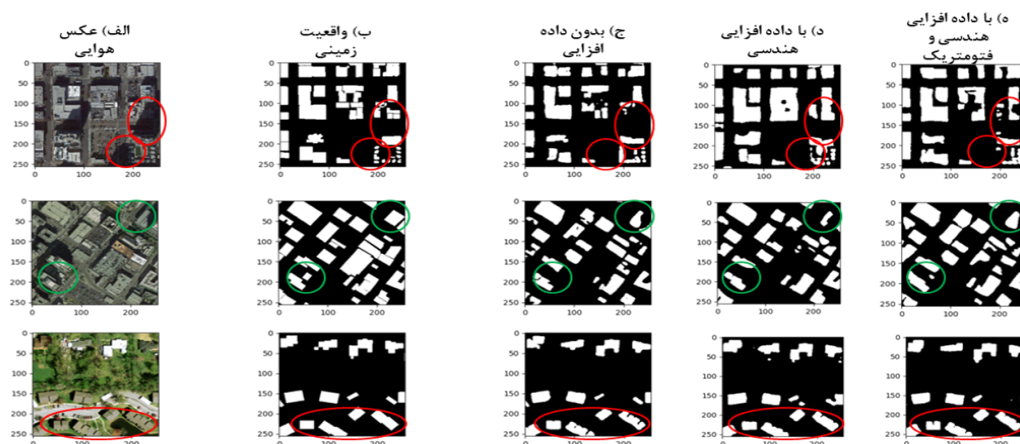
به‌منظور بررسی بصری نتایج حاصل از ارزیابی صحت مدل در سه مرحله آموزش مورد بررسی، نمونه‌های پیش‌بینی شده برای مجموعه داده آزمون بررسی شدند. همچنین ارزیابی جامعی مبتنی بر مهم‌ترین چالش‌های استخراج ساختمان در فتوگرامتری و سنجنش از دور، از جمله سایه‌ها، پوشش‌های گیاهی، مناطق متراکم ساختمان و ناپیوستگی در ساختمان‌های بزرگ مقیاس ارائه شدند.

### - سایه‌ها

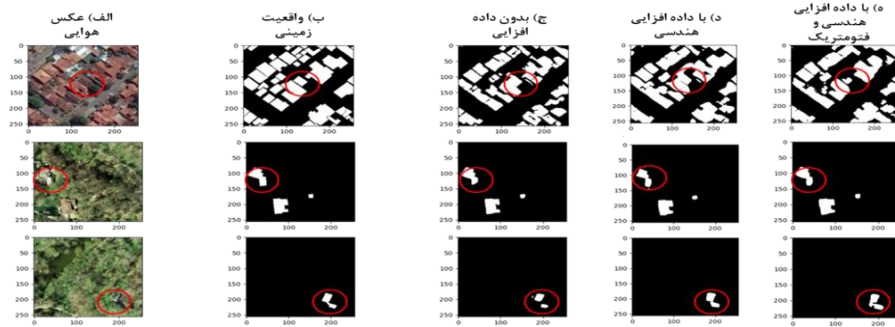
با توجه به نمونه‌های بصری که نتایج پیش‌بینی مدل را در سه مرحله آموزش مورد بررسی بر روی مجموعه داده آزمون نشان می‌دهد، در مورد سایه ساختمان‌های مرتفع که ساختمان‌های اطراف را می‌پوشاند، در مرحله دو نسبت به مرحله یک و سه بهبود عملکرد مدل را شاهد هستیم. در واقع ترکیب داده‌افزایی فتومتریک و هندسی باعث کاهش توانایی مدل در تشخیص این ساختمان‌ها شده است (شکل ۵، ردیف یک). در خصوص سقف ساختمان‌ها که سایه اجزای آن باعث عدم پیوستگی در قطعه‌بندی ساختمان می‌شود مشاهده شد که داده‌افزایی با تبدیلات فتومتریک باعث بارز شدن بیشتر این ناپیوستگی‌ها نسبت به تبدیلات هندسی شده است (شکل ۵، ردیف سه).

### - پوشش‌های گیاهی

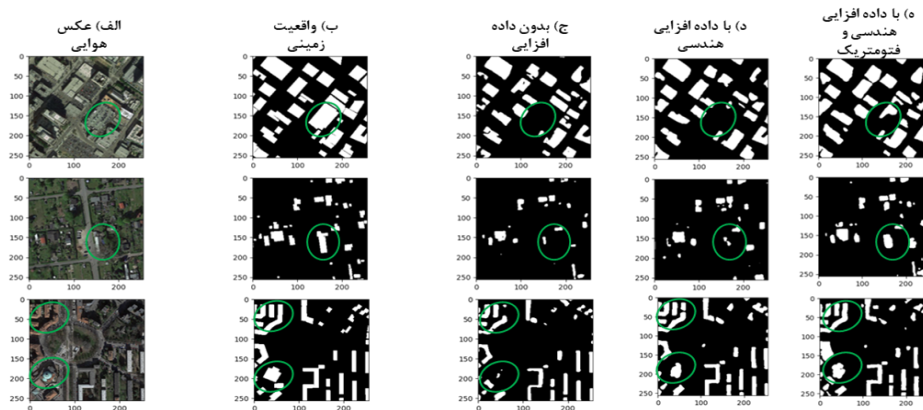
در نمونه‌های بصری حاصل از پیش‌بینی مدل در شکل (۶) مشاهده شد که در مرحله دو یعنی با استفاده از داده‌افزایی تبدیلات هندسی مدل توانسته است در برخی موارد قسمت‌هایی از ساختمان که با پوشش گیاهی پوشیده شده‌اند را شناسایی کند اما در مرحله سه یعنی با استفاده از ترکیب داده‌افزایی هندسی و فتومتریک، از توانایی مدل در تشخیص این مناطق کاسته شده است. بنابراین استفاده از ترکیب تبدیلات هندسی و فتومتریک به بهبود شناسایی ساختمان‌های



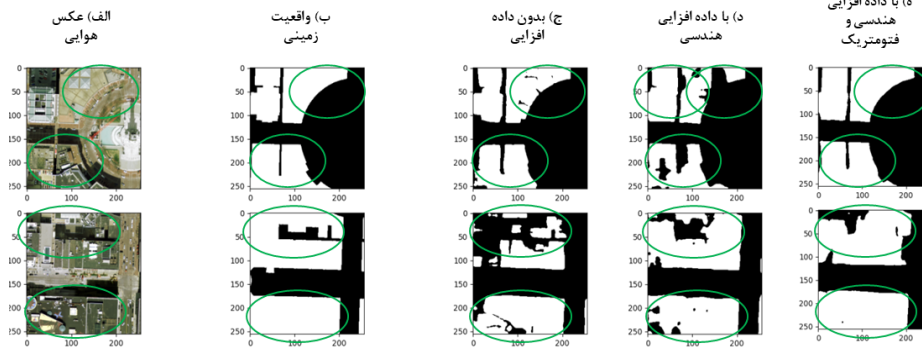
شکل ۵- سایه‌ها: نمونه‌هایی از نتایج ارزیابی مدل MapNet در سه مرحله آموزش مورد بررسی مجموعه آزمون. (الف) تصویر، (ب) واقعیت زمینی، (ج) نتیجه پیش‌بینی مدل بدون داده‌افزایی، (د) نتیجه پیش‌بینی مدل با داده‌افزایی تبدیلات هندسی و (ه) نتیجه پیش‌بینی مدل با داده‌افزایی تبدیلات هندسی و فتومتریک



شکل ۶- پوشش های گیاهی: نمونه هایی از نتایج ارزیابی مدل MapNet در سه مرحله آموزش مورد بررسی مجموعه آزمون. الف) تصویر، ب) واقعیت زمینی، ج) نتیجه پیش بینی مدل بدون داده افزایی، د) نتیجه پیش بینی مدل با داده افزایی تبدیلات هندسی و ه) نتیجه پیش بینی مدل با داده افزایی تبدیلات هندسی و فتومتریک



شکل ۷- مناطق مترکم ساختمان: نمونه هایی از نتایج ارزیابی مدل MapNet در سه مرحله آموزش مورد بررسی مجموعه آزمون. الف) تصویر، ب) واقعیت زمینی، ج) نتیجه پیش بینی مدل بدون داده افزایی، د) نتیجه پیش بینی مدل با داده افزایی تبدیلات هندسی و ه) نتیجه پیش بینی مدل با داده افزایی تبدیلات هندسی و فتومتریک



شکل ۸- ناپوستگی در ساختمان های بزرگ: نمونه هایی از نتایج ارزیابی مدل MapNet در سه مرحله آموزش مورد بررسی مجموعه آزمون. الف) تصویر، ب) واقعیت زمینی، ج) نتیجه پیش بینی مدل بدون داده افزایی، د) نتیجه پیش بینی مدل با داده افزایی تبدیلات هندسی و ه) نتیجه پیش بینی مدل با داده افزایی تبدیلات هندسی و فتومتریک

به ندرت در اختیار عموم قرار می گیرد و از طرف دیگر از آنجا که تهیه مجموعه داده با برچسب مورد نظر بسیار پرهزینه و زمان بر است یکی از چالش های اصلی در استفاده از روش یادگیری عمیق برای قطعه بندی ساختمان مجموعه داده بزرگ و استاندارد می باشد. چراکه عدم وجود مجموعه داده مناسب و کافی باعث مشکلاتی در آموزش مدل از جمله بیش برآزش و همچنین کاهش قدرت تعمیم پذیری مدل می گردد. لذا به منظور بهبود این مشکل در این تحقیق از تکنیک های داده افزایی استفاده شد. از آنجا که در اکثر پژوهش های مرتبط از تکنیک های داده افزایی تبدیلات هندسی استفاده شده بود در این تحقیق سعی شد علاوه بر تبدیلات هندسی از تبدیلات فتومتریک

## نتیجه گیری

اشیای شهری مانند ساختمان ها باید در فضاها و بازنمایی هایی با مقیاس های مختلف مورد تجزیه و تحلیل قرار بگیرند [27] بنابراین، به رغم روش های پیشرفته ای که سعی در استخراج رد پای ساختمان از یک مجموعه داده در یک مقیاس دارند، در این مطالعه از یک مجموعه داده جامع شامل تصاویر ماهواره ای و هوایی چند مقیاس از مکان های مختلف جغرافیایی و با قدرت تفکیک پذیری مکانی مختلف خارج از ایران، استفاده شد. از آنجا که مجموعه داده های هوایی و ماهواره ای با قدرت تفکیک پذیری مکانی بالا



- [8] M. Awrangjeb, C. Zhang, and C. S. Fraser, "Automatic extraction of building roofs using LIDAR data and multispectral imagery," *ISPR journal of photogrammetry and remote sensing*, vol. 83, p.p. 1–18, 2013.
- [9] D. Shouji, Y. Zhang, Z. Zou, S. Xu, X. He, and S. Chen, "Automatic building extraction from LIDAR data fusion of point and grid-based features," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, p.p. 294–307, 2017.
- [10] S. A. N. Gilani, M. Awrangjeb, and G. Lu, "An automatic building extraction and regularisation technique using lidar point cloud data and orthoimage," *Remote Sensing*, vol. 8, no. 3, p.p. 258, 2016.
- [11] S. Gunho, and I. Dowman, "Data fusion of igh-resolution satellite imagery and LIDAR data for automatic building extraction," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 62, no. 1, p.p. 43–63, 2007.
- [12] E. Simonetto, H. Oriot, and R. Garelo, "Rectangular building extraction from stereoscopic airborne radar images," *IEEE Transactions on Geoscience and remote Sensing*, vol. 43, no. 10, p.p. 2386–2395, 2005.
- [13] C. Chawda, J. Aghav, and S. Udar, "Extracting building footprints from satellite images using convolutional neural networks," *In 2018 International Conference on Advances in Computing, Communications and Informatics(ICACCI), IEEE*, p.p. 572-577, 2018.
- [14] Q. Bi, et al, "A multi-scale filtering building index for building extraction in very high-resolution satellite imagery," *Remote Sensing*, vol.11, no. 5, p. 482, 2019.
- [15] Z. Cai, H. Ma, and L. Zhang, "A building detection method based on semi-suppressed fuzzy C-means and restricted region growing using airborne LiDAR," *Remote Sensing*, vol. 11, no.7, p.p. 848, 2019.
- [16] F. Rottensteiner, et al., "Building detection by fusion of airborne laser scanner data and multi-spectral images: Performance evaluation and sensitivity analysis," *ISPRS J. Photogrammetry and Remote Sensing*, vol. 62, no. 2, p.p. 135–149, 2007.
- [17] T. T. Vu, F. Yamazaki, and M. Matsuoka, "Multi-scale solution for building extraction from LiDAR and image data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 11, no. 4, p.p. 281–289, 2009.
- [18] M. Khoshboresh-Masouleh, F. Alidoost, and H. Arefi. "Multiscale building segmentation based on deep learning for remote sensing RGB images from different sensors," *Journal of Applied Remote Sensing*, vol. 14, no. 3, 2020.
- [19] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *Jornal of Big Data*, vol. 6, no.1, p.p 1-48, 2019, DIO 10.1186/s40537-019-0197-0.

نیز استفاده شود. بنابراین این تحقیق در سه مرحله آموزش مدل بدون داده‌افزایی، با داده‌افزایی تبدیلات هندسی و با داده‌افزایی تبدیلات هندسی و فتومتریک انجام گرفت.

نتایج این بررسی نشان داد که با استفاده از داده‌افزایی تبدیلات هندسی و فتومتریک میزان امتیاز اف-یک ۱/۹۱ درصد و میزان معیار IoU ۲/۱۲ درصد افزایش پیدا کرده است. با توجه به نتایج بصری از لحاظ چهار چالش سایه، پوشش گیاهی، مناطق مترکم ساختمان و ناپیوستگی در ساختمان‌های بزرگ مقیاس مشخص شده‌است استفاده از تبدیلات هندسی و فتومتریک در شناسایی مناطق مترکم ساختمان و پیوستگی ساختمان‌های بزرگ مقیاس بهبود داشته است و در مورد سایه‌ها و هم‌چنین پوشش گیاهی بهبودی را نشان نداد.

## تعارض منافع

هیچگونه تعارض منافع توسط نویسندگان بیان نشده است.

## مراجع

- [1] W. Boonpooki, Y. Tan, and B. Xu. 2020. "Deep learning-based multi-feature semantic segmentation in building extraction from images of UAV photogrammetry". *International Journal of Remote Sensing*. DOI 10.1080/01431161.2020.1788742.
- [2] Zhu, Qing, C. Liao, H. Hu, X. Mei, and H. Li. "MAP-Net: Multiple Attending Path Neural Network for Building Footprint Extraction from Remote Sensed Imagery". *IEEE Transactions on geoscience and remote sensing*. <https://arxiv.org/abs/1910.12060>.
- [3] D. Jianli, D. Chen, R. Wang, J. Peethambaran, P. T. Mathiopoulous, L. Xie, and T. Yun, "A novel framework for 2.5-d building contouring from large-scale residential scenes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 6, p.p. 4121–4145, 2019.
- [4] L. Z hongbin, W. Shi, Q. Wang, and Z. Miao, "Extracting man-made objects from high spatial resolution remote sensing images via fast level set evolutions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 2, p.p. 883–899, 2014.
- [5] G. Nitin. L, and S. K. Ghosh, "Automatic building footprint extraction from high-resolution satellite image using mathematical morphology," *European Journal of Remote Sensing*, vol. 51, no. 1, p.p. 182–193, 2018.
- [6] J-P. Burochin, B. Vallet, M. Br'edif, C. Mallet, T. Brosset, and N. Paparoditis, "Detecting blind building fac,ades from highly overlapping wide angle aerial imagery," *ISPRS journal of photogrammetry and remote sensing*, vol. 96, p.p. 193–209, 2014.
- [7] M.Cote and P. Saeedi, "Automatic rooftop extraction in nadir aerial imagery of suburban regions using corners and variational level set evolution," *IEEE transactions on geoscience and remote sensing*, vol. 51, no. 1, p.p. 313–328, 2012.

- [22] J. Patterson, and A. Gibson, "Deep Learning, A Pr  
Actioner's Approach", *O'Reilly Media, Inc., 1005  
Gravenstein Highway North, Sebastopol, CA 95472. 532  
pages. <http://oreilly.com/safari>*
- [23] T. Lindeberg, "Scale-space theory: a basic tool for  
analyzing structures at different scales," *Journal of  
applied statistics*, vol. 21, no. 1-2, p.p 225-270,1994.
- [20] E. Chartock, L. Whitney, V. Singh, "Extraction of  
Building Footprints from Satellite Imagery".  
Stanford University Report, 2017.
- [21] S. Muruganandham, "Semantic Segmentation of  
Satellite Images using Deep Learning". Space  
Engineering, masters level. Luleå University of  
Technology, Department of Computer Science,  
Electrical and Space Engineering, 2016.