



ORIGINAL RESEARCH PAPER

Multi-Class 3D Change Detection in Urban Areas Using UNet++

A. Ebrahimi, M. Hasanlou*

School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

ABSTRACT

Received: 19 June 2025
Reviewed: 06 August 2025
Revised: 16 September 2025
Accepted: 10 January 2026

KEYWORDS:

Change Detection
Remote Sensing
Deep Learning
3D
CNN

* Corresponding author

✉ hasanlou@ut.ac.ir

☎ (+9821) 88008841

Background and Objectives: With the rapid expansion of urbanization, the need for automatic updating of change maps has become increasingly important. Accurate and up-to-date spatial information is essential for monitoring construction activities and tracking the development of urban areas. Traditional approaches to change detection are mostly limited to two-dimensional analysis and often lack sensitivity to vertical changes. This shortcoming fails to identify multi-story constructions, thereby limiting the completeness of monitoring outcomes. Recent advances in remote sensing and deep learning have enabled three-dimensional urban change detection, providing superior results compared to classical methods. This study aims to improve the performance of 3D urban change detection by introducing a deep learning approach that integrates multi-source data. The primary objective is to automatically identify and distinguish four types of building-related changes—new construction, complete demolition, height increase, and height decrease—alongside unchanged areas, to generate a comprehensive 3D change map.

Methods: The dataset employed in this research consists of high-resolution RGB aerial imagery and corresponding Digital Surface Model (DSM) data acquired from two different periods over Valladolid, Spain. The input data were prepared by stacking RGB images and DSMs from both epochs into an eight-band input, allowing the network to jointly analyze spectral and elevation information. The dataset was divided into training (90%) and testing (10%) subsets. To increase variability in the training data and reduce overfitting, augmentation techniques such as horizontal and vertical flipping, random rotation, and Gaussian blurring were applied. The proposed model architecture combines a ResNet-34 backbone for feature extraction with a UNet++ decoder for pixel-level change reconstruction. Model parameters were updated using the Adam optimizer. In the first stage, the deep network was trained in a binary setting (change/no-change) and evaluated against classical approaches, including Random Forest, image differencing/ratioing, and a PCA–K-Means hybrid method. In the second stage, the network was retrained for five-class classification, including the four change categories and the unchanged class, using a loss function optimized directly for the Intersection-over-Union (IoU) metric. Model performance was assessed using Accuracy, Recall, Precision, and F1-score.

Findings: In the binary classification stage, after 50 epochs of training, the network successfully identified most real changes while maintaining a low false alarm rate. Evaluation metrics confirmed this performance, with Recall and Accuracy both reaching 98.5% and an F1-score of 0.92, considerably outperforming the classical methods. Unlike traditional approaches, the deep learning model was able to detect almost all small-scale constructions and demolitions. In the five-class stage, the model effectively identified and classified change types, achieving a Recall of 96.32%, an Accuracy of 96%, and an F1-score of 0.95. All newly constructed and fully demolished buildings were correctly labeled in the output maps, and a large proportion of unchanged areas received no misclassification.

Conclusion: The findings demonstrate that combining elevation data with 2D imagery and leveraging deep learning architectures significantly mitigates the limitations of traditional change detection approaches and enhances accuracy. The developed model is capable of detecting not only the location but also the type of change. This approach has strong potential applications in monitoring unauthorized constructions, updating spatial databases, and assessing urban development. However, its effectiveness relies on the availability of accurate DSM data, which may not be consistently accessible for all urban areas. Additionally, the

training of deep networks requires extensive labeled datasets and considerable computational resources, which could limit their applicability in operational contexts.



NUMBER OF REFERENCES

48



NUMBER OF FIGURES

7



NUMBER OF TABLES

1

مقاله پژوهشی

آشکارسازی تغییرات سه بعدی در مناطق شهری با استفاده از مدل UNet++

علیرضا ابراهیمی، مهدی حسنلو*

دانشکده مهندسی نقشه برداری و اطلاعات مکانی، دانشکده گان فنی، دانشگاه تهران، تهران، ایران

چکیده

پیشینه و اهداف: امروزه با توسعه شهر نشینی، ضرورت به روزرسانی خودکار نقشه های تغییرات بیش از پیش اهمیت یافته است. به روزرسانی نقشه های مکانی، نظارت بر ساخت و ساز و پایش مناطق توسعه یافته نیازمند اطلاعات دقیق از بافت شهری و نقشه های جدید تغییرات می باشد. از آنجا که روش های سنتی پایش اغلب به تحلیل دو بعدی تغییرات پرداخته و حساسیت کافی در ثبت تغییرات بعد سوم یا ارتفاع را ندارند، لذا این محدودیت موجب عدم تشخیص ساخت و سازهای طبقه ای شده که در نتیجه آن، نقص اطلاعات در پایش را به همراه خواهد داشت. با پیشرفت در سنجش از دور و روش های نوین یادگیری عمیق، آشکارسازی تغییرات سه بعدی شهری امکان پذیر شده است که تاکنون نتایج برتری را نسبت به روش های کلاسیک ارائه داده اند. این پژوهش با هدف بهبود عملکرد در آشکارسازی تغییرات سه بعدی شهری، رویکردی بر مبنای یادگیری عمیق با استفاده از داده هایی با منابع مختلف ارائه می دهد تا انواع تغییرات ساختمانی به طور خودکار شناسایی و تفکیک نماید. هدف اصلی تفکیک چهار نوع تغییر ساختمانی از جمله ساخت جدید، تخریب کامل، افزایش ارتفاع و کاهش ارتفاع در کنار نواحی بدون تغییر می باشد تا نقشه جامعی از تغییرات سه بعدی شهری ارائه گردد.

روش ها: مجموعه داده مورد استفاده در این پژوهش شامل تصاویر هوایی رنگی (RGB) با وضوح مکانی بالا به همراه داده های مدل رقومی سطح (DSM) مربوط به دو بازه زمانی مختلف از محدوده شهری والادولید اسپانیا می باشد. داده ورودی شامل تصاویر رنگی قبل و بعد از تغییر به همراه مدل ارتفاعی متناظر آن به صورت یک ورودی هشت بانده با یکدیگر ترکیب شده تا شبکه به طور همزمان اطلاعات طیفی و ارتفاعی را مشاهده کند و جهت ورود به فرآیند آموزش مدل، داده ها به نسبت ۹۰ به ۱۰ درصد به دو دسته آموزش و آزمایش تقسیم بندی می شوند. جهت افزایش تنوع داده های آموزش و جلوگیری از برازش بیش از حد از روش هایی چون وارون سازی افقی و عمودی، چرخش تصادفی و تارای گاوسی بهره گرفته شده است. معماری مدل متشکل از یک شبکه ی عمیق ResNet-34 به عنوان بخش ویژگی یاب و از شبکه UNet++ جهت بازسازی پیکسل به پیکسل تغییرات می باشد. همچنین برای به روزرسانی پارامترهای مدل از الگوریتم Adam بهره گرفته شده است. در مرحله نخست مدل شبکه عمیق به صورت دودویی (تغییر/عدم تغییر) آموزش داده شده و عملکرد آن در مقایسه با روش کلاسیک جنگل تصادفی، تفاضل یا نسبت گیری و روش ترکیبی تحلیل مولفه های اصلی (PCA) و خوشه بندی K-Means مقایسه می شود. پس از آن شبکه با برچسب های پنج کلاسه شامل چهار نوع تغییر گفته شده و برچسب عدم تغییر مجدداً آموزش داده و جهت بهبود فرآیند یادگیری از یک تابع هزینه ساز بر مبنای بهینه سازی مستقیم معیار (IoU) استفاده شده است. جهت مقایسه عملکرد مدل ها از معیارهایی نظیر Accuracy، Recall، F1-score و Precision استفاده شده است.

یافته ها: در مرحله نخست این شبکه پس از ۵۰ اپک آموزش اغلب تغییرات واقعی را شناسایی کرده و در عین حال نرخ هشدار غلط را پایین نگه داشته است که معیارهای ارزیابی Accuracy، Recall، F1-score و مقادیر ۹۸.۵٪ و ۹۸.۵٪ و ۰.۹۲ این نتایج را تأکید می کنند و این ارقام به طور محسوسی بهتر از روش های کلاسیک بوده. همچنین در مدل یادگیری عمیق بر خلاف دیگر روش ها تقریباً تمامی موارد ساخت و ساز یا تخریب در مقیاس های کوچک تشخیص داده شده است. در مرحله پنج کلاسه، تغییرات به خوبی توسط مدل شناسایی و طبقه بندی شده اند و معیارهای Recall،

تاریخ دریافت: ۲۹ خرداد ۱۴۰۴
تاریخ داوری: ۱۵ مرداد ۱۴۰۴
تاریخ اصلاح: ۲۵ شهریور ۱۴۰۴
تاریخ پذیرش: ۲۰ دی ۱۴۰۴

واژگان کلیدی:

آشکارسازی تغییرات
سنجش از دور
یادگیری عمیق
سه بعدی
شبکه عصبی پیچشی

* نویسنده مسئول

hasanlou@ut.ac.ir

۰۲۱-۸۸۰۰۸۸۴۱

Accuracy و F1-score به ترتیب برابر با ۹۶.۳۲٪ و ۹۶٪ و ۰.۹۵ می‌باشد. تمامی سازه‌های جدید احداث شده و ساختمان‌های کاملاً تخریب شده در نقشه خروجی مدل با برچسب درست تشخیص داده شده اند و بخش وسیعی از مناطق بدون تغییر هیچ برچسب اشتباهی دریافت نکرده‌اند.

نتیجه‌گیری: نتایج نشان‌دهنده آن است که ترکیب داده‌های ارتفاعی با تصاویر دو بعدی و به کارگیری مدل‌های یادگیری عمیق به طرز چشمگیری کاستی‌های روش سنتی تشخیص تغییرات را برطرف کرده و افزایش دقت را به همراه خواهد داشت. این مدل توسعه یافته قادر است که نه تنها مکان تغییر بلکه نوع تغییر را شناسایی کند. لازم به ذکر است که از این شیوه نوین می‌توان جهت پایش ساخت و سازهای غیر مجاز، به روز رسانی پایگاه‌های داده مکانی و ارزیابی تغییرات شهری استفاده شود. مهمترین محدودیت این پژوهش وابستگی به داده‌های ارتفاعی دقیق می‌باشد که تهیه منظم آن در تمامی شهرها امکان پذیر نمی‌باشد. به علاوه آموزش مدل‌های عمیق نیازمند داده‌های برچسب دار گسترده و توان پردازشی بالا می‌باشد که ممکن است در کاربردهای عملی محدودیت ایجاد کند.

مقدمه

سهم‌بندی لازم را جهت تشخیص تغییرات سهم‌بندی فراهم کند [۵]. به کمک اطلاعات سهم‌بندی می‌توان عوارضی نظیر ساختمان‌ها را از سطح زمین تفکیک کرده و تغییرات قائم را شناسایی نمود. ادغام تصاویر هوایی با داده‌های ارتفاعی به طرز چشم‌گیری دقت شناسایی ساختمان‌های جدید و سایر تغییرات سازه‌ای همچون کاهش یا افزایش ارتفاع را افزایش داده است. بدین ترتیب، تلفیق داده‌های طیفی و اطلاعات ارتفاعی این امکان را ایجاد می‌کند که نه تنها وقوع تغییر در یک مکان مشخص، شناسایی بلکه ماهیت سهم‌بندی تغییر نیز (از نظر افزایش یا کاهش ارتفاع ساختمان) مشخص گردد [۶]. این رویکرد نسبت به روش‌های دوبعدی رایج، تصویری با اطلاعات بیشتر از تحولات شهری ارائه می‌دهد و قابلیت اطمینان و دقت نتایج را بهبود می‌بخشد [۷].

همزمان با پیشرفت در بهره‌برداری از داده‌های چندمنبعی، ظهور روش‌های یادگیری عمیق تحول شگرفی در حوزه تحلیل‌های مکانی و سنجش‌ازدور به‌وجود آورده است [۸]. در سال‌های اخیر، رویکردهای مبتنی بر شبکه‌های عصبی عمیق، به‌ویژه شبکه‌های عصبی پیچشی (CNN)، جایگاه اصلی را در روش‌های تشخیص تغییرات به خود اختصاص داده است. این شبکه‌ها با یادگیری الگوهای پیچیده از داده‌های تصویری، ویژگی‌های مهم تصویر را پیدا کرده و دقت آشکارسازی تغییرات را به میزان قابل توجهی نسبت به روش‌های سنتی افزایش داده‌اند [۹]. به عنوان نمونه، به‌کارگیری مدل‌های قطعه‌بندی مانند UNet در تصاویر هوایی چندزمانه نشان داده است که شبکه‌های عمیق قادر به شناسایی دقیق نواحی تغییر یافته در ساختمان‌ها هستند و عملکردی به مراتب بهتر از روش‌های سنتی مبتنی بر آستانه‌گذاری یا مقایسه تصاویر دارند [۱۰، ۱۱]. از اواسط دهه ۲۰۱۰ میلادی به دنبال پیشرفت معماری‌های CNN و در دسترس بودن داده‌های با وضوح مکانی بالا، پژوهش‌های بسیاری در زمینه سنجش‌ازدور مبتنی بر یادگیری عمیق انجام شده است که امکان تحلیل خودکار تصاویر شهری را با جزئیات و دقت بالایی فراهم کرده‌اند [۸، ۹]. در نتیجه این پیشرفت‌ها، اکنون ابزارهای نوینی در اختیار محققان و مدیران شهری قرار گرفته که می‌توانند تغییرات را نه تنها سریع‌تر بلکه با دقت مکانی بالاتر و در بعد سوم (ارتفاع) شناسایی کنند [۱۲].

امروزه آشکارسازی تغییرات در مناطق شهری یکی از نیازهای اساسی مدیریت شهری می‌باشد که در زمینه‌هایی مانند نظارت بر ساخت‌وسازها، به‌روزرسانی نقشه‌ها، ارزیابی خسارات و تصمیمات مدیریتی نقشی حیاتی دارد. از آنجا که به‌روزرسانی تغییرات شهری فرایندی زمان‌بر و پرهزینه است، خودکارسازی تشخیص تغییرات ساختمانی نقش مهمی در برنامه‌ریزی شهری و مدیریت کاربری زمین دارد. هدف این فرآیند شناسایی ساخت‌وسازهای جدید، تخریب ساختمان‌ها یا تغییرات دیگر در سازه‌های شهری از طریق مقایسه تصاویر چندزمانه می‌باشد [۱]. با توجه به رشد بالای شهرنشینی و قرار داشتن بسیاری از کشورها در پهنه‌های زلزله خیز یا سایر حوادث طبیعی، اهمیت پایش تغییرات ساختمانی در سامانه‌های اطلاعات مکانی و نظارت مبتنی بر سنجش از دور به یکی از دغدغه‌های اساسی مدیریت شهری تبدیل شده است. در این میان، بهره‌گیری از تصاویر هوایی و ماهواره‌ای با توان تفکیک مکانی بالا، امکان تحلیل دقیق‌تر جزئیات تغییرات شهری همچون تغییرات سهم‌بندی را فراهم کرده است که این موضوع به افزایش دقت و کارایی نظارت خودکار منجر می‌شود [۲].

روش‌های سنتی آشکارسازی تغییرات، نظیر تفاضل یا نسبت‌گیری دو تصویر معمولاً تغییرات دو بعدی را شناسایی کرده و در تشخیص تغییرات حجمی و ارتفاعی با محدودیت‌های جدی مواجه‌اند [۳]. در پایش توسعه شهری، صرف تشخیص یک تغییر در نقشه‌ی دوبعدی کافی نیست و باید مشخص شود که آیا ساختمانی جدید احداث شده است یا یک ساختمان موجود افزایش ارتفاع داشته یا تخریب شده است؛ چنین اطلاعاتی با رویکردهای صرفاً دوبعدی به‌دست نمی‌آید. از آنجاکه نقشه‌های تغییر حاصل از این روش‌ها عمدتاً محدود به تغییرات دو بعدی هستند و اطلاعاتی در مورد تغییرات ارتفاعی سازه‌ها ارائه نمی‌دهند بنابراین در تحلیل‌های حجمی یا توپوگرافی (مانند پایش توسعه عمودی شهرها، برآورد تغییرات ارتفاع ساختمان‌ها یا بررسی حجم آوار و خاک‌برداری) مورد استفاده نمی‌باشد [۴].

برای رفع این محدودیت‌ها، بهره‌گیری همزمان از داده‌های ارتفاعی در کنار اطلاعات دوبعدی مکانی پیشنهاد شده است. ترکیب مدل رقومی ارتفاع (DEM) یا مدل رقومی سطح (DSM) با تصاویر رنگی می‌تواند دید

بر گرفته و با تفکیک بین انواع مختلف سازه‌ها، دید جامعی نسبت به تحولات شهری ارائه می‌دهد [۱۷]. برای شناسایی الگوهای پیچیده‌ی تغییرات، یک چارچوب یادگیری عمیق به کار گرفته شده است که معماری ResNet-34 را با شبکه قطعه‌بندی پیشرفته UNet++ تلفیق می‌کند. استفاده از ساختار ResNet-34 به شبکه امکان یادگیری مؤثر ویژگی‌های عمیق و آشکارسازی تغییرات را می‌دهد و معماری UNet++ با بهره‌گیری از مسیرهای چندمقیاسه و اتصالات پرشی، امکان بازسازی دقیق نقشه‌های تغییر با حفظ جزئیات مکانی را فراهم می‌کند [۱۸، ۱۹]. این شبکه با دریافت همزمان اطلاعات طیفی و ارتفاعی، نقشه تغییرات ساختمانی را به صورت پیکسلی و در قالب چهار کلاس مذکور تولید می‌کند و می‌تواند به بهبود مدیریت شهری کمک شایانی نماید. نتایج حاصل از این پژوهش، اطلاعات ارزشمندی درباره میزان و نوع ساخت‌وسازها یا تخریب‌های صورت گرفته در شهر ارائه می‌دهد که در برنامه‌ریزی توسعه‌ی شهری، ارزیابی ایمنی سازه‌ها و مدیریت بحران (نظیر برآورد خسارات بعد از حوادث) کاربرد دارد. در نهایت، این دستاوردها نشان‌دهنده پتانسیل الگوریتم‌های یادگیری عمیق و تلفیق داده‌های چندبعدی در بهبود دقت و جزئیات تحلیل تغییرات مکانی در محیط‌های شهری است [۲۰].

روش تحقیق

جهت انجام قطعه‌بندی با دقت بالا از ترکیب داده‌های تصویری رنگی و مدل رقومی ارتفاع استفاده شده است که از مجموعه داده 3DCD به دست آمده است. این مجموعه داده شامل تصاویر طیفی و همچنین مدل رقومی زمین در دو تاریخ ۲۰۱۰ و ۲۰۱۷ با ابعاد 400×400 پیکسل و قدرت تفکیک مکانی ۰.۵ متر می‌باشد [۴۸]. بمنظور آماده‌سازی مجموعه داده، ابتدا تصاویر رنگی و اطلاعات ارتفاعی متناظر با یکدیگر برای دو زمان قبل و بعد با یکدیگر ترکیب شده و به صورت یک ورودی هشت کاناله در آمده‌اند؛ به عبارتی برای هر زمان باندهای قرمز، آبی، سبز به همراه لایه ارتفاع روی یکدیگر قرار گرفته و در نهایت اطلاعات زمان قبل و بعد جهت ورود به مدل با یکدیگر ترکیب و به صورت یک داده یکپارچه در آمده‌اند. این ادغام موجب می‌شود که شبکه عصبی به‌طور همزمان از ویژگی‌های بصری و ارتفاعی برای درک بهتر ساختارهای موجود در تصویر استفاده کند [۲۱]. برای آماده‌سازی داده‌ها جهت ورود به مدل، تمامی تصاویر به ابعاد 512×512 پیکسل بازنمونه‌گیری شده‌اند و جهت بهبود تعمیم‌پذیری مدل و جلوگیری از بیش‌برازش، از روش‌هایی چون وارون سازی افقی و عمودی، چرخش تصادفی و تارگی گوسی بر روی داده‌های آموزشی استفاده شده است که تنوع مجموعه داده را افزایش داده و مدل را نسبت به تغییرات مختلف در داده‌ها مقاوم‌تر می‌کنند.

ونگ و همکاران [۱۳] یک معماری جدید به نام FRCD-Net برای آشکارسازی تغییرات در تصاویر سنجش‌ازدور معرفی می‌کند. هدف اصلی این پژوهش، بهبود دقت در آشکارسازی از طریق تلفیق مؤثر ویژگی‌ها و تقویت جزئیات مکانی است. در طراحی این شبکه، از رویه‌هایی برای بازنمایی بهتر ویژگی‌های با سطح بالا و پایین استفاده شده تا اطلاعات متنی و فضایی به خوبی ترکیب شوند. نتایج تجربی بر روی چندین مجموعه داده، نشان می‌دهد که FRCD در مقایسه با مدل‌های پایه و دیگر روش‌های یادگیری عمیق، عملکرد بهتری در تفکیک نواحی تغییر کرده از ثابت ارائه می‌دهد. این پیشرفت به‌ویژه در آشکارسازی تغییرات ظریف و مناطق با نویز زیاد مشهود است. مشکینی و همکاران [۱۴] مدلی مبتنی بر 3D CNN پیشنهاد داده‌اند که توانایی استخراج ویژگی‌های زمانی-مکانی را از تصاویر ماهواره‌ای با وضوح بالا دارد. این مدل از یک شبکه‌ی 2D CNN از پیش‌آمخته‌شده به عنوان استخراج‌کننده ویژگی پایه استفاده کرده و آن را با ساختار سه‌بعدی تلفیق می‌کند تا تغییرات در طول زمان به‌صورت دقیق‌تری شناسایی شود. ترکیب بین ویژگی‌های مکانی از 2D CNN و ویژگی‌های زمانی از طریق ساختار سه‌بعدی موجب بهبود قابل توجهی در عملکرد مدل به‌ویژه در داده‌هایی که شامل تغییرات تدریجی هستند، می‌شود.

در پژوهشی که توسط هوآنگ [۱۰] صورت گرفته است از یک شبکه‌ی دوقلو (Siamese) با ترکیب معماری‌های EfficientNet و MANet برای آشکارسازی ساختمانی در تصاویر با توان تفکیک مکانی متوسط بهره‌گیری شده است. استفاده از معماری EfficientNet باعث کاهش بار محاسباتی و افزایش بازدهی شده و در کنار آن، MANet به پردازش بهتر نواحی مهم تصویر کمک می‌کند. این مدل با پردازش جفت تصاویر قبل و بعد از تغییرات، تفاوت‌های ساختاری را شناسایی کرده و با کمک مکانیزم‌های قابل توجه، نواحی تغییر یافته را با دقت بالا مشخص می‌کند. نتایج آزمایش‌ها نشان‌دهنده برتری این روش در مقابل بسیاری از مدل‌های مشابه در زمینه‌ی تشخیص تغییرات ساختمانی است.

با توجه به توضیحات ذکرشده، هدف این پژوهش رویکرد جدید برای آشکارسازی تغییرات سه‌بعدی در بافت شهری می‌باشد که از ترکیب داده‌های تصویری و ارتفاعی با بهره‌گیری از روش یادگیری عمیق استفاده شده است. داده‌های مورد استفاده شامل تصاویر RGB با وضوح مکانی بالا و مدل رقومی ارتفاع از شهر والادولید اسپانیا در دو زمان متفاوت است که امکان مقایسه دقیق ساختار سه‌بعدی شهری را فراهم می‌آورد [۱۵، ۱۶]. بر خلاف بسیاری از مطالعات پیشین که صرفاً به آشکارسازی وجود یا عدم وجود تغییر بسنده می‌کنند، روش پیشنهادی تغییرات را به چهار دسته (۱) ساخت‌وساز جدید، (۲) افزایش ارتفاع ساختمان، (۳) تخریب کامل و (۴) کاهش ارتفاع ساختمان دسته‌بندی می‌کند. این دسته‌بندی به‌طور ویژه تغییرات حجمی ساختمان‌ها را در



شکل ۱: منطقه مورد مطالعه

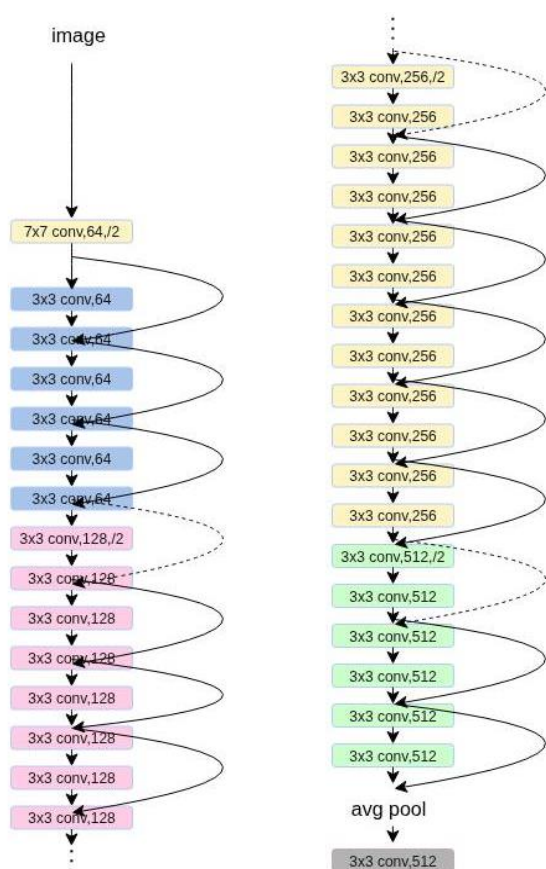
Fig. 1: Study area

فرآیند یادگیری ایجاد کند، اتصال میان بر تضمین می‌کند که خروجی آن بلوک حداقل می‌تواند همان ورودی (بدون تغییر) باشد و از این رو لایه‌های بیشتر به کاهش دقت مدل منجر نشوند. این راهبرد به ResNet اجازه می‌دهد تا شبکه‌هایی با عمق بسیار زیاد (حتی صدها لایه) را با موفقیت آموزش دهد [۲۶].

وزن‌های اولیه‌ی شبکه به صورت تصادفی مقدار دهی شده است و به منظور دستیابی به دقت بالاتر و اطمینان از آموزش اختصاصی بر روی داده‌های پژوهش از مدل‌های از پیش آموزش یافته استفاده نشده است [۲۲، ۲۳]. در معماری مدل، ساختار ResNet-34 به عنوان کدگذار (encoder) و شبکه UNet++ به عنوان بخش رمزگشا (decoder) تعریف شده است. این معماری به شبکه امکان می‌دهد که از اتصالات میان‌بر (skip connection) برای حفظ جزئیات در سطوح مختلف استفاده کند و با بلوک‌های باقیمانده (Residual Blocks) بدون آنکه مشکل محو شدن گرادیان به وجود آید به عمق زیاد دست یابد [۲۴]. برای تابع هزینه، تابع Lovasz-softmax استفاده شد تا معیار IoU به طور مستقیم بهینه‌سازی شود. همچنین بهینه‌سازی وزن‌های شبکه توسط الگوریتم Adam انجام گرفت که به دلیل سازوکار تطبیقی خود برای یادگیری نرخ‌ها در مسائل یادگیری عمیق کارایی بالایی دارد [۲۵]. برای ارزیابی عملکرد مدل از ماتریس ابهام (Confusion Matrix) و مشتقات آن شامل Accuracy، Recall، و F1-score استفاده شده است که در ادامه به توضیح هر یک از اجزای مذکور پرداخته شده است.

معماری شبکه‌های ResNet و ResNet-34

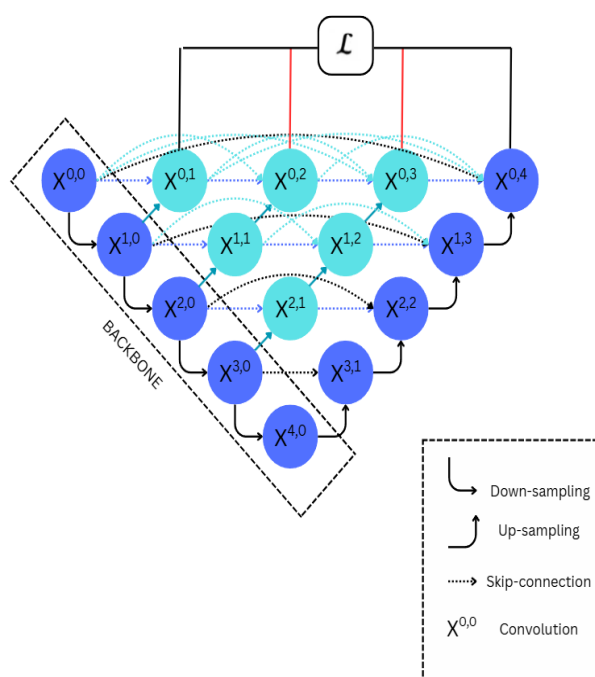
ResNet یا Residual Networks خانواده‌ای از شبکه‌های عصبی پیچشی بسیار عمیق هستند که با معرفی مفهوم بلوک‌های باقیمانده، تحول مهمی در یادگیری عمیق ایجاد کرده‌اند. ایده‌ی اصلی ResNet افزودن اتصالات میان‌بر (Skip Connections) بین لایه‌هاست؛ به طوری که خروجی یک لایه عمیق‌تر با خروجی چند لایه قبل‌تر جهت ورود مجدد به مدل جمع می‌شود. این سازوکار باعث می‌شود که شبکه در هنگام پس‌انتشار خطا (Backpropagation) با مشکل محو شدن گرادیان مواجه نشود و یادگیری در شبکه‌های بسیار عمیق امکان‌پذیر گردد. به عبارت دیگر این شبکه تلاش می‌کند تا در هر بلوک تفاوت بین خروجی و ورودی یادگرفته شود. در صورتیکه یک بلوک نتواند بهبود خاصی در



شکل ۲: معماری ResNet-34

Fig. 2: ResNet-34 architecture

معماری UNet معمولی دارای اتصالات مستقیم از هر لایه کدگذار به لایه‌ی متناظر در رمزگشا است تا ویژگی‌های جزئی‌تر لایه‌های کم‌عمق‌تر را به مراحل بازیابی وضوح (Decoding) منتقل کند. با این حال، تفاوت معنایی قابل توجهی بین ویژگی‌های سطح بالا (کدگذار عمیق) و ویژگی‌های سطح پایین (کدگذار کم‌عمق) وجود دارد که گاهی ادغام مستقیم آنها در UNet چالش‌برانگیز است که شبکه UNet++ با معرفی اتصالات تودرتو این مشکل را برطرف می‌کند؛ به جای یک اتصال مستقیم ساده، خروجی هر لایه‌ی کدگذار ابتدا با گذر از چند لایه‌ی پیچشی میانی در سطوح مختلف پالایش شده و سپس به لایه‌ی رمزگشا منتقل می‌گردد. بدین ترتیب، ویژگی‌های سطح پایین طی چند مرحله‌ی میانی با ویژگی‌های سطح بالاتر ترکیب شده و شکاف معنایی بین بخش کدگذار و رمزگشا پر می‌شود. علاوه بر این، UNet++ از مکانیزمی به نام نظارت عمیق (Deep Supervision) نیز بهره می‌گیرد که طی آن خروجی‌های میانی چندگانه شبکه نیز تحت نظارت تابع هزینه قرار می‌گیرند تا گرادینت‌های مؤثرتر و شبکه‌ی با قابلیت تعمیم بهتر حاصل شود [۱۱، ۲۸].



شکل ۳: معماری UNet++
Fig. 3: UNet++ architecture

شکل (۲) معماری UNet++ را نمایش می‌دهد. در این شکل، دایره‌های آبی تیره نشان‌دهنده‌ی خروجی بلوک‌های کدگذار و ویژگی‌های استخراج‌شده در اعماق شبکه‌ی backbone و دایره‌های آبی روشن نشان‌دهنده‌ی خروجی واحدهای میانی/رمزگشا در سطوح گوناگون هستند. فلش‌های توپر سیاه رو به پایین فرآیند Down-sampling یا کاهش وضوح (در مسیر کدگذار) و فلش‌های رو به بالا عملیات Up-sampling یا افزایش وضوح (در مسیر رمزگشا) را نشان

شکل ۱ معماری کلی ResNet-34 را نشان می‌دهد که یک شبکه ۳۴ لایه‌ای از خانواده ResNet است. این معماری یک شبکه‌ی نسبتاً عمیق با حدود ۲۱ میلیون پارامتر است که به خاطر به‌کارگیری ساختار باقیمانده، قدرت یادگیری ویژگی‌های پیچیده را بدون مشکل ناشی از افزایش عمق دارد. در این معماری ابتدا یک لایه پیچش با کرنل 7×7 و ۶۴ فیلتر (با اندازه گام ۲) به همراه یک لایه Pooling برای کاهش ابعاد تصویر به‌کار می‌رود. پس از آن شبکه وارد چهار مرحله‌ی متوالی بلوک‌های Conv2_x تا Conv5_x می‌شود. در هر مرحله، چند بلوک باقیمانده شامل دو لایه‌ی پیچشی 3×3 وجود دارد. در این معماری، تعداد بلوک‌های هر مرحله به ترتیب ۳، ۴، ۶ و ۳ بلوک است که مجموعاً ۳۲ لایه‌ی پیچشی را تشکیل می‌دهند. با ورود به هر مرحله جدید، تعداد فیلترها دو برابر می‌شود (به عنوان مثال از ۶۴ به ۱۲۸، ۱۲۸ به ۲۵۶ و ۲۵۶ به ۵۱۲) و اولین بلوک آن مرحله با افزایش گام اندازه‌ی ویژگی‌ها را نصف می‌کند. هر بلوک باقیمانده دارای یک اتصال میان‌بر است که خروجی دو لایه‌ی 3×3 آن بلوک را با ورودی بلوک جمع می‌کند. این اتصال مستقیم (مشخص شده با فلش‌های منحنی در شکل) موجب می‌شود خروجی هر بلوک حاوی اطلاعات لایه‌های قبلی نیز باشد و در نتیجه گرادینت خطا در جهت معکوس، بدون افت شدید، می‌تواند به لایه‌های ابتدایی برسد. در انتها نیز پس از آخرین بلوک باقیمانده، یک لایه میانگین‌گیری سراسری (Global Average Pooling) و یک لایه Fully connected قرار می‌گیرد که نقش دسته‌بندی نهایی را ایفا می‌کند [۲۷].

مزایای ResNet و نقش آن در مدل حاضر

معماری ResNet به دلیل توانایی خود در یادگیری هویت نگاشت (Identity Mapping)، این امکان را فراهم می‌سازد که لایه‌های اضافی نه تنها به مدل آسیب نزنند بلکه بهبود دهنده مدل نیز باشند. در مدل حاضر، از ResNet-34 به عنوان بخش کدگذار شبکه قطعه‌بندی استفاده شده است. نقش ResNet-34 استخراج ویژگی‌های چند مقیاسی و عمیق از تصویر ورودی که شامل اطلاعات طیفی و ارتفاعی است، می‌باشد. در سطوح اولیه این شبکه ویژگی‌های محلی و جزئی مانند بافت و لبه شناسایی می‌شوند این در حالی است که در لایه‌های عمیق‌تر روابط گسترده و الگوهای پیچیده‌تر در مقیاس بزرگتر استخراج می‌شوند. با این کار، بردارهای ویژگی نهایی که به بخش رمزگشای UNet++ تحویل داده می‌شوند، دارای اطلاعات محلی و الگوهای کلی از تصویر می‌باشند. ترکیب ResNet-34 با معماری UNet++ در این پروژه امکان یادگیری مدل با دقت بالا و تعمیم مدل را فراهم کرده است [۱۸].

شبکه UNet++ و ساختار آن

UNet++ یک ساختار پیشرفته برای شبکه‌های کدگذار-رمزگشا مخصوص قطعه‌بندی تصاویر است که نسبت به UNet کلاسیک بهبودهایی در بخش اتصالات بین کدگذار و رمزگشا ارائه می‌دهد.

بهینه‌ساز Adam

برای به‌روزرسانی وزن‌های شبکه در طی آموزش، از روش Adam (Adaptive Moment Estimation) استفاده شده است که برای هر وزن شبکه دو میانگین نمایی متحرک نگهداری می‌شود. اولین میانگین، میانگین درجه یک گرادیان‌ها بوده که نقشی مشابه Momentum دارد و نوسانات گرادیان را هموار می‌کند و دیگری میانگین درجه دوم گرادیان‌ها که نقشی شبیه RMSProp دارد و مقیاسی از بزرگی مربع گرادیان‌ها ارائه می‌دهد [۳۳].

روش Adam به‌صورت خودتنظیم برای هر پارامتر یک نرخ یادگیری موثر تعیین می‌کند؛ پارامترهایی با گرادیان‌های بزرگ نرخ یادگیری کوچک‌تری می‌گیرند و بالعکس که این ویژگی باعث می‌شود Adam در مسائل دارای داده‌های بزرگ و مدل‌های پیچیده به سرعت همگرا شود و نیاز به تنظیم دستی مکرر نرخ یادگیری یا سایر ابرپارامترها کمتر باشد. در پژوهش حاضر نیز به دلیل وجود پارامترهای فراوان و اهمیت همگرایی پایدار، از این بهینه‌ساز استفاده شده است و با نرخ یادگیری اولیه مناسب با کاهش تدریجی آن طی دوران آموزش به مدل کمک می‌کند تا به دقت بالا رسیده و از نوسانات یادگیری جلوگیری شود، به طور کلی به دلیل کارایی حافظه و تطبیق‌پذیری خود در تنظیم گام‌های حرکت در فضای پارامترها، یک انتخاب استاندارد و قدرتمند برای بهینه‌سازی شبکه‌های عمیق محسوب می‌شود [۳۴].

شاخص‌های ارزیابی

برای ارزیابی عملکرد مدل‌های شناسایی تغییرات از شاخص‌های متعددی استفاده می‌شود که در این پژوهش به سه معیار Accuracy، F1-score و Recall بهره گرفته شده است.

○ دقت کلی (Accuracy): معیاری است برای درصد پیش‌بینی‌های صحیح مدل روی کل داده‌ها به عبارتی Accuracy برابر با نسبت مجموع موارد پیش‌بینی‌شده صحیح (TP + TN) به کل موارد می‌باشد:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

○ Recall: این معیار که نرخ تشخیص صحیح مثبت نیز خوانده می‌شود، نسبت تعداد موارد مثبت درست شناسایی شده به کل موارد مثبت واقعی است. این معیار مشخص می‌کند که مدل تا چه میزان در شناسایی تمامی موارد مثبت موفق بوده است.

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

○ F1-score: این معیار یک شاخص کلی از عملکرد مدل ارائه می‌دهد که هم به کیفیت پیش‌بینی‌های مثبت یا precision و هم به پوشش موارد مثبت یا recall توجه دارد. این میانگین هارمونیک به گونه‌ای عمل می‌کند که در صورت پایین بودن هر کدام از معیارهای ذکر شده، مقدار F1-score نیز پایین خواهد بود.

$$F1 - Score = \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

لازم به ذکر است که به دلیل نامتوازن بودن داده‌ها و زیاد بودن تعداد پیکسل‌های بدون تغییر نسبت به پیکسل‌های تغییر، استفاده از معیار

می‌دهند. اتصالات نقطه‌چین، همان اتصالات skip سنتی UNet می‌باشند که در UNet++ تودرتو شده‌اند. خطوط نقطه‌چین مشکی اتصال مستقیم UNet اصلی بین یک لایه کدگذار و رمزگشا را نشان می‌دهد، در حالی که خطوط آبی نقطه‌چین اتصالات چندگانه‌ی جدید UNet++ هستند. همان‌طور که ملاحظه می‌شود، یک خروجی میانی (به طور مثال $x^{0,2}$) در شکل نه تنها از خروجی یک لایه‌ی کدگذار $x^{0,0}$ بهره می‌برد، بلکه خروجی لایه‌های کدگذار/رمزگشا در سطوح پایین‌تر $x^{1,1}$ و $x^{2,0}$ نیز پس از Upsample شدن و کانولوشن وارد ترکیب می‌شوند. به بیان ساده، UNet++ به جای یک اتصال منفرد بین کدگذار و رمزگشا، مسیرهای متعددی ایجاد می‌کند که به‌طور تدریجی ویژگی‌های با وضوح‌های مختلف را ترکیب کرده و اطلاعات بیشتری را به بخش رمزگشا می‌رساند [۲۹،۳۰]. در نهایت خروجی آخرین لایه‌ی رمزگشا که با دایره‌ی آبی تیره نمایش داده شده است در بالاترین سطح، مثلاً $x^{0,4}$ نقشه‌ی پیش‌بینی نهایی را تشکیل می‌دهد. تحقیقات نشان داده‌اند که این طراحی تودرتو می‌تواند دقت قطعه‌بندی را به طرز چشم‌گیری بهبود دهد و جزئیات ریزتری را در مرزبندی‌ها بازبایی نماید. در مدل حاضر، UNet++ به عنوان ساختار اصلی بخش رمزگشا مورد استفاده قرار گرفته است و خروجی ویژگی‌های عمیق ResNet-34 (با کانولوشن 1×1 برای تطبیق کانال‌ها) به عنوان ورودی این ساختار برای بازسازی نقشه‌ی قطعه‌بندی به کار می‌رود [۱۱].

تابع هزینه Lovász

در مسائل قطعه‌بندی، معمولاً معیار ارزیابی اصلی دقت ناحیه‌ای مثل IoU است اما توابع هزینه‌ی متداول مانند Cross Entropy به صورت غیرمستقیم این معیار را بهبود می‌دهند و ممکن است تحت شرایطی (مثل عدم توازن کلاس‌ها) مدل را به سمت بهینه‌ای ببرند که IoU چندان بالایی نداشته باشد. برای حل این مسئله، تابع زیان لاواژ-سافت‌مکس (Lovász-Softmax) معرفی شده است که یک جانشین قابل مشتق‌گیری برای بهینه‌سازی مستقیم IoU محسوب می‌شود [۳۱]. ایده‌ی اصلی این تابع آن است که به‌جای محاسبه‌ی مستقیم IoU (که غیرقابل مشتق است)، از مفهوم بسط لاواژ (Lovász Extension) در ریاضیات استفاده کند تا نسخه‌ای پیوسته و زیرتفاضلی (sub-differentiable) از IoU را به دست آورد. به عبارتی، طی محاسبه‌ی این زیان، پیش‌بینی‌های شبکه بر اساس احتمال تعلق به کلاس مرتب‌سازی می‌شوند و سپس اختلاف ترتیب آن‌ها با ترتیب ایده‌آل برچسب‌های واقعی سنجیده می‌شود. این عدم تطابق ترتیبی، مقدار Lovász Loss را تشکیل می‌دهد که در فرآیند یادگیری کمینه می‌گردد. ویژگی مهم Lovász-Softmax آن است که مستقیماً IoU میانگین را بهبود داده و در عمل مشاهده شده که آموزش با این تابع می‌تواند به امتیاز IoU بالاتری نسبت به Cross Entropy منجر شود. مزیت دیگر این تابع، مقاومت در برابر عدم توازن کلاس است؛ زیرا بر اساس رتبه‌بندی پیش‌بینی‌ها، تاثیر نواحی کوچک ولی مهم حفظ می‌شود [۳۲].

پایاده سازی

در این پژوهش به منظور تشخیص تغییرات سه بعدی در ساختمان‌های شهری، از مجموعه داده 3DCD استفاده شده است. این مجموعه داده شامل تصاویر هوایی با وضوح بالا به همراه داده‌های مدل رقومی سطح (DSM) مربوط به دو بازه‌ی زمانی مختلف می‌باشد. تمامی تغییرات مورد بررسی مربوط به ساختمان‌ها بوده و به پنج کلاس (۱) بدون تغییر، (۲) ساخت ساختمان جدید در مکان فاقد بنا، (۳) تخریب کامل ساختمان، (۴) افزایش ارتفاع یک ساختمان موجود و (۵) کاهش ارتفاع ساختمان تقسیم شده‌اند. هدف از این تقسیم‌بندی، استخراج جزئیات دقیق‌تر از تحولات شهری در بعد ارتفاعی می‌باشد.

مجموع کل داده‌های مورد استفاده ۴۲۳ جفت تصویر می‌باشد که شامل اطلاعات طیفی (RGB) و رقومی (DSM) در دو بازه زمانی می‌باشد. هر تصویر اولیه دارای ابعاد 400×400 پیکسل بود که به منظور افزایش سازگاری با معماری مدل و بهره‌گیری بهتر از منابع سخت‌افزاری، به ابعاد 512×512 پیکسل بازنمونه‌برداری شده‌اند. برای هر نمونه، تصاویر به همراه ارتفاع متناظر در هر تاریخ با یکدیگر ادغام شده و سپس اطلاعات دو زمان به صورت کانال‌های پیاپی روی هم قرار گرفته‌اند. در نهایت، تصویری با هشت باند به عنوان ورودی مدل ساخته شده است. برای افزایش تنوع داده‌های آموزشی و بهبود عملکرد مدل، از تکنیک‌های داده‌افزایی شامل چرخش تصادفی، وارونگی افقی و وارونگی عمودی استفاده شد و این عملیات به صورت تصادفی بر روی تصاویر آموزشی اعمال گردید تا مدل در برابر تغییرات ظاهری مقاوم‌تر عمل کند.

فرآیند پایاده‌سازی مدل در دو مرحله مجزا انجام شده است که در ادامه هر مرحله توضیح داده شده است.

مرحله اول: تشخیص دودویی تغییرات و مقایسه با روش‌های سنتی

در این مرحله، برچسب‌های داده‌ی حقیقت-زمینی (Ground Truth) به صورت دودویی تنظیم شدند؛ به این معنی که صرفاً دو کلاس «تغییر» و «عدم تغییر» در نظر گرفته شده است و مدل پیشنهادی نیز به عنوان یک طبقه‌بند دودویی آموزش داده شده است. هدف از این مرحله، بررسی عملکرد پایه‌ای مدل و مقایسه آن با روش‌های سنتی تشخیص تغییرات بوده است [۳۹].

برای ارزیابی عملکرد مدل با روش‌های سنتی از سه مدل جنگل تصادفی، تقسیم دو تصویر و روش ترکیبی PCA و KMeans استفاده شده است.

- طبقه‌بند جنگل تصادفی (Random Forest): این روش یکی از الگوریتم‌های شناخته‌شده‌ی یادگیری نظارت‌شده می‌باشد که بر پایه‌ی مجموعه‌ای از درختان تصمیم‌گیری عمل می‌کند. در این پژوهش، ویژگی‌های استخراج‌شده از تصاویر ورودی برای هر پیکسل (یا ناحیه) به عنوان ورودی به طبقه‌بندی داده شده و خروجی مدل مشخص می‌کند که آن پیکسل به کلاس «تغییر» یا «عدم تغییر» تعلق دارد.

Accuracy می‌تواند گمراه کننده باشد و در این موارد معیارهایی مثل Recall و F1-score اهمیت بیشتری پیدا می‌کنند [۳۵].

افزایش داده (Data Augmentation)

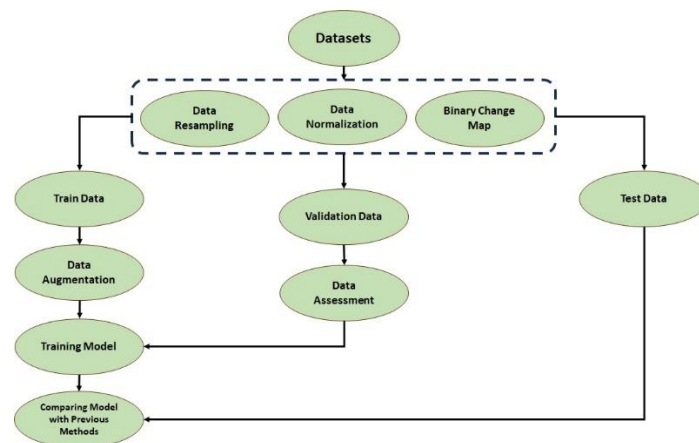
یکی از اجزای مهم در مرحله‌ی آموزش مدل، به کارگیری افزایش داده است که هدف آن بالا بردن تنوع داده‌های آموزشی و در نتیجه بهبود تعمیم‌پذیری مدل می‌باشد. با توجه به محدودیت تعداد تصاویر آموزشی، اعمال تغییرات تصادفی بر روی تصاویر موجود می‌تواند به مدل کمک کند تا نسبت به تغییرات مختلف غیرحساس‌تر شود و از حفظ شدن بیش از حد روی الگوهای خاص داده‌های آموزشی اجتناب شود [۲۲]. در این پژوهش از چند نوع تغییر داده مرسوم استفاده شده که در ادامه به توضیح آن‌ها پرداخته شده است.

○ وارون‌سازی افقی و عمودی (Horizontal/Vertical Flip): در این روش تصویر به ترتیب نسبت به محور عمودی یا افقی قرینه می‌شود. برای مثال، در وارون‌سازی افقی، سمت چپ و راست تصویر جابجا می‌شوند. این تکنیک به ویژه زمانی مفید است که جهت‌گیری اشیاء در تصاویر نباید روی برچسب تأثیر بگذارد. مدل با دیدن تصاویر پشت و رو قادر خواهد بود ویژگی‌ها را مستقل از جهت‌شناسایی کند. در این پروژه، تصاویر به طور تصادفی با احتمال مشخصی (مثلاً ۵۰٪) افقی یا عمودی برگردانده شده‌اند تا مجموعه‌ی داده آموزشی حاوی حالت‌های آینه‌ای نیز باشد [۳۶].

○ دوران (Rotation): در این روش تصویر حول مرکز خود با یک زاویه‌ی تصادفی دوران می‌یابد و این چرخش می‌تواند محدود به زوایای کوچک (مثلاً $15 \pm$ درجه) یا شامل زوایای ۹۰ درجه‌ای باشد. این افزونه باعث می‌شود مدل نسبت به چرخش شیء حساسیت کمتری داشته باشد. به عنوان نمونه، در داده‌های ماهواره‌ای یا هوایی، چرخش تصویر معادل تغییر جهت شمال نقشه است که نباید ماهیت شیء (مثلاً یک ساختمان) را برای مدل متفاوت جلوه دهد. در این پروژه از چرخش‌های کوچک استفاده شده است تا مدل انواع جهت‌ها را ببیند [۲۳].

○ تاری گاوسی (Gaussian Blur): در این روش، یک فیلتر Gaussian با سطح محوشدگی معین روی تصویر اعمال می‌شود تا تصویر کمی مات شود. اعمال تاری کنترل‌شده می‌تواند مدل را در برابر تصاویر کمی خارج از فوکوس یا نویز سنسور مقاوم کند. علاوه بر این، تا حدی مانند یک فیلتر پایین‌گذر عمل کرده و از وابستگی بیش از حد مدل به لبه‌های تیز نیز جلوگیری می‌کند [۳۷].

نتیجه‌ی کلی استفاده از افزونگی داده این است که مدل آموزش‌دیده پایدارتر می‌شود؛ به این معنی که در برابر تغییرات هندسی (مانند چرخش و معکوس شدن) و تغییرات کیفی (مانند تاری یا نویز) عملکرد خود را حفظ می‌کند و در مواجهه با تصاویر جدید (داده‌های تستی) که دقیقاً شبیه به تصاویر آموزشی نیستند، دقت بالاتری نشان خواهد داد [۳۸].



شکل ۴: فرآیند تشخیص دودویی تغییرات

Fig. 4: Binary change detection flowchart

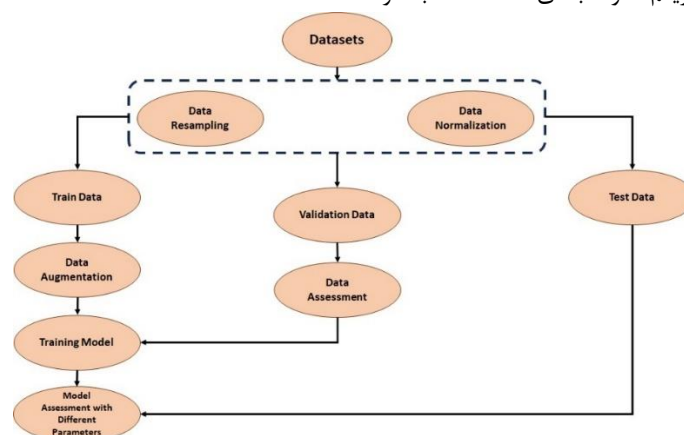
خوشه اجرا می‌شود تا پیکسل‌ها به دو دسته‌ی تغییر و عدم تغییر تفکیک شوند. این روش در مقالاتی معرفی شده و به‌عنوان یک روش مرجع در مطالعات تشخیص تغییر نظارت نشده شناخته می‌شود [۴۲].

مرحله دوم: آموزش مدل برای پنج کلاس و تنظیم ابرپارامترها در این مرحله، پس از اثبات عملکرد برتر مدل در تشخیص تغییرات دودویی، مدل نهایی برای شناسایی پنج کلاس مختلف تغییرات ساختمانی، آموزش داده شد. ساختار مدل همانند مرحله اول شامل ترکیب شبکه ResNet-34 به‌عنوان کدگذار و UNet++ به‌عنوان رمزگشا بود. ورودی مدل نیز همچنان شامل هشت باند (ترکیب RGB و DSM قبل و بعد) بوده و از داده‌افزایی مشابه مرحله اول بهره گرفته شد. برای یافتن ترکیب بهینه ابرپارامترها و تابع زیان، از مجموعه داده‌ی تست جهت ارزیابی تنظیمات مختلف استفاده شد. مقادیر گوناگونی برای نرخ یادگیری، نوع تابع زیان و سایر پارامترهای آموزش بررسی شدند و در نهایت بهترین ترکیب شامل تابع Lovasz Loss و نرخ یادگیری ۰.۰۰۰۰۱ انتخاب گردید. مدل با پارامترهای بهینه روی مجموعه آموزشی (۳۸۳ تصویر) آموزش داده شد و بر اساس عملکرد آن روی داده‌های تست (۴۰ تصویر)، نتایج نهایی مدل پنج کلاسه استخراج گردید.

استفاده از جنگل تصادفی به دلیل مقاومت بالا در برابر بیش‌برازش و قابلیت مدل‌سازی روابط غیرخطی بین ویژگی‌ها، به‌عنوان یک مرجع مناسب در مقایسه با مدل یادگیری عمیق می‌باشد [۴۰].

روش Image Ratio (یا تقسیم تصاویر): این روش یک روش ساده و نظارت نشده برای شناسایی تغییرات در تصاویر سنجش‌ازدور است. در این رویکرد، ابتدا اختلاف پیکسل‌به‌پیکسل بین تصاویر قبل و بعد به صورت نسبت‌گیری محاسبه و سپس بر روی تصویر حاصل یک آستانه‌گذاری انجام می‌شود تا مناطقی که تغییرات معنی‌داری در آن‌ها رخ داده، مشخص گردند. مزیت این روش سادگی و سرعت بالای اجرای آن است، اما ضعف اصلی آن در وابستگی آن به حدآستانه‌گذاری در تمایز بین نوع تغییر و حساسیت پایین به نویزهای تصویری می‌باشد [۴۱].

روش ترکیبی PCA و KMeans: این روش یک الگوریتم نظارت نشده مبتنی بر تحلیل تغییرات آماری بین دو تصویر است. ابتدا تفاضل بین دو تصویر (تاریخ قبل و بعد) به‌صورت پیکسل به پیکسل محاسبه می‌شود و تصویر اختلاف حاصل، نمایانگر نقاط تغییر احتمالی است. سپس برای کاهش نویز و کاهش ابعاد داده، تحلیل مؤلفه‌های اصلی (PCA) بر روی تصویر اختلافی اعمال می‌شود تا مؤثرترین مؤلفه‌ها استخراج گردند. در نهایت، بر روی خروجی PCA، الگوریتم خوشه‌بندی K-Means با دو



شکل ۵: فرآیند تشخیص چند کلاس تغییرات

Fig. 5: Multi-class change detection flowchart

نتایج و بحث

ارزیابی مرحله اول (دودویی)

در مرحله اول، مدل یادگیری عمیق پیشنهادی به صورت دودویی (دو کلاس «تغییر» و «عدم تغییر») ارزیابی شد تا توانایی آن در آشکارسازی کلی تغییرات سنجیده شود. در این مرحله مدل تنها به تعداد ۵۰ اپک آموزش داده شد و نتایج حاصل از این ارزیابی در جدول ۱ خلاصه شده است که شامل معیارهای دقت، F1-score و Recall برای مدل پیشنهادی و روش‌های مرجع می‌باشد. مدل عمیق توانست با دقت کلی بالا نواحی تغییر یافته را به خوبی از نواحی بدون تغییر تفکیک کند. به علاوه، F1-score قابل توجهی برای این مدل به دست آمد که بیانگر تعادل مناسب میان دقت آشکارسازی تغییرات و نرخ خطای پایین در اعلام تغییر نادرست است. مقدار بالای معیار Recall نیز نشان‌دهنده این است که بخش اعظم تغییرات واقعی شناسایی کرده و موارد کمتری از تغییرات از توسط مدل شناسایی نشده‌اند. از سوی دیگر، مقدار تابع زیان مدل بر روی داده‌های آزمایشی به سطح پایینی همگرا شد که نشان‌دهنده یادگیری موثر مدل و تطابق خوب خروجی‌های آن با برچسب‌های واقعی می‌باشد.

نکته حائز اهمیت آن است که به دلیل عدم توازن ذاتی داده‌ها در مسئله آشکارسازی تغییر (تعداد بسیار بیشتر پیکسل‌های بدون تغییر نسبت به پیکسل‌های تغییر کرده)، استفاده از معیار دقت کلی به تنهایی برای ارزیابی کافی نمی‌باشد؛ از این رو، استفاده از F1-score و Recall برای ارزیابی عملکرد مدل ضروری است [۴۳، ۱۳]. در مدل پیشنهادی، مقدار بالای Recall حاکی از توانایی آن در آشکارسازی اکثر تغییرات (حساسیت بالا) بوده و F1-score بالای حاصل شده تأیید می‌کند که نرخ اخطار اشتباه (False Alarm) نیز پایین نگه داشته شده است. به بیان دیگر، مدل با حفظ دقت، قادر به شناسایی تغییرات جزئی در مناطق وسیع بدون تغییر بوده است. این نتایج اولیه نشان می‌دهد که روش یادگیری عمیق توسعه یافته به خوبی از عهده‌ی آشکارسازی کلی تغییرات برآمده و پایه‌ی محکمی برای مرحله بعدی (تفکیک انواع تغییر) فراهم ساخته است [۴۴].

مقایسه با روش‌های مرجع

به منظور ارزیابی دقیق‌تر عملکرد مدل یادگیری عمیق، نتایج مرحله دودویی آن با سه روش مرجع مقایسه جنگل تصادفی (Random Forest)، روش نسبت تصویری (Image Ratio) و روش ترکیبی PCA و KMeans گردید. در جدول ۱ معیارهای دقت، F1-score و Recall برای هر یک از این روش‌ها در کنار مدل پیشنهادی ارائه شده است و این مقایسه‌ها نشان می‌دهد که مدل یادگیری عمیق پیشنهادی در تمامی معیارها عملکرد برتری نسبت به روش‌های مرجع داشته است. دقت مدل پیشنهادی به طور محسوس بالاتر از سایرین بود؛ به طوری که نسبت به بهترین روش کلاسیک (جنگل تصادفی) نیز چندین درصد افزایش دقت نشان داده است. همچنین پارامتر F1-score مدل به طرز چشمگیری از

روش‌های کلاسیک بالاتر است که بیانگر کارایی بهتر آن در تشخیص تغییرات واقعی می‌باشد.

با وجود عملکرد قابل توجه جنگل تصادفی، نتیجه آن در مقایسه با مدل یادگیری عمیق به میزان معناداری ضعیف‌تر بود. روش نسبت تصویری، ضعیف‌ترین عملکرد را داشت که دقت و F1-score پایین آن نشان‌دهنده این موضوع می‌باشد که در این روش تغییرات شدت روشنایی یا سایر تغییرات غیرساختاری به اشتباه به عنوان تغییر شناسایی شده و نرخ بالایی از تشخیص‌های نادرست را موجب گردیده است. روش KMeans+PCA به عنوان یک روش تغییریابی بدون نظارت، عملکرد بهتری نسبت به روش نسبت تصویری داشت ولی همچنان در مقایسه با روش‌های یادگیری نظارت‌شده ضعیف‌تر عمل کرد. این روش در تفکیک تغییرات کوچک و جزئی ناتوان‌تر بود و برخی از تغییرات با دامنه کم (مثلاً کاهش جزئی ارتفاع) را تشخیص نداده که باعث کاهش Recall آن نسبت به مدل عمیق گردید.

با توجه به معیار Recall، مدل پیشنهادی عملکرد بسیار بالاتری را کسب کرد که نشان می‌دهد تقریباً تمامی تغییرات واقعی توسط این مدل کشف شده است این در حالی می‌باشد که روش‌های مرجع (خصوصاً روش‌های بدون نظارت) تعدادی از تغییرات را کشف نکرده و Recall کمتری داشتند. به عنوان مثال، جنگل تصادفی تعدادی از تغییرات با مقیاس کوچک را از دست داد و روش‌های بدون نظارت مانند PCA+KMeans نیز تنها تغییرات با شدت بالا را به خوبی شناسایی کردند. برتری قابل ملاحظه‌ی مدل یادگیری عمیق در این مرحله، حاکی از توانایی آن در یادگیری الگوهای پیچیده‌ی تغییرات شهری است که روش‌های سنتی فاقد آن هستند. نتایج نیز تأکید می‌کند که رویکرد یادگیری عمیق قادر است با استخراج ویژگی‌های سطح بالا و تلفیق مؤثر اطلاعات طیفی و ارتفاعی، تغییرات ساختمانی و سازه‌ای را با دقت بیشتر تشخیص دهد؛ در حالی که روش‌های مرجع مبتنی بر ویژگی‌های دستی یا آستانه‌گذاری ساده در مواجهه با پیچیدگی داده‌های شهری کارآمدی کمتری از خود نشان می‌دهند [۴۵، ۹].

جدول ۱: معیارهای ارزیابی روش‌ها بر روی داده اعتبارسنجی
Table 1: Evaluation metrics for methods based on validation dataset

F1-SCORE	ACCURACY	RECALL	METHOD NAME
0.2236	0.8353	0.3597	IMAGE RATIO
0.2800	0.8932	0.3767	KMEANS+PCA
0.2426	0.9125	0.4920	RF
0.9258	0.9850	0.9860	UNET++

نتایج مرحله دوم (پنج کلاس)

پس از اثبات عملکرد برتر مدل در تشخیص دودویی تغییرات، مدل نهایی برای طبقه‌بندی پنج‌کلاسه آموزش داده شد. در این مرحله علاوه بر آشکارسازی تغییر، نوع تغییر نیز طبقه‌بندی می‌شود. پنج کلاس تعریف شده شامل عدم تغییر، ظهور سازه‌های جدید مانند احداث ساختمان، تخریب کامل، افزایش ارتفاع (مثلاً اضافه شدن طبقات به

تغییرات دودویی، قابلیت تعمیم و تشخیص انواع مختلف تغییرات را نیز کسب کرده است و می‌تواند تغییرات پیچیده‌تری را در محیط شهری شناسایی و طبقه‌بندی کند [۴۷].

تحلیل کیفی نتایج (تصاویر نمونه)

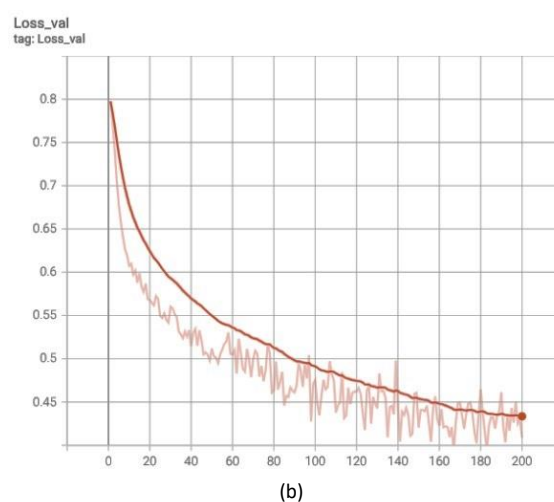
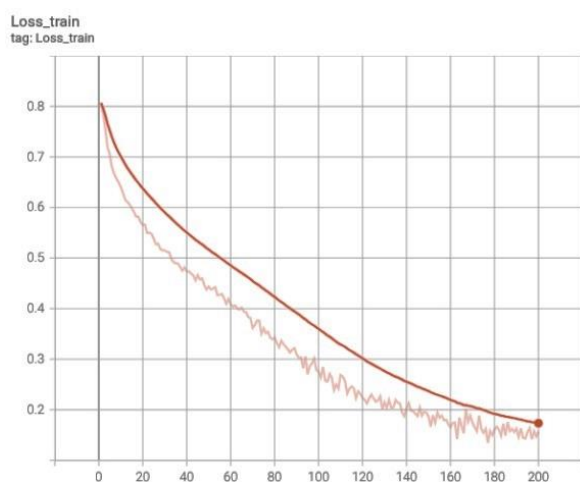
علاوه بر ارزیابی‌های کمی، به منظور درک بهتر کیفیت عملکرد مدل، تحلیل کیفی نتایج نیز انجام گرفته است. بدین منظور، خروجی مدل بر روی نمونه‌هایی از داده‌های آزمایشی به صورت بصری بررسی شد. شکل ۶ و ۷ نمونه‌هایی از نقشه تغییرات پیش‌بینی شده توسط مدل را نشان می‌دهد. در این تصاویر، نواحی بدون تغییر با رنگ بنفش و سایر کلاس‌ها با رنگ‌های برجسته نمایش داده شده‌اند.

بررسی کیفی شکل ۶ نشان می‌دهد که مدل پیشنهادی به خوبی توانسته تغییرات واقعی را در تصاویر دو زمانه آشکارسازی و تفکیک کند. به عنوان نمونه، سازه‌های جدیدی که در بازه زمانی بین دو تصویر احداث شده‌اند، توسط مدل به درستی در نقشه خروجی به عنوان «ساخت جدید» مشخص شده‌اند. این سازه‌ها در تصویر مربوط به زمان اول وجود نداشته و در تصویر زمان دوم پدیدار شده‌اند که مدل دقیقاً همین نواحی را به عنوان ساخت جدید شناسایی کرده است.

بررسی شکل ۷ نشان می‌دهد که مدل تغییرات ظریف ارتفاعی را تا حد زیادی به درستی شناسایی کرده است. هرچند مرزبندی این تغییرات ارتفاعی به دلیل ماهیت تدریجی ترشان کمی پراکنده‌تر از تغییرات کاملاً متمایز (مانند ساخت یا تخریب) به نظر می‌رسد، اما حضور رنگ مربوط به افزایش ارتفاع در مکان صحیح حاکی از آن است که مدل این تغییرات را نیز تشخیص داده است. به طور مشابه، کاهش ارتفاع یا نیمه‌تخریب‌های رخ داده در برخی ساختمان‌ها نیز توسط مدل شناسایی شده و با رنگ کلاس «کاهش ارتفاع» در نقشه خروجی منعکس گردیده است.

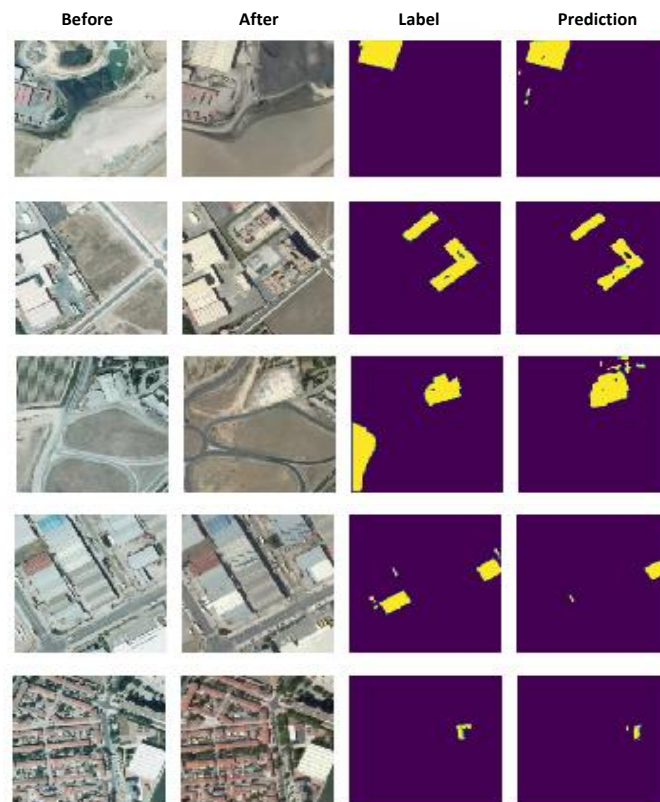
ساختمان‌های موجود و کاهش ارتفاع می‌باشند. نتایج به دست آمده حاکی از آن است که مدل پیشنهادی در طبقه‌بندی پنج کلاس نیز عملکرد مناسبی داشته و بیشتر الگوهای گوناگون تغییر را آموزش دیده است. معیارهای Accuracy، F1-score و Recall بر روی مجموعه اعتبارسنجی به ترتیب برابر ۹۶.۳۲، ۰.۹۵ و ۰.۹۶ می‌باشد که نشان‌دهنده توانایی تعمیم مناسب مدل و عدم بیش‌برازش است. با نگاهی دقیق‌تر به عملکرد مدل در طبقه‌بندی تصاویر، مشخص می‌شود که کلاس‌های با تغییرات بارز (ساخت جدید و تخریب کامل) با دقت بالایی شناسایی شده‌اند. مدل تقریباً تمامی موارد ساخت سازه‌های جدید را به درستی برچسب‌گذاری کرده و تخریب‌های کامل را نیز با قطعیت بالا تشخیص داده است. این دو نوع تغییر به دلیل ماهیت مشخص خود (ظهور یا حذف کامل یک ساختار) به خوبی توسط مدل آموخته شده و حتی در داده‌های اعتبارسنجی نیز تقریباً بدون خطا شناسایی شده‌اند. در مقابل، در کلاس‌های افزایش ارتفاع و کاهش ارتفاع که ماهیت تغییرات جزئی‌تر و تدریجی‌تر است، عملکرد مدل اندکی ضعیف‌تر از دو کلاس مذکور بوده است.

مقادیر تابع زیان در طی آموزش مدل پنج کلاس نشان می‌دهد که فرآیند بهینه‌سازی شبکه به خوبی صورت گرفته است. تابع زیان بر روی داده‌های آموزشی پس از چند دوره آموزش به مقدار کمی کاهش یافته و همزمان تابع زیان مجموعه اعتبارسنجی نیز به طور پایدار پایین باقی مانده است. نزدیک بودن مقادیر زیان آموزش و اعتبارسنجی بیانگر آن است که مدل دچار بیش‌برازش نشده و عملکرد آن در مواجهه با داده‌های جدید، مشابه عملکرد روی داده‌های آموزشی است. به طور کلی، اندکی کاهش در دقت و سایر شاخص‌ها در مجموعه اعتبارسنجی نسبت به آموزش طبیعی بوده و نشان‌دهنده سخت‌تر بودن مسئله چندکلاس در مقایسه با حالت دودویی است [۴۶]. با این حال، نتایج مرحله دوم تأیید می‌کند که مدل عمیق پیشنهادی پس از یادگیری

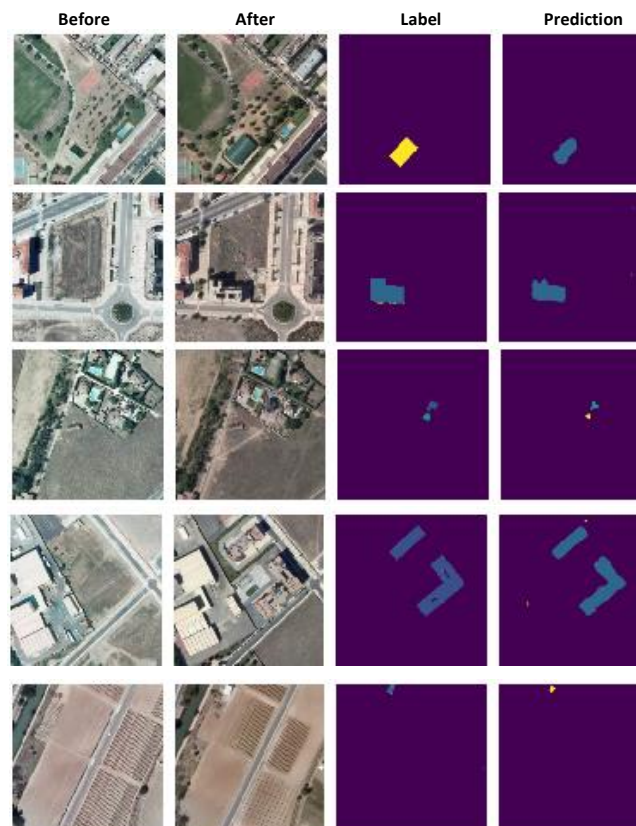


شکل ۶: نمودارهای تغییرات Loss در فرآیند آموزش مدل: (الف) آموزش (ب) اعتبارسنجی

Fig 6: Loss changes during model training: (a) Train (b) Validation



شکل ۷: نتایج خروجی شناسایی تغییرات دودویی
Fig. 7: Binary change detection results



شکل ۸: نتایج خروجی شناسایی تغییرات چند کلاس
Fig. 8: Multi-class change detection results

تغییر، بلکه در تعیین نوع کلاس آن است که برای تحلیل دقیق تغییرات شهری مزیتی قابل توجه محسوب می‌شود.

این بهبود عملکرد را می‌توان ناشی از قدرت یادگیری مدل عمیق در به دست آوردن ویژگی‌های پیچیده مکانی-طیفی دانست. الگوریتم‌های کلاسیکی نظیر جنگل تصادفی یا روش‌های مبتنی بر نسبت‌گیری یا خوشه‌بندی در استخراج خودکار الگوهای تغییر محدودیت داشته و معمولاً در مواجهه با تغییرات جزئی یا غیرخطی دقت کمتری نشان می‌دهند. از سوی دیگر، شبکه‌های عصبی عمیق به راحتی قادر به ترکیب و تحلیل همزمان انواع داده (مانند تصاویر چندزمانه و مدل‌های ارتفاعی) هستند که این موضوع باعث افزایش دقت و انعطاف‌پذیری در آشکارسازی تغییرات می‌شود [۲۰، ۹].

در راستای ارتقای مدل پیشنهادی و رفع چالش‌های مذکور، چند محور اصلی برای تحقیقات آینده پیشنهاد می‌گردد:

- بهبود دقت مدل با بازطراحی ساختار کدگذار و رمزگشا در شبکه
- ارائه معماری‌های جدید یادگیری عمیق مانند شبکه‌های تبدیل‌کننده
- بکارگیری داده‌های سری زمانی جهت آشکارسازی تغییرات در لایه شهری
- پیاده‌سازی مدل‌های Double-Stream برای استخراج ویژگی در دو لایه

مشارکت نویسندگان

علیرضا ابراهیمی در جمع‌آوری داده‌ها، تحلیل داده‌ها و نگارش پیش‌نویس متن مشارکت داشته است. مهدی حسنلو در ایده‌پردازی، طراحی پژوهش و اصلاح اساسی متن مقاله مشارکت داشته است.

تشکر و قدردانی

بدین وسیله از تمامی کسانی که ما را در انجام این تحقیق یاری نموده‌اند، تشکر و قدردانی می‌کنیم.

تعارض منافع

«هیچ‌گونه تعارض منافع توسط نویسندگان بیان نشده است.»

منابع و مأخذ

[1] Wen D, Huang X, Bovolo F, Li J, Ke X, Zhang A, et al. Change detection from very-high-spatial-resolution optical remote sensing images: Methods, applications, and future directions. *IEEE Geoscience and Remote Sensing Magazine*. 2021;9(4):68-101. DOI: 10.1109/MGRS.2021.3063465.

[2] Viana CM, Oliveira S, Oliveira SC, Rocha J. Land use/land cover change detection and urban sprawl analysis. *Spatial modeling in GIS and R for earth and environmental sciences*: Elsevier; 2019. p. 621-51. DOI: 10.1016/B978-0-12-815226-3.00029-6.

نواحی گسترده بدون تغییر در هر دو شکل به درستی توسط مدل دست‌نخورده باقی مانده‌اند و هیچ برچسب تغییری دریافت نکرده‌اند که این نشان‌دهنده نرخ پایین هشدار غلط مدل است. به بیان دیگر، مدل صرفاً در مکان‌هایی اعلام تغییر کرده که شواهد معتبری از وقوع تغییر وجود داشته و برای بخش‌های وسیعی که وضعیت سابق خود را حفظ کرده‌اند، خروجی مدل «عدم تغییر» باقی مانده است. این رفتار مطلوب، مؤید آن است که مدل یادگیری عمیق ما دچار بیش‌آشکارسازی تغییر (over-detection) نشده و به خوبی میان نواحی ثابت و نواحی متحول‌شده تمایز قائل می‌شود [۱۰].

در جمع‌بندی بخش تحلیل کیفی می‌توان گفت که مشاهدات بصری کاملاً مؤید نتایج کمی به‌دست‌آمده هستند. مدل پیشنهادی نه تنها در آمار بلکه در خروجی بصری نیز عملکرد قابل ملاحظه‌ای از خود نشان داده است و توانایی مدل در شناسایی صحیح انواع تغییرات ساختمانی در محیط شهری به صورت عینی در شکل‌های نمونه قابل مشاهده می‌باشد. این امر بیانگر آن است که مدل یادگیری عمیق توسعه‌یافته ابزار قابل اعتمادی برای پایش تغییرات سه‌بعدی شهری بوده و امکان تشخیص تغییرات گوناگون با دقت بالا و خطای پایین را دارد. برتری مدل نسبت به روش‌های کلاسیک در هر دو بعد کمی و بصری آشکار بوده و بدین ترتیب کاربرد عملی آن در مطالعات آتی پایش شهری و برنامه‌ریزی شهری نویدبخش می‌باشد.

نتیجه‌گیری

در این پژوهش، یک مدل مبتنی بر یادگیری عمیق به منظور تشخیص تغییرات سه‌بعدی در مناطق شهری توسعه داده شد. ارزیابی عملکرد مدل طی یک فرآیند دو مرحله‌ای صورت گرفته است به طوریکه در مرحله نخست، توانایی مدل در آشکارسازی دو کلاسه «تغییر» یا «عدم تغییر» ارزیابی و نتایج آن با روش‌های کلاسیک (از جمله جنگل تصادفی، روش نسبت‌گیری تصاویر و ترکیب PCA و KMeans) مقایسه گردید. در مرحله دوم، خروجی مدل به صورت تفصیلی‌تر به پنج کلاس شامل (۱) عدم تغییر، (۲) ساخت جدید، (۳) تخریب کامل، (۴) افزایش ارتفاع و (۵) کاهش ارتفاع ساختمان‌ها دسته‌بندی شد. نتایج کلیدی به دست آمده نشان داد که مدل یادگیری عمیق پیشنهادی در مرحله نخست با دقت بالایی تغییرات را شناسایی کرده و به‌طور محسوسی عملکرد بهتری نسبت به روش‌های کلاسیک یادشده از خود نشان داده است. بدین ترتیب، مدل توانست مناطق دچار تغییر را حتی در مناطق پیچیده شهری به‌خوبی تفکیک کند این درحالی است که روش‌های سنتی در آن با چالش مواجه بودند. همچنین در مرحله دوم، مدل توسعه‌یافته قادر بود نوع تغییر رخ‌داده در هر ساختمان را به‌درستی طبقه‌بندی نماید و نمونه‌های مربوط به ساخت‌وسازهای جدید، تخریب‌ها و تغییرات ارتفاعی را با صحت قابل قبولی تشخیص دهد. این عملکرد چندکلاسه نشان‌دهنده توانایی مدل نه تنها در آشکارسازی وقوع

- Transactions on Geoscience and Remote Sensing. 2013;52(1):406-17. DOI:10.1109/TGRS.2013.2240692.
- [16] Wen D, Huang X, Zhang A, Ke X. Monitoring 3D building change and urban redevelopment patterns in inner city areas of Chinese megacities using multi-view satellite imagery. *Remote Sensing*. 2019;11(7):763. DOI:10.3390/rs11070763.
- [17] Hussain M, Chen D, Cheng A, Wei H, Stanley D. Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of photogrammetry and remote sensing*. 2013;80:91-106. DOI:10.1016/j.isprsjprs.2013.03.006.
- [18] Gao M, Qi D, Mu H, Chen J. A transfer residual neural network based on ResNet-34 for detection of wood knot defects. *Forests*. 2021;12(2):212. DOI:10.3390/f12020212.
- [19] Brahim E, Amri E, Barhoumi W, Bouzidi S. Fusion of UNet and ResNet decisions for change detection using low and high spectral resolution images. *Signal, Image and Video Processing*. 2024;18(Suppl 1):695-702. DOI:10.1007/s11760-024-03185-2.
- [20] Zhu XX, Tuia D, Mou L, Xia G-S, Zhang L, Xu F, et al. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE geoscience and remote sensing magazine*. 2017;5(4):8-36. DOI:10.1109/MGRS.2017.2762307.
- [21] Zhang J, Lin X. Advances in fusion of optical imagery and LiDAR point cloud applied to photogrammetry and remote sensing. *International Journal of Image and Data Fusion*. 2017;8(1):1-31. DOI:10.1080/19479832.2016.1160960.
- [22] Nanni L, Paci M, Brahnam S, Lumini A. Comparison of different image data augmentation approaches. *Journal of imaging*. 2021;7(12):254. DOI:10.3390/jimaging7120254.
- [23] Van Dyk DA, Meng X-L. The art of data augmentation. *Journal of Computational and Graphical Statistics*. 2001;10(1):1-50. DOI:10.1198/10618600152418584.
- [24] Stoyanov D, Taylor Z, Carneiro G, Syeda-Mahmood T, Martel A, Maier-Hein L, et al. Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th international workshop, dlmia 2018, and 8th international workshop, ml-cds 2018, held in conjunction with miccai 2018, granada, spain, september 20, 2018, proceedings: Springer; 2018. DOI:10.1007/978-3-319-67558-9.
- [25] Zhang Z, editor Improved adam optimizer for deep neural networks. 2018 IEEE/ACM 26th international symposium on quality of service (IWQoS); 2018: IEEE. DOI:10.1109/IWQoS.2018.8624183.
- [26] He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2016. DOI:10.1109/CVPR.2016.90.
- [27] Wu S, Zhong S, Liu Y. Deep residual learning for image steganalysis. *Multimed Tools Appl*. 2017:1-17. DOI:10.1007/s11042-017-4440-4.
- [3] Jensen JR, Im J. Remote sensing change detection in urban environments. *Geo-spatial technologies in urban environments: policy, practice, and pixels*: Springer; 2007. p. 7-31. DOI:10.1007/978-3-540-69417-5_2.
- [4] Gstaiger V, Tian J, Kiefl R, Kurz F. 2D vs. 3D change detection using aerial imagery to support crisis management of large-scale events. *Remote Sensing*. 2018;10(12):2054. DOI:10.3390/rs10122054.
- [5] Champion N, Boldo D, Pierrot-Deseilligny M, Stamon G. 2D building change detection from high resolution satellite imagery: A two-step hierarchical method based on 3D invariant primitives. *Pattern Recognition Letters*. 2010;31(10):1138-47. DOI:10.1016/j.patrec.2009.10.012.
- [6] Qin R. Change detection on LOD 2 building models with very high resolution spaceborne stereo imagery. *ISPRS journal of photogrammetry and remote sensing*. 2014;96:179-92. DOI:10.1016/j.isprsjprs.2014.07.007.
- [7] Qin R, Tian J, Reinartz P. 3D change detection—approaches and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2016;122:41-56. DOI:10.1016/j.isprsjprs.2016.09.013.
- [8] Jiang H, Peng M, Zhong Y, Xie H, Hao Z, Lin J, et al. A survey on deep learning-based change detection from high-resolution remote sensing images. *Remote Sensing*. 2022;14(7):1552. DOI:10.3390/rs14071552.
- [9] Song A, Choi J. Fully convolutional networks with multiscale 3D filters and transfer learning for change detection in high spatial resolution satellite images. *Remote Sensing*. 2020;12(5):799. DOI:10.3390/rs12050799.
- [10] Huang L, Tian Q, Tang B-H, Le W, Wang M, Ma X. Siam-EMNet: A Siamese EfficientNet–MANet Network for Building Change Detection in Very High Resolution Images. *Remote Sensing*. 2023;15(16):3972. DOI:10.3390/rs15163972.
- [11] Peng D, Zhang Y, Guan H. End-to-end change detection for high resolution satellite images using improved UNet++. *Remote Sensing*. 2019;11(11):1382. DOI:10.3390/rs11111382.
- [12] Lv Z, Huang H, Gao L, Benediktsson JA, Zhao M, Shi C. Simple multiscale UNet for change detection with heterogeneous remote sensing images. *IEEE Geoscience and Remote Sensing Letters*. 2022;19:1-5. DOI:10.1109/LGRS.2022.3173300.
- [13] Wang Z, Pan Z, Hu Y, Lei B. FRCD: Feature Refine Change Detection Network for Remote Sensing Images. *IEEE Geoscience and Remote Sensing Letters*. 2023. DOI:10.1109/LGRS.2023.3303200.
- [14] Meshkini K, Bovolo F, Bruzzone L. A 3D CNN approach for change detection in HR satellite image time series based on a pretrained 2D CNN. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2022;43:143-50. DOI:10.5194/isprs-archives-XLIII-B3-2022-143-2022.
- [15] Tian J, Cui S, Reinartz P. Building change detection based on satellite stereo imagery and digital surface models. *IEEE*

- [40] Belgiu M, Drăguț L. Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*. 2016;114:24-31. DOI:10.1016/j.isprsjprs.2016.01.011.
- [41] Lu D, Mausel P, Brondizio E, Moran E. Change detection techniques. *International journal of remote sensing*. 2004;25(12):2365-401. DOI:10.1080/0143116031000139863.
- [42] JST C, editor Change detection from a street image pair using cnn features and superpixel segmentation. *Proc Brit Mach Vis Conf*; 2015. DOI:10.5244/C.29.61.
- [43] Maggiori E, Tarabalka Y, Charpiat G, Alliez P. Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Transactions on geoscience and remote sensing*. 2016;55(2):645-57. DOI:10.1109/TGRS.2016.2612821.
- [44] Daudt RC, Le Saux B, Boulch A, editors. Fully convolutional siamese networks for change detection. 2018 25th IEEE international conference on image processing (ICIP); 2018: IEEE. DOI:10.1109/ICIP.2018.8451652.
- [45] Gharibbafghi Z, Tian J, Reinartz P, editors. Superpixel-based 3D building model refinement and change detection, using VHR stereo satellite imagery. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*; 2019. DOI:10.1109/CVPRW.2019.00069.
- [46] Li J, Zhu S, Gao Y, Zhang G, Xu Y. Change detection for high-resolution remote sensing images based on a multi-scale attention Siamese network. *Remote Sensing*. 2022;14(14):3464. DOI:10.3390/rs14143464.
- [47] Gholamrezaie H, Hasanlou M, Amani M, Mirmazloumi M. Automatic Mapping of Burned Areas Using Landsat 8 Time-Series Images in Google Earth Engine: A Case Study from Iran. *Remote Sensing*. 2022. DOI:10.3390/rs14246376.
- [48] Coletta, V., et al. 3DCD: A new dataset for 2D and 3D change detection using deep learning techniques. 2022, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 43: 1349-1354. DOI:10.5194/isprs-archives-XLIII-B3-2022-1349-2022.
- [28] Cardoso MJ, Arbel T, Carneiro G, Syeda-Mahmood T, Tavares JMR, Moradi M, et al. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings: Springer; 2017. DOI:10.1007/978-3-319-67558-9.
- [29] Zhang X, Yue Y, Gao W, Yun S, Su Q, Yin H, et al. DifUnet++: A satellite images change detection network based on Unet++ and differential pyramid. *IEEE Geoscience and Remote Sensing Letters*. 2021;19:1-5. DOI:10.1109/LGRS.2021.3049370.
- [30] Deng L, Wang Y, Lan Q, Chen F. Remote sensing image building change detection based on efficient-UNet++. *Journal of Applied Remote Sensing*. 2023;17(3):034501-. DOI:10.1117/1.JRS.17.034501.
- [31] Berman M, Triki AR, Blaschko MB, editors. The lovasz-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2018. DOI: 10.1109/CVPR.2018.00464.
- [32] Bousias Alexakis E, Armenakis C. Evaluation of UNet and UNet++ architectures in high resolution image change detection applications. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2020;43:1507-14. DOI:10.5194/isprs-archives-XLIII-B3-2020-1507-2020.
- [33] Kingma D, Ba J. Adam: A method for stochastic optimization: *International Conference on Learning Representations* 5. 2015. DOI:10.48550/arXiv.1412.6980.
- [34] Kingma DP, Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:14126980*. 2014. DOI:10.48550/arXiv.1412.6980.
- [35] Ji S, Wei S, Lu M. Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set. *IEEE Transactions on geoscience and remote sensing*. 2018;57(1):574-86. DOI:10.1109/TGRS.2018.2858817.
- [36] Mikołajczyk A, Grochowski M, editors. Data augmentation for improving deep learning in image classification problem. 2018 *international interdisciplinary PhD workshop (IIPhDW)*; 2018: IEEE. DOI:10.1109/IIPHDW.2018.8388338.
- [37] Chen H, Li W, Shi Z. Adversarial instance augmentation for building change detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*. 2021;60:1-16. DOI:10.1109/TGRS.2021.3066802.
- [38] Wang J, Perez L. The effectiveness of data augmentation in image classification using deep learning. *Convolutional Neural Networks Vis Recognit*. 2017;11(2017):1-8. DOI:10.48550/arXiv.1712.04621.
- [39] Radke RJ, Andra S, Al-Kofahi O, Roysam B. Image change detection algorithms: a systematic survey. *IEEE transactions on image processing*. 2005;14(3):294-307. DOI:10.1109/TIP.2004.838698.

معرفی نویسندگان

AUTHOR(S) BIOSKETCHES



علیرضا ابراهیمی دارای مدرک کارشناسی در رشته مهندسی نقشه برداری و اطلاعات مکانی و دانشجوی کارشناسی ارشد سنجش از دور از دانشگاه تهران است. وی از سال ۱۴۰۱ تاکنون در زمینه هوش مصنوعی و سنجش از دور فعالیت دارد.

Ebrahimi, A. MS.c student at the School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

✉ ebrahimi.alireza@ut.ac.ir

منابع طبیعی و پایش محیط زیست می باشد. وی سرپرست آزمایشگاه سنجش از دور چندطیفی و ابرطیفی دانشگاه تهران است و علاوه بر تدریس دروس تخصصی، در نگارش جزوات و معرفی منابع علمی متنوع در این حوزه نیز فعال می باشد.

Hasanlou, M. Associate Professor at the School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

✉ hasanlou@ut.ac.ir



مهدی حسنلو دانشیار و عضو هیئت علمی دانشکده مهندسی نقشه برداری و اطلاعات مکانی، در دانشگاه تهران می باشند. زمینه های تخصصی ایشان به طور کلی شامل توسعه و به کارگیری الگوریتم های پیشرفته در سنجش از دور، تحلیل تغییرات محیطی، کاربردهای علمی و عملی داده های ماهواره ای در مدیریت

Citation (Vancouver): Ebrahimi A, Hasanlou M. [Multi-Class 3D Change Detection in Urban Areas Using UNet++]. *J. RS. GEOINF. RES.* 2025; 3(2): 363-378

 <https://doi.org/10.22061/jrsgr.2026.12415.1107>



COPYRIGHTS

© 2025 The Author(s). This is an open-access article distributed under the terms and conditions of the Creative Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)
(<https://creativecommons.org/licenses/by-nc/4.0/>)