



# بهبود الگوریتم تناظریابی SIFT جهت تطبیق تصاویر ماهواره‌ای مرئی با استفاده از شبکه عصبی عمیق دوقلو

احمد رضا زارعی<sup>۱</sup>، پیمان معلم<sup>۲\*</sup>

دانش‌آموخته کارشناسی ارشد رشته مهندسی برق، دانشکده فنی و مهندسی، دانشگاه اصفهان، اصفهان، ایران<sup>۱</sup>

استاد گروه مهندسی برق، دانشکده فنی و مهندسی، دانشگاه اصفهان، اصفهان، ایران<sup>۲\*</sup>

## چکیده

تطبیق تصاویر سنجش‌ازدور یک مرحله اساسی و پایه‌ای در بسیاری از کاربردهای پردازش تصویر است؛ برخلاف تصاویر معمولی، تصاویر سنجش‌ازدور به‌طور معمول تحت تأثیر تغییرات پیچیده پس‌زمینه هستند که باعث دشواری تطبیق این دسته از تصاویر می‌شود. تصاویر مورد استفاده علاوه بر تغییرات شدید و غیرخطی در پس‌زمینه، شامل چالش‌هایی مانند تفاوت مقیاس، چرخش و زاویه دید متفاوت است. یکی از الگوریتم‌های پرکاربرد برای یافتن نقاط متناظر میان تصاویر، الگوریتم SIFT است که استفاده از این الگوریتم برای تطبیق تصاویر با ویژگی‌های گفته‌شده، در بسیاری از مواقع باعث تولید تطابق‌های اشتباه فراوان و تطابق‌های صحیح کم می‌شود؛ از طرفی روش‌های مبتنی بر یادگیری عمیق قادر به استخراج ویژگی‌های سطح متوسط و بالا و مقایسه آن‌ها برای تطبیق تصاویرند. با الهام از توانایی‌ها و پیشرفت‌های انجام‌شده در یادگیری عمیق، روشی برای استفاده هم‌زمان از الگوریتم SIFT و شبکه‌های عصبی عمیق برای تطبیق تصاویر سنجش‌ازدور معرفی شده است؛ بر اساس آزمایش‌های صورت گرفته بر روی مجموعه‌ای از تصاویر سنجش‌ازدور شامل ۳۵ جفت تصویر و مقایسه نتایج نسبت به الگوریتم SIFT و الگوریتم‌های مبتنی بر یادگیری عمیق، روش پیشنهادی توانسته با کاهش تطابق‌های اشتباه و افزایش تطابق‌های صحیح، به دقت ۸۴۹ درصد برسد.

واژگان کلیدی: الگوریتم SIFT، تطبیق تصاویر، سنجش‌ازدور، شبکه‌های عصبی کانولوشنی دوقلو.

## Improvement of SIFT matching algorithm for matching visible satellite images using Siamese deep neural network

Ahmad Reza Zarei<sup>1</sup> & Payman Moallem<sup>2\*</sup>

Master, Department of Electrical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran<sup>1</sup>

Professor, Department of Electrical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran<sup>2\*</sup>

### Abstract

Matching remote sensing images is a fundamental step in many image processing applications. Unlike regular images, remote sensing images often undergo complex and nonlinear background changes, making them difficult to match. They also pose challenges such as scale variations, rotation, and different viewing angles. One commonly used method for finding corresponding points between images is the Scale-Invariant Feature Transform (SIFT) algorithm; however, it often produces many incorrect matches when applied to such data. In contrast, deep learning-based approaches can extract and compare medium and high-level features for more accurate matching. Inspired by these advances, this work introduces a method that combines the SIFT algorithm with a Siamese deep neural network to improve the matching of remote sensing images.

The proposed method modifies the conventional SIFT by adjusting its parameters to increase the proportion of correct to incorrect correspondences. After keypoints are extracted and described, initial correspondences are established. Then, for each matched point, a local patch is extracted based on the keypoint's position, scale, and orientation. These patch pairs are input to a trained Siamese network that estimates the probability of a correct match. Matches with confidence below a threshold are rejected. This hybrid approach leverages the strengths of both traditional and deep learning-based techniques to enhance accuracy. The proposed approach introduces several key innovations, including

\* Corresponding author

\* نویسنده عهده‌دار مکاتبات

optimized keypoint extraction to maximize true matches, patch-based feature representation aligned with local image geometry, and a neural network-based verification step to suppress incorrect matches. Based on experiments conducted on a dataset of 35 pairs of remote sensing images, and comparing the results with the SIFT algorithm and deep learning-based methods, the proposed approach achieved an accuracy of 0.849 by reducing false matches and increasing correct ones.

**Keywords:** Image Matching, Remote Sensing, SIFT Algorithm, Deep Learning, Siamese Convolutional Neural Networks.

نخستین روش برای یافتن خودکار نقاط گوشه در تصویر را می‌توان روش موراک<sup>۲</sup> [۲] در نظر گرفت. روش موراک در صورت وجود نوفه از حساسیت بالایی برخوردار است و ضعف دیگر این روش این است که تنها در چهار جهت اصلی به بررسی تغییرات شدت روشنایی می‌پردازد؛ در صورتی که برای نقاط گوشه باید در تمامی جهتها تغییرات شدت روشنایی زیاد باشد.

برای برطرف کردن نقاط ضعف روش موراک، تشخیص‌گر شناخته‌شده و معروفی به نام هریس<sup>۳</sup> [۳] معرفی شد. در این روش برای شناسایی گوشه‌ها از یک پنجره هموار برای مثال پنجره گوسی استفاده می‌شود و با استفاده از بسط تیلور این پنجره را در تمامی جهتها حرکت داده و مقادیر مربوط به شدت روشنایی را برای تمامی این جهتها به دست می‌آورد. روش هریس مستقل از چرخش و روشنایی است و قابلیت تکرارپذیری و تمایز قابل قبولی دارد. ضعف اصلی روش هریس را می‌توان حساسیت آن نسبت به تغییرات مقیاس دانست.

در پژوهش [۴] الگوریتم شناخته‌شده و بسیار پرکاربرد تحت عنوان ویژگی‌های مستقل از تغییر مقیاس که با نام الگوریتم SIFT<sup>۴</sup> شناخته می‌شود، معرفی شد. الگوریتم SIFT هم شناساگر و هم توصیف‌گر نقاط است. مراحل الگوریتم SIFT را می‌توان به سه مرحله کلی یافتن نقاط کلیدی، ایجاد توصیف‌گر برای نقاط کلیدی و تطابق نقاط با استفاده از توصیف‌گرها تقسیم کرد. الگوریتم SIFT با وجود کاربردهای بسیاری که در زمینه‌های مختلف پردازش تصویری دارد، دارای محدودیت‌هایی است که متناسب با نوع تصاویر و کاربرد، بهبودهای زیادی برای این الگوریتم ارائه شده است.

در پژوهش [۵] الگوریتم SURF<sup>۵</sup> معرفی شد که این الگوریتم نسخه‌ای بهبودیافته از الگوریتم SIFT است که دارای سرعت بالاتری است و توصیف‌گر آن نسبت به الگوریتم SIFT ابعاد کوچک‌تری دارد. مراحل الگوریتم SURF همانند الگوریتم SIFT است؛ اما در قسمت‌های مربوط به اختصاص جهت، ایجاد توصیف‌گر و تطابق نقاط دارای تفاوت‌هایی با الگوریتم SIFT است؛ از جمله بهبودهای ارائه‌شده برای الگوریتم SIFT در زمینه تطبیق تصاویر سنجش‌ازدور می‌توان به دو الگوریتم SAR-SIFT [۶] و BFSIFT [۷] اشاره کرد؛

<sup>۲</sup> Moravec

<sup>۳</sup> Harris

<sup>۴</sup> Scale-Invariant Features Transform

<sup>۵</sup> Speeded-Up Robust features

## ۱- مقدمه

تطبیق تصاویر سنجش‌ازدور فرایندی است که در آن تصاویر دریافت‌شده از منابع مختلف مانند ماهواره، هواپیما و هواپیماهای بدون سرنشین با یکدیگر مقایسه می‌شوند تا بهترین تطبیق مکانی بین دو یا چند تصویر تعیین شود. باتوجه به ماهیت تصاویر سنجش‌ازدور، تطبیق این تصاویر دارای چالش‌های بسیاری است؛ از جمله کاربردهای تطبیق تصاویر سنجش‌ازدور شامل نظارت شهری، کاربردهای نظامی، تشخیص اهداف، نظارت بر تغییرات سطح زمین مانند جنگل‌زدایی و بیابان‌زایی و غیره است.

روش‌های متنوع و بسیاری برای تطبیق تصاویر مورد استفاده قرار می‌گیرد که در سال‌های گذشته با توسعه و پیشرفت روش‌های مبتنی بر یادگیری عمیق، بهبود عملکرد الگوریتم‌های تطبیق و تنوع زیادی در روش‌ها به وجود آمده است. بسیاری از روش‌ها از ویژگی‌های سطح پایین تصویر استفاده می‌کنند که در این روش‌ها بسیاری از ویژگی‌های سطح متوسط و بالا بدون استفاده می‌ماند؛ در مقابل ویژگی‌های به دست آمده از طریق شبکه‌های عصبی کانولوشنی<sup>۱</sup> (CNN) می‌توانند به خوبی ویژگی‌های سطح متوسط و بالا را استخراج کرده و از آن‌ها استفاده کنند.

روش‌های تطبیق تصویر را می‌توان به دو دسته روش‌های مبتنی بر ناحیه و روش‌های مبتنی بر ویژگی تقسیم کرد. روش‌های مبتنی بر ناحیه بیشتر از شدت پیکسل‌های دو تصویر و اندازه‌گیری میزان شباهت آن‌ها برای تطبیق تصویر استفاده می‌کند؛ در مقابل روش‌های مبتنی بر ویژگی با استخراج ویژگی‌های مناسب و توصیف این ویژگی‌ها، تطبیق ویژگی‌ها، تخمین تابع تبدیل با استفاده از ویژگی‌های تطبیق‌یافته و در نهایت تغییر شکل و اندازه تصویر، وظیفه تطبیق تصویر را انجام می‌دهند [۱].

در بیشتر موارد روش‌های تطبیق مبتنی بر ویژگی کارایی بهتری دارند و می‌توانند تغییر شکل‌های هندسی را به نحو بهتری کنترل کنند. این دسته از روش‌ها بر اساس تشخیص ویژگی‌ها، توصیف ویژگی‌ها، تطبیق ویژگی‌ها و تخمین مدل هندسی استوار هستند که هر کدام از این موارد دارای چالش‌های مخصوص به خود است [۱].

گوشه‌ها را می‌توان به عنوان مهم‌ترین و پرکاربردترین ویژگی‌های نقطه‌ای برای تطبیق تصاویر در نظر گرفت.

<sup>۱</sup> Convolutional Neural Network

استفاده از ویژگی‌های محلی و عمیق انجام شده‌است. ابتدا هر دو تصویر مرجع و هدف را به‌عنوان ورودی به شبکه VGG16 داده شده و نقاط متناظر اولیه با محاسبه فاصله اقلیدسی ویژگی‌ها و سپس استفاده از روشی مشابه با SIFT، تعیین می‌شوند. در ادامه ماتریس انتقال اولیه تخمین زده می‌شود؛ در نهایت با استخراج توصیف‌گر SIFT برای هر نقطه و تطبیق آن‌ها، ماتریس انتقال دقیق بر اساس نقاط متناظر نهایی تعیین می‌شود. نتایج به‌دست‌آمده با روش‌هایی که تنها از ویژگی‌های محلی استفاده می‌کنند مقایسه شده‌است که نشان‌دهنده دقت بالاتر و تعداد تطابق صحیح بیشتر در این روش است.

در پژوهش [۲۰] یک چهارچوب مبتنی بر شبکه‌های عصبی کانولوشنی شبه‌سیامی<sup>۱</sup> برای تشخیص تکه‌تصویرهای متناظر در تصاویر سنجنش‌ازدور نوری با وضوح بسیار بالا (VHR<sup>۲</sup>) و تصاویر SAR معرفی شده‌است؛ به‌دلیل تفاوت‌های فراوان میان تصاویر SAR و بصری، در این روش استفاده از شبکه کانولوشنی شبه‌سیامی که از دو شاخه به‌طور کامل مجزا و بدون اشتراک‌گذاری وزن‌ها استفاده می‌کنند، پیشنهاد شده‌است. در این پژوهش نشان داده شد که با افزایش ابعاد تکه‌های تصویری، شبکه عصبی کانولوشنی با دقت بالاتری می‌تواند تناظر آن‌ها را تشخیص دهد.

از دیگر روش‌های مبتنی بر یادگیری عمیق می‌توان به مراجع [۲۱] تا [۲۶] اشاره کرد. در پژوهش [۲۱] با بهره‌گیری از ماژول توجه بلوک کانولوشنی CBAM<sup>۳</sup> روشی ارائه شده‌است که دقت تطبیق را تا ۴۹.۲ درصد افزایش داده و نسبت به نسخه‌های پیشین تا ۳۰ درصد بهبود داشته‌است؛ همچنین در پژوهش [۲۲] با معرفی شبکه LM-Net که یک معماری سبک‌وزن است، سعی شد پیچیدگی محاسباتی بدون افت محسوس در عملکرد، کاهش داده شود. در پژوهش [۲۳] مدل LightGlue معرفی شده‌است که بر پایه یک شبکه عصبی عمیق طراحی شده‌است. هدف این روش تطبیق ویژگی‌های محلی آموزش‌دیده است و علاوه‌براین از نظر مصرف حافظه و سرعت محاسبات نیز عملکرد مناسبی دارد. از ویژگی‌های مهم این مدل، سازگاری پویای آن با میزان دشواری تطبیق تصاویر است؛ به‌طوری که در شرایطی با هم‌پوشانی بالا یا تفاوت‌های کم میان تصاویر، تطبیق با سرعت بیشتری انجام می‌شود.

یکی از مهم‌ترین چالش‌های استفاده از روش‌های مبتنی بر یادگیری و به‌ویژه یادگیری عمیق، محدودیت در تعداد داده‌های آموزشی است. برای غلبه بر چالش کمبود داده‌های آموزشی برچسب‌خورده، در پژوهش [۲۴] مدل LTFormer

همچنین از دیگر بهبودهای ارائه‌شده برای الگوریتم SIFT، می‌توان از پژوهش‌های [۸] تا [۱۲] نام برد.

در [۱۳]، پژوهش‌گران با بهره‌گیری از خروجی لایه‌های خاصی از شبکه VGG-16 معرفی‌شده در پژوهش [۱۴]، یک توصیف‌گر ویژگی طراحی کردند. در این روش، لایه‌های کانولوشن ساده، ادغام (Pooling) و لایه‌های به‌طور کامل متصل، بر پایه پالایه‌های تجسم پیچیدگی و نتایج حاصل از آزمایش‌های ردیابی و ارزیابی خطا، به‌گونه‌ای انتخاب شدند که خروجی یک لایه خاص به‌عنوان توصیف‌گر ویژگی استفاده شود. در راستای همین رویکرد، چندین شبکه دیگر نیز برای استخراج توصیف‌گرهای ویژگی مورد بررسی و انتخاب قرار گرفتند. در پژوهش [۱۵] تطبیق تصاویر سنجنش‌ازدور با استفاده از یک شبکه عصبی کانولوشنی دوقلو انجام می‌شود. تعداد لایه‌های به‌کاررفته در شبکه دوقلو متناسب با عواملی که در تطبیق تصویر تأثیرگذارند تعیین شده‌است؛ این روش می‌تواند تطبیق‌های با توزیع مناسب و با دقت و صحت مطابقت بالاتری برای تصاویر سنجنش‌ازدور دارای پس‌زمینه پیچیده به‌دست آورد. این روش در مواقعی که از کیفیت مکانی تصاویر مرجع و هدف اطلاعی در دست نیست، نمی‌تواند عملکرد خوبی داشته باشد؛ علاوه‌براین، این روش ممکن است، برای تصاویری که دارای بافت تکرار‌شونده و ضعیف‌اند، مناسب نباشد.

در پژوهش [۱۶] با آموزش یک شبکه عصبی کانولوشنی به نام DescNet و ایجاد توصیف‌گر پایدار برای ویژگی‌ها با استفاده از این شبکه، سعی در تطبیق تصاویر سنجنش‌ازدور دارد. در این روش ابتدا با استفاده از DoG نقاط کلیدی تصاویر مرجع و هدف استخراج می‌شود؛ سپس اطراف هر کدام از این نقاط یک تکه تصویر به مرکزیت نقطه کلیدی و متناسب با مقیاس و جهت آن نقطه استخراج می‌شود. در ادامه با استفاده از شبکه کانولوشنی آموزش‌داده‌شده برای هر کدام از نقاط توصیف‌گر محاسبه می‌شود؛ در نهایت توصیف‌گرهای به‌دست‌آمده برای نقاط کلیدی با روشی مشابه با روش استفاده‌شده در الگوریتم SIFT مورد مقایسه و مطابقت قرار می‌گیرند و نقاط متناظر در دو تصویر تعیین می‌شوند.

در پژوهش [۱۷] برای ثبت تصاویر سنجنش‌ازدور از ترکیب ویژگی‌های الگوریتم SIFT و شبکه‌های عصبی کانولوشنی استفاده می‌کند. در این پژوهش برای استخراج ویژگی‌های CNN از شبکه VGG16 استفاده می‌شود. برای هر نقطه کلیدی توصیف‌گر SIFT و توصیف‌گر مبتنی بر شبکه عصبی استخراج می‌شود و سپس با استفاده جمع وزنی فاصله اقلیدسی مربوط به توصیف‌گر SIFT و شبکه عصبی عمیق، میزان شباهت نقاط محاسبه می‌شود. در ادامه همانند روش استفاده‌شده در الگوریتم PSO-SIFT [۱۸] ثبت تصاویر انجام می‌شود. در پژوهش [۱۹] ثبت تصاویر سنجنش‌ازدور با

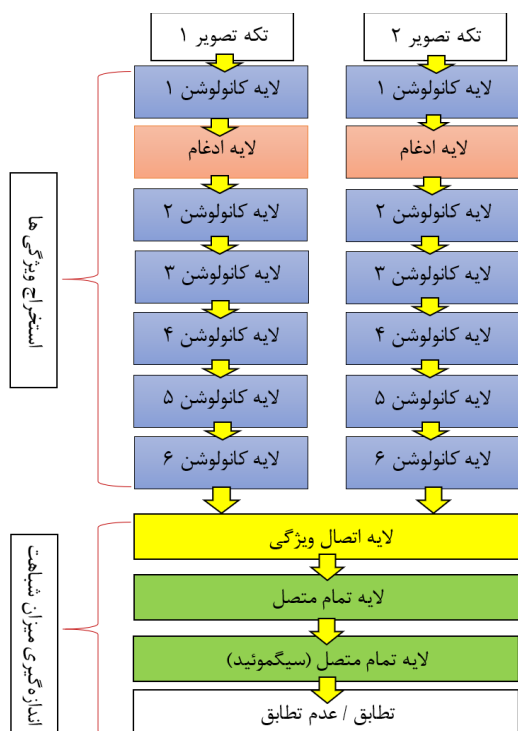
<sup>۱</sup> Siamese

<sup>۲</sup> Very High Resolution

<sup>۳</sup> Convolutional Block Attention Module

## ۲- شبکه‌های عصبی کانولوشنی دوقلو

شبکه‌های عصبی دوقلو برای نخستین بار در اوایل دهه ۱۹۹۰ در پژوهش [۲۷] برای تشخیص اعتبار امضا معرفی شد. این شبکه‌ها از دو شاخه موازی و یکسان استفاده می‌کنند که توانایی یادگیری و استخراج ویژگی‌های ورودی به شبکه را دارند. خروجی ایجادشده در شبکه‌های عصبی از نوع دوقلو را می‌توان به عنوان معیاری برای اندازه‌گیری میزان شباهت ورودی‌ها در نظر گرفت. استفاده از این ساختار دارای کاربردهای فراوانی در پردازش تصاویر و ویدیو، پردازش صوت، رباتیک و غیره است. کاربردهای این نوع شبکه در پردازش تصویر شامل شناخت چهره<sup>۱</sup>، طبقه‌بندی تصاویر، تطبیق تصاویر، اندازه‌گیری تغییرات<sup>۲</sup> و غیره است؛ هنگامی که بتوان بازنمایی از این دو شیء را در یک فضای ویژگی مشخص نشان داد، می‌توان از رویکردهای استاندارد هندسی و آماری مانند فاصله اقلیدسی و یا سایر موارد استفاده کرد؛ اما وقتی دو شیء از داده‌هایی از فضاهای ویژگی متفاوتی باشند، برای پیش‌بینی و نمایش نیاز است از روش‌های دیگری استفاده شود و در صورتی که هدف اندازه‌گیری شباهت میان اشیایی باشد که در میان داده‌های آموزشی نمونه‌ای از آن‌ها وجود ندارد، شبکه‌های عصبی دوقلو می‌توانند گزینه مناسبی باشند [۲۸].



(شکل-۱): ساختار شبکه دوقلو برای تطبیق تصاویر [۱۵].  
(Figure-1): Siamese convolutional network structure for image matching [15].

شکل (۱) نشان‌دهنده یک شبکه دوقلو است که در این پژوهش مورد استفاده قرار گرفته است. این ساختار

معرفی شده است که با بهره‌گیری از رویکرد خودنظارتی، توصیف‌گرهای مقاوم و قابل اطمینانی تولید کرده است. در این روش بدون نیاز به برچسب‌گذاری گسترده، تطبیق تصویر را تسهیل می‌کند.

با وجود این پیشرفت‌ها، ارزیابی‌های جامع نشان می‌دهند که در برخی سناریوها، به‌ویژه در شرایطی با منابع محاسباتی محدود، روش‌های کلاسیک نظیر SIFT همچنان از نظر پایداری و سرعت، عملکرد مناسبی دارند [۲۵]. در پژوهش [۲۶] یک رویکرد ترکیبی استخراج ویژگی مطرح شده است که در آن از سه نوع ویژگی، شامل دو نوع ویژگی سراسری، همچنین یک نوع از ویژگی‌های محلی دستی بهره برده و به‌منظور بهبود کارایی تشخیص صحنه‌ها، آن‌ها را با هم ترکیب کرده است.

مجموعه این یافته‌ها بیان‌گر آن است که اگرچه روش‌های مبتنی بر یادگیری عمیق نویدبخش هستند، اما برای دستیابی به عملکرد پایدار و همه‌جانبه، توسعه بیشتر و بهره‌گیری از رویکردهای ترکیبی همچنان ضروری است. در روش‌های موجود تطبیق تصویر، به‌ویژه در حوزه تصاویر سنجش‌ازدور، چالش‌هایی نظیر حساسیت به تغییرات غیرخطی روشنایی، استخراج ویژگی‌های ناپایدار در بافت‌های تکرارشونده و دقت پایین در تطابق‌های نهایی به چشم می‌خورد. روش‌های سنتی مانند SIFT با وجود مزایای زیاد، در مواجهه با این پیچیدگی‌ها ضعف دارند و روش‌های به‌طور کامل مبتنی بر یادگیری عمیق نیز برحسب معمول نیازمند داده‌های برچسب‌خورده گسترده و هزینه‌بر هستند؛ همچنین این روش‌ها برحسب معمول در مواقعی که تفاوت شدید میان دو تصویر وجود دارد؛ مانند تفاوت مقیاس و چرخش شدید، با شکست روبه‌رو می‌شوند یا دقت کمی دارند. در روش پیشنهادی این پژوهش، با تنظیم هدفمند پارامترهای الگوریتم SIFT، نقاط کلیدی باکیفیت‌تری استخراج شده و نسبت تطابق‌های درست افزایش یافته است؛ همچنین، با استخراج تکه‌تصاویر محلی حول نقاط متناظر و استفاده از شبکه عصبی دوقلو برای ارزیابی صحت تطابق، تطابق‌های نادرست حذف و دقت نهایی بهبود یافته است. این ترکیب بهینه از روش‌های کلاسیک و یادگیری عمیق، دقت و پایداری تطبیق تصاویر سنجش‌ازدور را افزایش چشم‌گیری داده است.

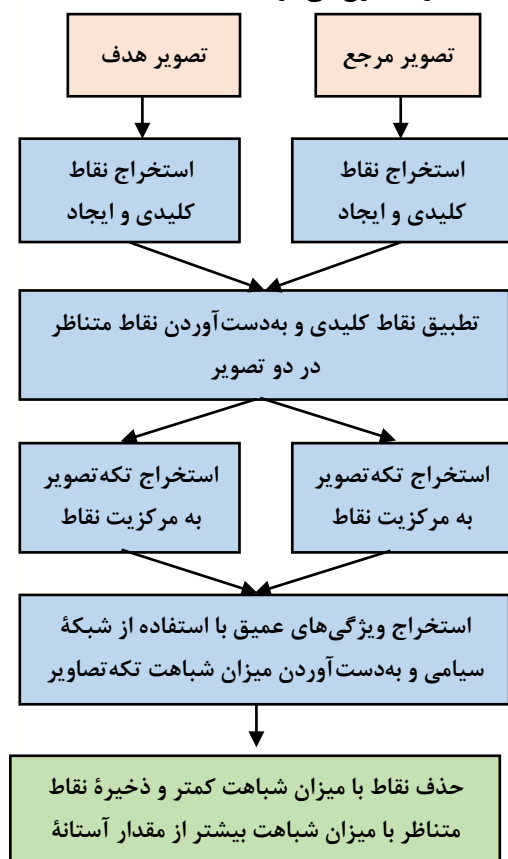
در بخش دوم به اختصار به ساختار شبکه عصبی کانولوشنی دوقلو پرداخته خواهد شد؛ سپس در بخش سوم، به تشریح الگوریتم پیشنهادی و مراحل آن پرداخته می‌شود. در بخش چهارم، با آموزش شبکه دوقلو و شبیه‌سازی الگوریتم پیشنهادی، به بررسی عملکرد روش پیشنهادی و مقایسه آن می‌پردازیم. در نهایت در بخش پنجم نتیجه‌گیری و پیشنهادها برای کارهای آینده ارائه خواهد شد.

<sup>1</sup> Face recognition

<sup>2</sup> Change detection



تصویری را تحلیل کرده و صحت تطابق آن‌ها را با احتمال عددی تخمین بزند؛ درنهایت، با اعمال آستانه بر خروجی شبکه دوقلو، تطابق‌های نادرست حذف شده و مجموعه‌ای دقیق‌تر از نقاط متناظر استخراج می‌شود.



(شکل-۲): روندنمای روش پیشنهادی برای تطبیق تصاویر

سنجش‌ازدور

(Figure-2): Process of the proposed method for matching of remote sensing images

این ترکیب از روش‌های کلاسیک و یادگیری عمیق باعث بهبود قابل توجه دقت در تطابق تصاویر ماهواره‌ای شده‌است.

### ۳-۱- استخراج نقاط کلیدی و ایجاد توصیف‌گر

در این مرحله با استفاده از الگوریتم SIFT، نقاط کلیدی را برای دو تصویر به‌دست می‌آوریم و سپس برای هر نقطه یک بردار توصیف‌گر ایجاد می‌کنیم. در الگوریتم SIFT نقاط کلیدی ناپایداری که از شدت نور پایینی برخوردارند حذف می‌شود. برای حذف این نقاط در الگوریتم SIFT پیشنهاد شد که نقاط اکسترممی که مقدار تابع  $DOG^1$  برای آن نقطه کمتر از مقدار آستانه  $0.03$  است، حذف شوند. این کار با وجود اینکه تعداد تطابق‌های اشتباه را کاهش می‌دهد، باعث کاهش تعداد تطابق‌های صحیح نیز می‌شود؛ بر اساس آزمایش‌های نشان‌داده‌شده در بخش چهارم و بررسی نتایج به‌دست‌آمده، با کاهش این مقدار می‌توان تعداد تطابق‌های صحیح را افزایش داد. در این مرحله تعداد تطابق‌های اشتباه نیز افزایش خواهد

<sup>1</sup> Difference of Gaussian

شامل دو قسمت کلی استخراج ویژگی‌ها و اندازه‌گیری میزان شباهت است.

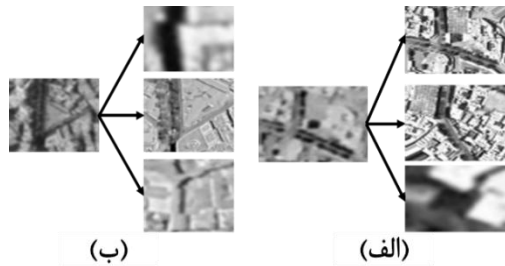
در قسمت استخراج ویژگی با استفاده از لایه‌های کانولوشنی ویژگی‌های هرکدام از ورودی‌ها استخراج می‌شود. در مرحله استخراج ویژگی با توجه به این که از هر دو شاخه موازی به‌طورکامل یکسان هستند، پس در صورت یکسان‌بودن ورودی‌ها، ویژگی‌های یکسانی نیز استخراج خواهد شد. برای اندازه‌گیری میزان شباهت ورودی‌ها، ویژگی‌های استخراج‌شده از طریق لایه‌های کانولوشنی با استفاده از لایه concatenate در کنار هم قرار گرفته و درنهایت با استفاده از لایه‌های تمام متصل خروجی نهایی که نشان‌دهنده میزان شباهت و یا تناظر و عدم تناظر ورودی‌ها است، ایجاد می‌شود.

### ۳- الگوریتم پیشنهادی

در این روش که مراحل آن در شکل (۲) نشان داده شده‌است، نخست برای هر دو تصویر به روشی مشابه با الگوریتم SIFT نقاط کلیدی را به‌دست آورده و برای هرکدام از این نقاط بردار توصیف‌گر محاسبه می‌شود. استخراج نقاط کلیدی در تصاویر از طریق الگوریتم SIFT انجام می‌شود با این تفاوت که با تغییر در پارامترهای این الگوریتم به‌دنبال تولید نقاط متناظر صحیح بیشتری خواهیم بود. پس از استخراج نقاط کلیدی و استخراج توصیف‌گر برای هر نقطه، این نقاط را به صورتی که در بخش‌های بعد گفته خواهد شد، مطابقت داده و نقاط متناظر برای دو تصویر به‌دست می‌آید. در ادامه برای نقاط متناظر یک تکه تصویر به مرکزیت نقطه کلیدی و متناسب با مقیاس و جهت محاسبه‌شده به‌وسیله الگوریتم SIFT استخراج خواهد شد. در مرحله آخر با استفاده از شبکه دوقلو آموزش‌داده‌شده صحت تناظر را بررسی و در صورتی که احتمال متناظر بودن دو نقطه از یک مقدار آستانه کمتر باشد این تناظر حذف خواهد شد. در ادامه هرکدام از این مراحل جداگانه بررسی خواهند شد.

به‌اختصار، نوآوری‌های روش پیشنهادی به‌صورت زیر قابل بیان است: نخست، برخلاف استفاده مرسوم از الگوریتم SIFT با پارامترهای پیش‌فرض، در این روش با تغییر هدفمند پارامترهای استخراج ویژگی، تلاش شد تا نقاط کلیدی به‌گونه‌ای انتخاب شوند که نسبت تطابق‌های درست به تطابق‌های اشتباه افزایش یابد یا به بیان دیگر، تطابق‌های درست افزایش یافته و درعین حال تطابق‌های اشتباه افزایش کمتری داشته باشد. دوم، پس از استخراج نقاط متناظر اولیه، برای هر نقطه یک تکه تصویر با در نظر گرفتن موقعیت، مقیاس و جهت محلی محاسبه‌شده به‌وسیله SIFT جدا می‌شود. این کار باعث می‌شود ویژگی‌های بافت و ساختار محلی تصویر با دقت بیشتری به مدل یادگیری داده شود. سوم، از یک شبکه عصبی دوقلوی آموزش‌دیده استفاده شده تا جفت تکه‌های

از تصویر هدف متناظر است؛ در صورتی که نخستین نقطه داوطلب فاصله اقلیدسی کمتری دارد که در این حالت نیز روش تطابق الگوریتم SIFT نقطه متناظری تشخیص نمی‌دهد.



(شکل-۳): تکه‌های تصویر اطراف دو مورد از نقاط کلیدی تصویر مرجع به همراه نزدیک‌ترین نقاط کلیدی از تصویر هدف  
(Figure-3): Image patches around two keypoints of the reference image along with the closest keypoints of the target image

برای حل این مشکل ابتدا با کاهش مقدار آستانه  $T$  بر اساس آزمایش‌های صورت گرفته در بخش چهار به دنبال افزایش تعداد تطابق‌های صحیح خواهیم بود؛ سپس به جای استفاده از نسبت فاصله اقلیدسی میان نقطه کلیدی از تصویر مرجع و دو نقطه کلیدی با کمترین فاصله نسبت به این نقطه از تصویر هدف، این نسبت برای چندین نقطه محاسبه خواهد شد. اگر  $P$  توصیف‌گر یکی از نقاط مربوط به تصویر مرجع،  $Q = [q_1, q_2, \dots, q_n]$  بردار شامل توصیف‌گرهای نقاط کلیدی تصویر هدف و  $D = [d_1, d_2, \dots, d_n]$  بردار فاصله نقاط  $Q$  نسبت به  $P$  باشد و به صورت صعودی مرتب شده باشد؛ آن‌گاه همانند الگوریتم شکل (۴) می‌توان چند نقطه از بردار  $Q$  را متناظر با  $P$  در نظر گرفت.

استفاده از الگوریتم نشان داده شده در شکل (۴) در عین حال که باعث افزایش تطابق‌های صحیح می‌شود ممکن است باعث شود در این مرحله برای یک نقطه کلیدی از تصویر مرجع چند نقطه کلیدی متناظر از تصویر هدف داشته باشیم و همین‌طور باعث افزایش تعداد تطابق‌های نادرست می‌شود که در مراحل بعد با استفاده از شبکه دوقلو صحت هر کدام از تناظرها بررسی می‌شود و در بالاترین حد، یکی از نقاط به عنوان نقطه متناظر باقی می‌ماند. تعداد بیشینه نقاط متناظر قابل قبول برای یک نقطه (پارامتر  $\max\_MP$ ) و تأثیر آن بر تعداد تطابق‌های صحیح و اشتباه با بررسی نتایج حاصل از آزمایشات نشان داده شده در بخش چهار تعیین می‌شود.

#### الگوریتم (۱):

- ۱: توصیف‌گر یکی از نقاط کلیدی تصویر مرجع  $P$ : ورودی ۱:
- ۲: بردار توصیف‌گر نقاط کلیدی  $Q = [q_1, q_2, \dots, q_n]$ : ورودی ۲: تصویر هدف
- ۳: بردار فاصله نقاط کلیدی  $P$ :  $D = [d_1, d_2, \dots, d_n]$ : ورودی ۳: تصویر هدف نسبت به

یافت که با استفاده از شبکه دوقلو بسیاری از این تطابق‌ها حذف خواهند شد.

### ۲-۳- تطبیق نقاط کلیدی و به دست آوردن نقاط متناظر

در الگوریتم SIFT برای مقایسه توصیف‌گرهای دو تصویر از فاصله اقلیدسی استفاده می‌شود. اگر  $P = [p_1, p_2, \dots, p_n]$  و  $Q = [q_1, q_2, \dots, q_n]$  به ترتیب دو توصیف‌گر مربوط به نقاط کلیدی از تصاویر مرجع و هدف باشد آن‌گاه فاصله این دو توصیف‌گر را می‌توان به صورت رابطه (۱) تعریف کرد:

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (1)$$

این فاصله میان هر کدام از نقاط کلیدی مربوط به تصویر نخست و تمام نقاط کلیدی مربوط به تصویر دوم محاسبه می‌شود؛ در نهایت در الگوریتم SIFT پیشنهاد شده است که تنها زمانی می‌توان دو نقطه  $f_1$  و  $f_2$  را به عنوان نقاط متناظر در نظر گرفت که در رابطه (۲) صدق کنند:

$$\frac{d(f_1, f_2)}{d(f_1, f'_2)} < T \quad (2)$$

که  $T$  مقدار آستانه،  $f_2$  بهترین نقطه کلیدی داوطلب از تصویر هدف به عنوان نقطه متناظر برای  $f_1$  از تصویر مرجع و  $f'_2$  دومین داوطلب مناسب از تصویر هدف به عنوان نقطه متناظر برای  $f_1$  است. این روش باعث حذف تناظرهای مبهم و افزایش دقت الگوریتم SIFT می‌شود.

استفاده از روش پیشنهادی الگوریتم SIFT برای مقایسه توصیف‌گرها و یافتن نقاط متناظر در تصاویر سنجش از دور، به ویژه در شرایطی که تصاویر دارای پس‌زمینه پیچیده، تغییرات شدید و غیرخطی شدت روشنایی، یا تفاوت در محتوا و پوشش سطح زمین باشند، ممکن است با چالش مواجه شود. در چنین شرایطی، این روش برحسب معمول قادر به شناسایی تعداد کافی از نقاط متناظر نیست و بسیاری از تطابق‌های صحیح ممکن است نادیده گرفته یا حذف شوند. شکل (۳) دو مورد از تکه‌های تصویر اطراف نقاط کلیدی از تصویر مرجع به همراه سه تکه تصویر اطراف نزدیک‌ترین نقاط کلیدی از تصویر هدف به این نقاط را که از بالا به پایین بر حسب میزان فاصله اقلیدسی نسبت به نقطه کلیدی تصویر مرجع است را نشان می‌دهد. استفاده از روش گفته شده برای تطابق نقاط در الگوریتم SIFT باعث حذف و یا شناسایی نشدن نقاط متناظر در این شرایط می‌شود.

در شکل (۳-الف) با وجود این که دو نقطه کلیدی دارای کمترین فاصله از تصاویر مرجع و هدف به درستی پیدا شده‌اند و متناظرند، اما به دلیل نزدیک بودن فاصله دومین نقطه داوطلب، الگوریتم SIFT این تناظر را رد کرده است. در شکل (۳-ب) نقطه کلیدی در تصویر مرجع با دومین نقطه داوطلب

### ۳-۴- ساختار پیشنهادی برای شبکه دوقلو

ساختار کلی شبکه دوقلو استفاده شده در این پژوهش در شکل (۱) نشان داده شد. این شبکه از دو بخش استخراج کننده ویژگی و اندازه گیری میزان شباهت تشکیل شده است. قسمت استخراج کننده ویژگی بر اساس مدل ارائه شده در پژوهش [۱۵] است.

همان گونه که نشان داده شده است ورودی های شبکه همان تکه های تصویر است که اطراف نقاط متناظر استخراج کردیم. در هر کانال شش لایه کانولوشنی وجود دارد که وظیفه استخراج ویژگی از تکه های تصویر را برعهده دارند. یک لایه Max Pooling پس از نخستین لایه کانولوشنی به منظور کاهش پیچیدگی و بار محاسباتی استفاده شده است، علاوه بر این پس از هر لایه کانولوشنی از لایه batch normalization برای تسریع در فرایند آموزشی استفاده شده است [۳۰]؛ سپس با استفاده از یک لایه concatenate ویژگی های استخراج شده در خروجی کانال ها کنار هم قرار داده می شود. در انتها برای تعیین میزان شباهت ورودی ها از دو لایه تمام متصل استفاده می شود. خروجی شبکه عددی بین صفر تا یک است که هرچه این عدد بزرگ تر باشد احتمال متناظر بودن ورودی ها بیشتر است.

### ۳-۵- آموزش شبکه دوقلو

برای آموزش شبکه دوقلو یک مجموعه از تصاویر گوگل ارث جمع آوری شده است. این مجموعه آموزشی شامل پنجاه طبقه مختلف است که هر طبقه شامل ده تصویر از یک ناحیه یکسان و مربوط به زمان های مختلف است. از این پنجاه طبقه تعداد ۴۴ طبقه برای آموزش شبکه دوقلو و تعداد شش طبقه برای ارزیابی شبکه دوقلو استفاده شده است. مجموعه داده آموزشی شامل مناطق شهری و روستایی، نواحی جنگلی، زمین های کشاورزی، جاده ها و غیره است. تصاویر مربوط به هر کلاس آموزشی همان گونه که در شکل (۶) نشان داده شد دارای تفاوت های پیچیده خطی و غیرخطی نسبت به هم هستند که باعث می شود شبکه آموزش داده شده نسبت به این عوامل پایدار باشد.

از هر کدام از تصاویر مربوط به طبقه های مختلف تکه های تصویر با ابعاد متفاوت استخراج می شود که به این صورت هر کلاس خود به چندین طبقه تقسیم می شود و سپس این تکه های تصویر متناسب با ورودی شبکه دوقلو تغییر ابعاد داده می شوند. در شکل (۷) چند نمونه از مجموعه تکه های تصویر متناظر استخراج شده از این تصاویر نشان داده شده است. برای ایجاد نمونه تکه های تصویر آموزشی غیرمتناظر، از طریق انتخاب تصادفی تکه های تصویر متعلق به طبقه های مختلف این بخش از نمونه های آموزشی نیز ایجاد می شود. انتخاب تصادفی جفت تکه تصاویر غیرمتناظر ممکن است در برخی موارد موجب شود که تشخیص غیرمتناظر بودن برای شبکه

مقدار آستانه  $T$ : ورودی ۴:

تعداد بیشینه نقاط متناظر قابل  $P$ :  $max-MP$ : ورودی ۵:

قبول برای نقطه

نقاط متناظر با  $P$ :  $MP$ : خروجی ۶:

۷: for  $i=1$  to  $max-MP$  do

۸: if  $(d_i / d_{i+1}) < T$  Then

۹:  $MP = q_l$  to  $q_i$ ;

۱۰: return  $MP$ ;

۱۱: break;

۱۲: end if

۱۳: end for

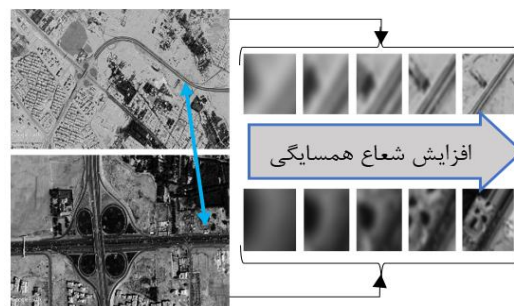
(شکل-۴): الگوریتم پیشنهادی برای مطابقت دادن نقاط کلیدی

تصاویر مرجع و هدف

(Figure-4): Proposed algorithm for matching key points of reference and target images

### ۳-۳- استخراج تکه تصویر به مرکزیت نقاط کلیدی متناظر

هر نقطه در الگوریتم SIFT را با استفاده از اطلاعات مکانی  $(x, y)$ ، مقیاس  $(\sigma)$  و زاویه  $(\theta)$  می توان مشخص کرد. با استفاده از این اطلاعات اطراف هر کدام از نقاط متناظر شناسایی شده در مراحل پیشین، یک تکه تصویر استخراج می شود؛ سپس ابعاد آن متناسب با ابعاد ورودی شبکه دوقلو تغییر داده خواهد شد. در این مرحله متناسب با مقیاس و زاویه مربوط به هر کدام از جفت نقاط متناظر و بر اساس اطلاعات مکانی هر کدام از این دو نقطه در تصاویر مربوطه، شعاع همسایگی متفاوتی برای آن ها می توان در نظر گرفت. در شکل (۵) یکی از جفت نقاط متناظر شناسایی شده به وسیله الگوریتم SIFT برای دو تصویر هدف و مرجع نشان داده شده است. برای این دو نقطه تکه های تصویر با شعاع همسایگی متفاوت و متناسب با زاویه و مقیاس هر نقطه استخراج شده و متناسب با ورودی شبکه دوقلو ابعاد هر کدام تغییر داده شده است؛ همان گونه که در تصویر مشخص است هرچه شعاع همسایگی بزرگ تر باشد با دقت بالاتری می توان در مورد تناظر و یا عدم تناظر قضاوت کرد.



(شکل-۵): استخراج تکه های تصویر با شعاع همسایگی

متفاوت اطراف یکی از جفت نقاط متناظر شناسایی شده توسط

الگوریتم SIFT

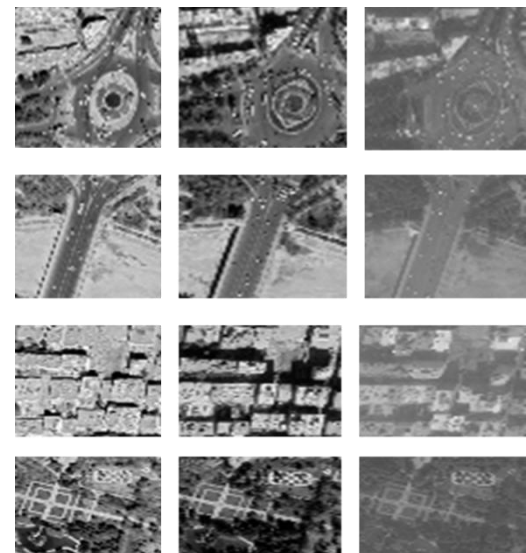
(Figure-5): Extracting image patches with different neighborhood radius around one of the pairs of corresponding points detected by SIFT algorithm



دوقلو به سادگی انجام شود؛ از سوی دیگر با توجه به ماهیت تصاویر سنجش از دور و وجود نواحی با شباهت فراوان در تصاویر و همچنین بخش‌های تکرارشونده در تصاویر، موجب می‌شود در بسیاری از جفت تکه‌تصاویر غیرمتناظر که به‌صورت تصادفی نیز انتخاب شده‌اند، شباهت زیادی وجود داشته باشد و رسیدن به دقت مناسب مستلزم آموزش دقیق شبکه دوقلو است.



(شکل-۶): چند نمونه از تصاویر مربوط به یک ناحیه شهری  
(Figure-6): Some examples of images related to an urban area



(شکل-۷): چند نمونه از مجموعه تکه‌های تصویر متناظر استخراج شده از تصاویر شکل (۶)  
(Figure-7): Some examples of image patches extracted from the images of Figure 6

که با استفاده از سنجنده‌های مختلف، مربوط به زمان‌های متفاوت، شرایط جوی مختلف، انواع شرایط محیطی، تفاوت در وضوح و کیفیت و تحت تأثیر عواملی مانند تغییرات مقیاس، چرخش و زاویه دید متفاوت قرار دارند به بررسی پارامترهای الگوریتم SIFT و تنظیم این پارامترها پرداخته خواهد شد. در ادامه با استفاده از شبکه دوقلو آموزش داده شده و به کمک الگوریتم SIFT با پارامترهای تنظیم شده، نتایج حاصل از روش پیشنهادی را مورد ارزیابی قرار داده و این نتایج با چند مورد از روش‌های تطبیق تصاویر سنجش از دور مقایسه می‌شود.

#### ۴-۱- آموزش و ارزیابی شبکه دوقلو

ساختار کلی استفاده شده برای شبکه دوقلو در شکل (۱) نشان داده شد. دو کانال موازی مربوط به استخراج ویژگی‌ها درواقع یک شاخه یکسان است که ویژگی‌های مربوط به هر کدام از ورودی‌ها را جداگانه استخراج می‌کند. این شاخه به ترتیب به صورت  $P(2,2) - C(64,7,1) - C(256,5,1) - C(128,5,1) - C(128,5,1) - C(256,5,1)$  است که  $C(i,j,k)$  نشان‌دهنده لایه کانولوشنی به تعداد  $i$  پالایه با ابعاد  $j \times j$  و تعداد کانال  $k$  و  $P(j,s)$  نشان‌دهنده یک لایه Max-Pooling با ابعاد  $j \times j$  و گام  $s$  است. در قسمت مربوط به اندازه‌گیری میزان شباهت، ابتدا با استفاده از یک لایه Concatenate ویژگی‌های استخراج شده کنار هم قرار می‌گیرد. با استفاده از لایه‌های تمام متصل  $F(512)$  و  $F(1)$  که  $F(n)$  نشان‌دهنده یک لایه تمام متصل با  $n$  خروجی است، خروجی نهایی شبکه ایجاد می‌شود. به جز آخرین لایه تمام متصل، پس از هر لایه کانولوشنی و تمام متصل از تابع فعال‌سازی ReLU و لایه batch-normalization استفاده شده‌است. در آخرین لایه تمام متصل از تابع فعال‌سازی sigmoid استفاده شده‌است که در خروجی احتمال متناظر بودن ورودی‌ها را به صورت عددی بین صفر تا یک تولید می‌کند.

همان‌گونه که در بخش پیشین گفته شد، مجموعه تصاویر آموزشی با استفاده از تصاویر گوگل ارث جمع‌آوری شده‌است و از این تصاویر تکه‌های تصویر متناظر استخراج می‌شود و در نهایت به ابعاد  $105 \times 105$  به عنوان ورودی شبکه تغییر اندازه داده می‌شوند. با استفاده از تصاویر جمع‌آوری شده تعداد ۴۳۰ هزار جفت تکه تصویر متناظر به منظور جلوگیری از بیش‌برازش در فرایند آموزش شبکه مورد استفاده قرار گرفت که از این تعداد ۳۸۴ هزار برای آموزش و ۴۶ هزار برای ارزیابی شبکه در حین فرایند آموزش استفاده شده‌است؛ همان‌طور که در بخش سه گفته شد برای ایجاد نمونه‌های آموزشی غیرمتناظر از طریق انتخاب تصادفی میان کلاس‌های متفاوت از نمونه‌های متناظر عمل می‌شود. فرایند

#### ۴- پیاده‌سازی و نتایج

در این فصل ابتدا شبکه دوقلو آموزش داده شده‌است و دقت این شبکه برای تصاویر سنجش از دور مورد بررسی قرار می‌گیرد؛ سپس یک مجموعه از تصاویر سنجش از دور



کانولوشنی، هدف حذف ورودی های غیرمتناظر با دقت بالا و سپس حفظ ورودی های متناظر با دقتی مناسب خواهد بود. با توجه به نتایج به دست آمده مقدار آستانه برابر با ۹۱ درصد برای شبکه کانولوشنی در نظر گرفته شد که در این حالت ورودی های غیرمتناظر با دقت ۹۸/۸ درصد حذف و ورودی های متناظر با دقت ۹۰/۲ درصد تشخیص داده شدند.

#### ۴-۲- تنظیم پارامترهای الگوریتم SIFT

در این بخش به تنظیم برخی از پارامترهای الگوریتم SIFT که تنظیم آن ها تأثیر زیادی می تواند در تعداد تطابق های صحیح داشته باشد، پرداخته خواهد شد. برای تنظیم این پارامترها تعداد تطابق های اشتباه ایجاد شده نیز مدنظر قرار خواهد گرفت. برای انجام آزمایش هایی در این بخش یک مجموعه شامل ۳۵ جفت تصویر از تصاویر سنجنش ازدور جمع آوری شده است. تصاویر این مجموعه با استفاده از گوگل ارث و همچنین مراجع [۳۱]، [۳۲]، [۳۳] و [۳۴] جمع آوری شد. این مجموعه از تصاویر سنجنش ازدور با استفاده از سنجنده های مختلف شامل Triple Sat، SkySat-2، Sentinel-2، IKONOS، CartoSat-1 و Sat است و وضوح مکانی از ۰/۸ متر (تصاویر مربوط به سنجنده Triple Sat) تا ده متر (تصاویر مربوط به سنجنده Sentinel-2) را شامل می شود؛ همچنین این مجموعه تصاویر مربوط به زمان های متفاوت، شرایط جوی مختلف، انواع شرایط محیطی، تفاوت در وضوح و کیفیت و تحت تأثیر عواملی مانند تغییرات مقیاس، چرخش و زاویه دید متفاوت قرار دارند.

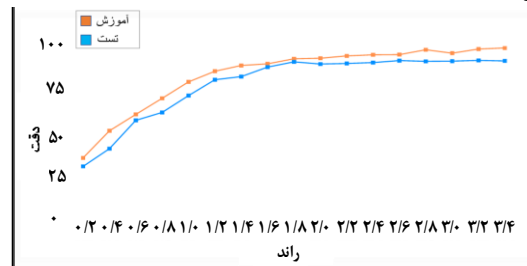
برای پیاده سازی الگوریتم SIFT و تغییر پارامترهای آن از نرم افزار Matlab 2015b و همچنین تابع vl\_sift مربوط به کتابخانه منبع باز بینایی کامپیوتر VLFeat [۳۵] استفاده شده است. پارامترهای مورد بررسی شامل پارامتر مربوط به حذف نقاط دارای شدت نور پایین (PeakThresh)، مقدار آستانه استفاده شده برای تطابق نقاط (T) و بیشینه نقاط متناظر قابل قبول برای یک نقطه (max\_MP) است. مقادیر پیشنهاد شده در الگوریتم SIFT برای پارامترهای PeakThresh، T و max\_MP به ترتیب برابر با ۰.۰۳، ۰.۸ و ۱ است.

از آنجایی که پارامتر PeakThresh تأثیر بسیار زیادی بر روی تعداد نقاط کلیدی در تصاویر دارد، ابتدا این پارامتر تنظیم خواهد شد. پس از تنظیم پارامتر PeakThresh، دو پارامتر دیگر یعنی T و max\_MP که در مرحله تطبیق نقاط کلیدی کاربرد دارند، به ترتیب تنظیم می شوند؛ درواقع می توان گفت ترتیب بهینه کردن این سه پارامتر بر اساس ترتیب اعمال هر کدام در الگوریتم SIFT است.

برای تنظیم هر پارامتر در مرحله نخست با در نظر گرفتن مقادیر پیشنهادی در الگوریتم SIFT، نقاط

پیاده سازی و آموزش و تست شبکه دوقلو با استفاده از زبان برنامه نویسی پایتون<sup>۱</sup> و کتابخانه کراس<sup>۲</sup> انجام شده است.

وزن های شبکه با استفاده از توزیع تصادفی گوسی [۳۰] مقداردهی اولیه شده اند. برای کامپایل شبکه از بهینه ساز Adam و تابع هزینه binary cross entropy استفاده شده است. نرخ آموزش اولیه برابر با ۰/۰۱ قرار داده شده است و بر اساس روش گفته شده در پژوهش [۳۲] کاهش داده می شود. شبکه به تعداد بیست هزار تکرار<sup>۳</sup> و اندازه دسته<sup>۴</sup> ۱۲۸ تایی که نیمی از آن ها متناظر و نیمی غیرمتناظرند، مورد آموزش قرار می گیرد؛ بنابراین به تعداد حدود ۳.۳ راند<sup>۵</sup> فرایند آموزش بر روی تمامی مجموعه داده های آموزشی انجام می شود. در شکل (۸) نمودار مربوط به میانگین دقت در حین فرایند آموزش شبکه دوقلو نشان داده شده است. میانگین دقت بر اساس تکرارهای موجود در هر ۰.۲ راند محاسبه شده و در صورتی که خروجی شبکه دوقلو بالای ۰.۵ باشد به معنای تشخیص جفت تصاویر ورودی به عنوان متناظر و در غیر این صورت تشخیص شبکه غیرمتناظر بودن تصاویر ورودی است.



(شکل-۸): میانگین دقت در حین فرایند آموزش شبکه دوقلو (Figure-8): Average accuracy during training process of Siamese network

خروجی شبکه دوقلو همان طور که گفته شد، عددی بین صفر تا یک است که هرچه این عدد به یک نزدیک تر باشد احتمال متناظر بودن ورودی ها بیشتر است. برای این خروجی لازم است مقداری به عنوان مقدار آستانه تعیین شود و در صورتی که خروجی از این مقدار بیشتر باشد، ورودی ها را متناظر و در صورتی که خروجی کمتر از این مقدار آستانه باشد، ورودی ها غیرمتناظر در نظر گرفته شوند. برای تعیین این مقدار آستانه به این صورت عمل شد که تعداد سه هزار جفت تکه تصویر که نیمی از آن ها متناظر و نیمی غیرمتناظرند از تصاویر سنجنش ازدور که در فرایند آموزش شبکه استفاده نشده اند، جمع آوری و شبکه دوقلو از این طریق مورد ارزیابی قرار گرفت. با توجه به اینکه در خروجی الگوریتم SIFT تعداد زیادی تطابق اشتباه ایجاد می شود، برای تعیین مقدار آستانه برای شبکه

<sup>۱</sup> Python

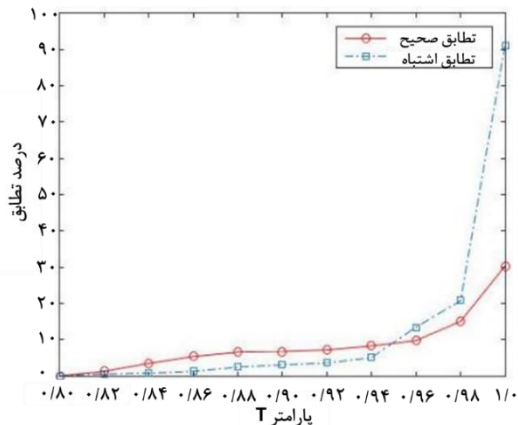
<sup>۲</sup> Keras

<sup>۳</sup> Iteration

<sup>۴</sup> Batch size

<sup>۵</sup> Round

می‌شود. درصد افزایش و یا کاهش تطابق‌های صحیح و اشتباه ایجادشده را نسبت به نتایج به‌دست‌آمده برای مقدار نهایی در نظر گرفته شده برای PeakTresh که ۰/۰۰۵ است مورد بررسی قرار می‌دهیم. شکل (۱۰) نتایج به‌دست‌آمده بر روی جفت تصاویر سنجش‌ازدور را نشان می‌دهد.



(شکل-۱۰): درصد تغییر تطابق‌های صحیح و اشتباه

با تغییر پارامتر T

(Figure-10): change of correct and incorrect matches by changing the T parameter

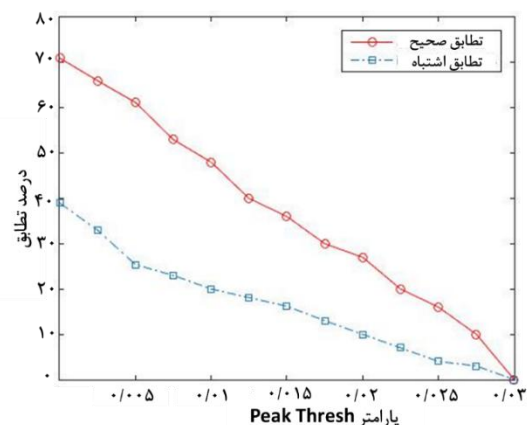
همان‌طور که در شکل (۱۰) مشاهده می‌شود با افزایش پارامتر T درصد تطابق‌های صحیح افزایش یافته است؛ همچنین هرچه به عدد یک نزدیک‌تر شده‌ایم رشد تطابق‌های اشتباه زیادتر شده‌است. دلیل افزایش شدید تطابق‌های اشتباه در حالتی که T برابر با یک در نظر گرفته شده‌است این است که در این حالت برای هر نقطه کلیدی یک نقطه متناظر وجود دارد و به‌دلیل ماهیت تصاویر سنجش‌ازدور و وجود ساختار تکرارشونده و نواحی بدون بافت یا دارای بافت ضعیف، تعداد تطابق‌های اشتباه افزایش زیادی خواهد داشت؛ در صورتی که پارامتر T را برابر با ۰/۸۸ در نظر بگیریم تعداد تطابق‌های صحیح نسبت به حالتی که این مقدار برابر با ۰/۸ است ۶/۵ درصد افزایش خواهد داشت و تطابق‌های اشتباه ۲/۳۵ درصد افزایش خواهد داشت؛ بنابراین در الگوریتم پیشنهادی از مقدار ۰/۸۸ برای پارامتر T استفاده خواهد شد.

برای بررسی و تنظیم پارامتر max\_MP مقدار این پارامتر از یک تا مقدار شش تغییر داده می‌شود و درصد افزایش و یا کاهش تطابق‌های صحیح و اشتباه با زمانی که این پارامتر یک است مقایسه می‌شود. در این مرحله از مقادیر نهایی در نظر گرفته‌شده برای پارامترهای PeakThresh و T استفاده خواهد شد. شکل (۱۱) نتایج به‌دست‌آمده بر روی جفت تصاویر سنجش‌ازدور را نشان می‌دهد.

با توجه به نمودار با افزایش پارامتر max\_MP تطابق‌های صحیح و اشتباه افزایش یافته است. بهترین شرایط در حالتی است که این پارامتر برابر با دو باشد که

متناظر شناسایی شده به‌وسیله این الگوریتم برای تمامی جفت تصاویر تعیین خواهد شد؛ سپس با استفاده از شبکه کانولوشنی دوقلو آموزش داده شده و صحت تطابق‌ها مورد بررسی قرار خواهد گرفت. در پایان مرحله نخست تعداد تطابق‌های صحیح و اشتباه نهایی تعیین می‌شود. در مرحله بعد با تغییر پارامتر موردنظر و ثابت نگه‌داشتن دیگر پارامترها، دوباره با استفاده از الگوریتم پیشنهادی تعداد تطابق‌های صحیح و اشتباه برای هر جفت از تصاویر به‌دست خواهد آمد و درصد افزایش و یا کاهش تطابق‌های صحیح و اشتباه نسبت به حالتی که از مقادیر پیشنهادی الگوریتم SIFT استفاده می‌شد، تعیین می‌شود. در انتها با در نظر گرفتن درصد افزایش و یا کاهش ایجادشده در تعداد تطابق‌های صحیح و اشتباه به‌دست‌آمده، مقدار مناسب برای هر پارامتر تعیین می‌شود.

ابتدا به بررسی و تنظیم پارامتر مربوط به حذف نقاط دارای شدت نور پایین (PeakThresh) می‌پردازیم. مقدار این پارامتر از صفر تا ۰/۳ تغییر داده می‌شود و درصد تغییر ایجادشده در تعداد تطابق‌های صحیح و اشتباه بر اساس مراحل گفته‌شده تعیین می‌شود. شکل (۹) نتایج به‌دست‌آمده برای پارامتر PeakThresh را بر روی تصاویر سنجش‌ازدور نشان می‌دهد.



(شکل-۹): درصد تغییر تطابق‌های صحیح و اشتباه با تغییر

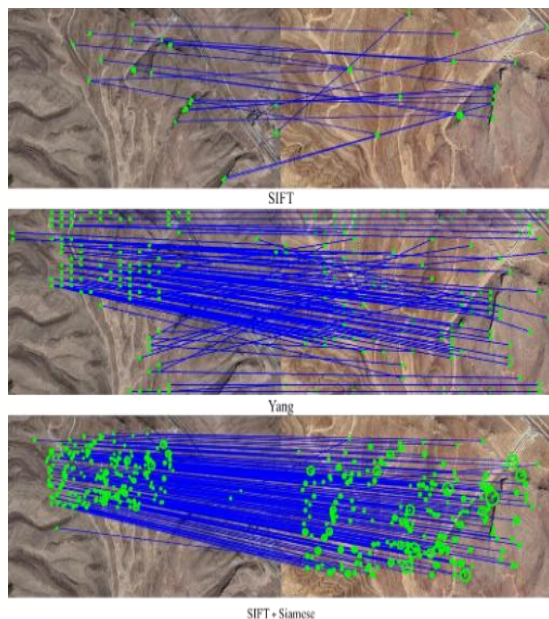
پارامتر PeakThresh

(Figure-9): change of correct and incorrect matches by changing the PeakThresh parameter

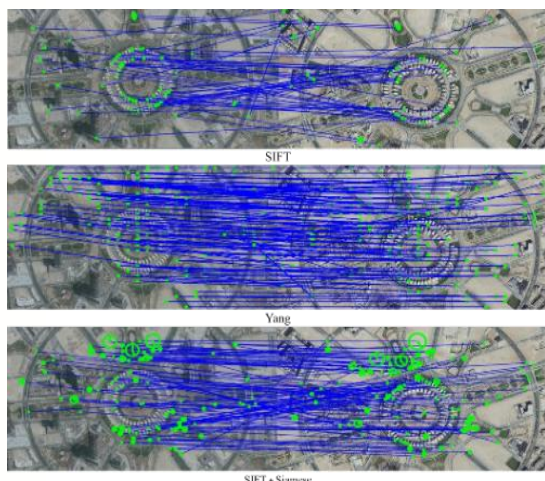
همان‌طور که در شکل (۹) مشاهده می‌شود با کاهش پارامتر PeakThresh درصد تطابق‌های صحیح افزایش قابل توجهی یافته است؛ در صورتی که این پارامتر را برابر با ۰/۰۰۵ در نظر بگیریم تعداد تطابق‌های صحیح نسبت به حالتی که این مقدار برابر با ۰/۳ است افزایش ۶۱/۲ درصد خواهد داشت و تطابق‌های اشتباه ۲۵/۴ درصد افزایش خواهد داشت؛ بنابراین در الگوریتم پیشنهادی از مقدار ۰/۰۰۵ برای حذف نقاط دارای شدت نور پایین استفاده خواهد شد.

برای بررسی و تنظیم پارامتر T مقدار این پارامتر از ۰/۸ که پیشنهاد الگوریتم SIFT است تا مقدار یک تغییر داده

از الگوریتم RANSAC نیز برای بررسی مجدد و کاهش احتمال خطا در شمارش تعداد تطابق‌های صحیح و اشتباه استفاده شده‌است.



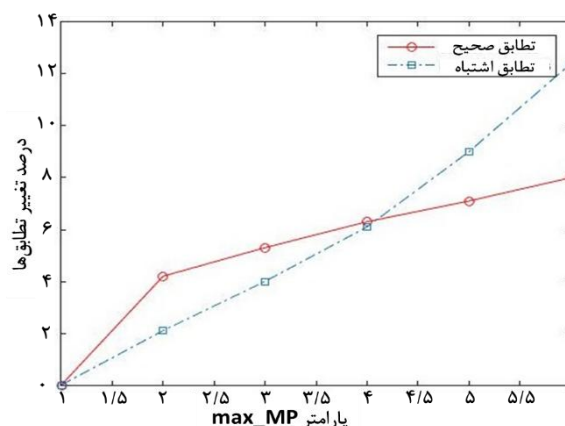
(شکل-۱۲): نتایج حاصل از اعمال روش پیشنهادی و دو روش دیگر به تصاویر سنجش‌ازدور از یک منطقه بیابانی  
(Figure-12): The results of applying the proposed method and two other methods to remote sensing images of a desert area



(شکل-۱۳): نتایج حاصل از اعمال روش پیشنهادی و دو روش دیگر به تصاویر سنجش‌ازدور از یک منطقه شهری  
(Figure-13): The results of applying the proposed method and two other methods to remote sensing images of an urban area

جدول (۱) نشان‌دهنده عملکرد مناسب الگوریتم پیشنهادی است. تعداد تطابق‌های صحیح و دقت تطابق‌ها در تمامی موارد نسبت به الگوریتم SIFT و در بسیاری از موارد نسبت به الگوریتم Yang بهبود قابل توجهی داشته مشخص است الگوریتم پیشنهادی موفق شده در بیشتر موارد، تعداد تطابق‌های صحیح و همچنین دقت تطابق‌ها را نسبت به الگوریتم‌های دیگر افزایش قابل توجهی دهد.

در این شرایط تطابق‌های صحیح ۴/۲ درصد و تطابق‌های اشتباه ۲/۱ درصد افزایش خواهد یافت؛ بنابراین در الگوریتم پیشنهادی از مقدار دو برای پارامتر  $max\_MP$  استفاده خواهد شد.



(شکل-۱۱): درصد تغییر تطابق‌های صحیح و اشتباه

با تغییر پارامتر  $max\_MP$

(Figure-11): change of correct and incorrect matches by changing the  $max\_MP$  parameter

### ۳-۴- بررسی عملکرد الگوریتم پیشنهادی و مقایسه

در این بخش به مقایسه عملکرد الگوریتم پیشنهادی با الگوریتم SIFT و الگوریتم ارائه‌شده در [۱۲] پرداخته خواهد شد. عملکرد روش پیشنهادی نسبت به دیگر روش‌ها با شمارش تعداد تطابق‌های صحیح و دقت تطابق‌ها ( $MP^1$ ) مورد ارزیابی قرار خواهد گرفت.  $MP$  به‌صورت رابطه (۳) محاسبه می‌شود.

$$MP = \frac{NCM}{NTM} \quad (3)$$

که  $NTM^2$  تعداد کل تطابق‌های یافت‌شده بین دو تصویر و  $NCM^3$  نشان‌دهنده تعداد تناظرهای صحیح است.

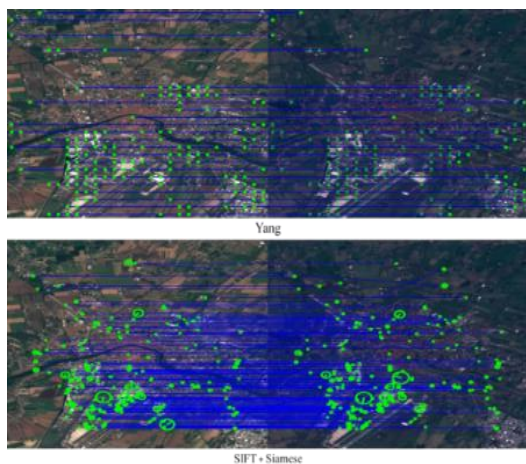
شکل‌های (۱۲)، (۱۳) و (۱۴) نتایج اعمال الگوریتم SIFT، الگوریتم [۱۳] (Yang) و روش پیشنهادی (SIFT+Siamese) بر روی سه جفت تصویر از مجموعه تصاویر سنجش‌ازدور جمع‌آوری‌شده قابل مشاهده است. نتایج نشان‌داده‌شده برای هرکدام از جفت تصاویر شامل تطابق‌های صحیح و اشتباه است. گفتنی است که جفت تصاویر ماهواره‌ای زمین مرجع نبوده و بررسی تناظر و یا عدم تناظر نقاط بر حسب پیکسل است. تعداد تطابق‌های صحیح و دقت تطابق‌ها برای هر یک از الگوریتم‌ها در جدول (۱) ذکر شده‌است. شمارش تعداد تطابق‌های صحیح و اشتباه به‌صورت بصری انجام شده‌است؛ همچنین

<sup>1</sup> Matching Precision

<sup>2</sup> Number of Total Matches

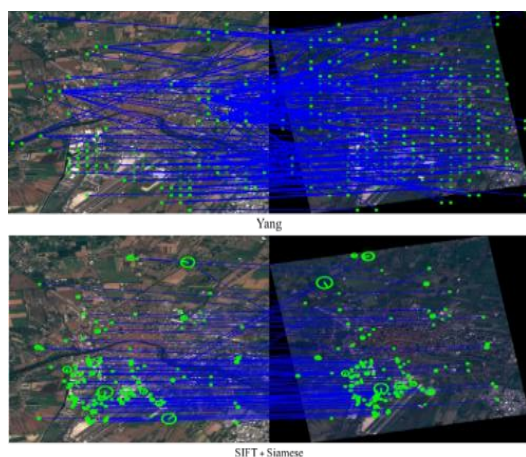
<sup>3</sup> Number of Correct Matches





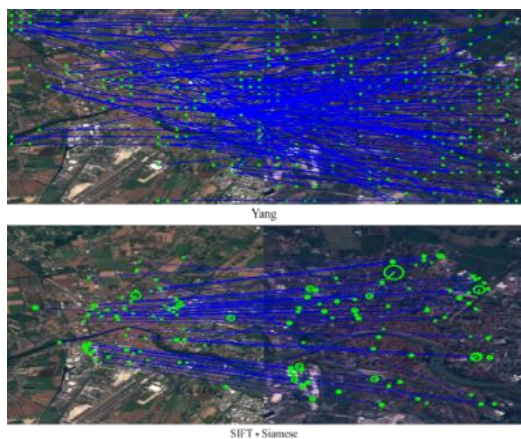
(شکل-۱۵): نتایج به دست آمده در شرایط عدم چرخش و تغییر مقیاس

(Figure-15): The results obtained in the case of no rotation and no scale changes



(شکل-۱۶): نتایج به دست آمده پس از ایجاد بیست درجه چرخش

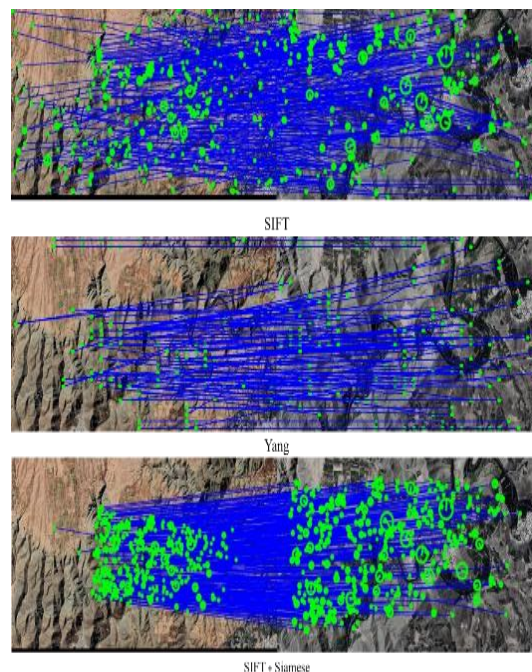
(Figure-16): The results obtained after applying 20 degrees of rotation



(شکل-۱۷): نتایج به دست آمده پس از ایجاد تغییر مقیاس به نسبت دو

(Figure-17): The results obtained when the scale is different with a ratio of 2

در اینجا همچنین روش پیشنهادی با روش LightGlue که یک روش مبتنی بر شبکه عصبی است



(شکل-۱۴): نتایج حاصل از اعمال روش پیشنهادی و دو روش

دیگر به تصاویر سنجش از دور از یک منطقه کوهستانی

(Figure-14): The results of applying the proposed method and two other methods to remote sensing images of a mountainous area

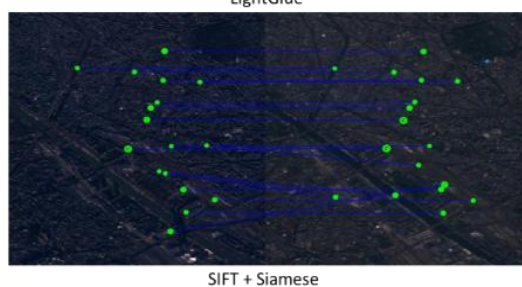
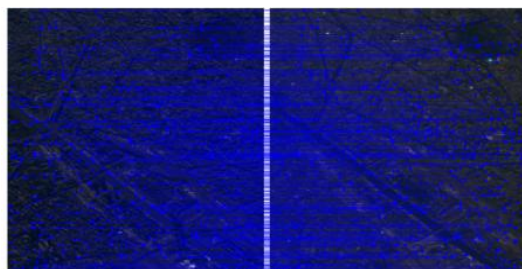
(جدول-۱): نتایج به دست آمده برای شکل‌های (۱۲)، (۱۳) و (۱۴)

(Table-1): The results obtained for figures (12), (13) and (14)

| شماره شکل | روش تطابق |      |      |      |                |      |
|-----------|-----------|------|------|------|----------------|------|
|           | SIFT      |      | Yang |      | SIFT + Siamese |      |
|           | NCM       | MP   | NCM  | MP   | NCM            | MP   |
| (۱۲)      | ۱۳        | ۰.۵۹ | ۹۵   | ۰.۷۱ | ۲۰۲            | ۰.۹۶ |
| (۱۳)      | ۱۰        | ۰.۱۷ | ۱۰۱  | ۰.۷۴ | ۷۶             | ۰.۶۷ |
| (۱۴)      | ۱۲۳       | ۰.۴۳ | ۷۰   | ۰.۵۲ | ۳۱۰            | ۰.۹۲ |

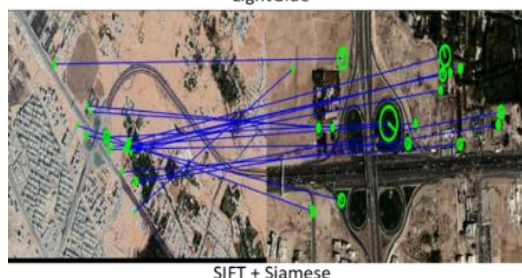
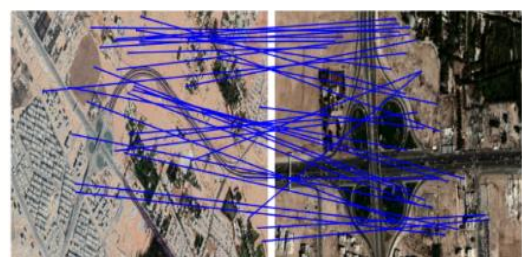
الگوریتم Yang در مواردی که تغییرات شدیدی در پس‌زمینه وجود دارد عملکرد خوبی دارد؛ اما در صورت وجود عواملی مانند تغییرات مقیاس، چرخش و زاویه دید متفاوت، دقت این الگوریتم کاهش خواهد یافت. برای بررسی این موضوع در ادامه با ایجاد چرخش و تغییر مقیاس در یکی از جفت تصاویر آزمایشی به مقایسه عملکرد الگوریتم پیشنهادی و الگوریتم Yang پرداخته می‌شود. در شکل (۱۵) حالتی که تغییرات مقیاس و چرخش در تصاویر وجود ندارد را نشان می‌دهد؛ سپس با ایجاد چرخش، تغییر مقیاس و وجود هم‌زمان چرخش و تغییر مقیاس عملکرد الگوریتم پیشنهادی و الگوریتم Yang مورد ارزیابی قرار می‌گیرد. نتایج به دست آمده در شکل‌های (۱۶)، (۱۷) و (۱۸) نشان داده شده است. تعداد تطابق‌های صحیح و دقت تطابق‌ها در جدول (۲) آورده شده است؛ همان‌طور که مشخص است، روش پیشنهادی در صورت وجود تغییرات مقیاس و چرخش میان تصاویر عملکرد بهتری نسبت به روش Yang دارد.





(شکل-۱۹): نتایج حاصل از اعمال روش پیشنهادی و روش LightGlue به تصاویر سنجش‌ازدور دارای چرخش و تفاوت مقیاس کم

(Figure-19): The results of applying the proposed method and the LightGlue method to remote sensing images with low rotation and low scale difference



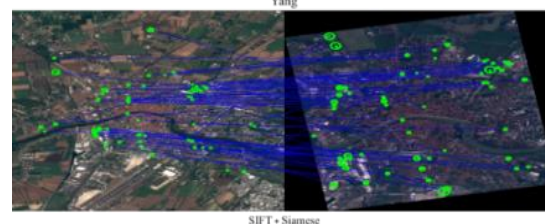
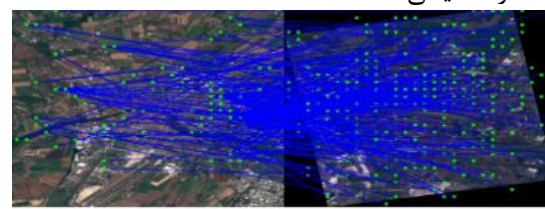
(شکل-۲۰): نتایج حاصل از اعمال روش پیشنهادی و روش LightGlue به تصاویر سنجش‌ازدور دارای چرخش و تفاوت مقیاس شدید

(Figure-20): The results of applying the proposed method and the LightGlue method to remote sensing images with high rotation and large scale difference

(جدول-۳): نتایج به‌دست آمده برای شکل‌های (۱۹) و (۲۰)  
(Table-3): The results obtained for figures (19) and (20)

| شماره<br>شکل | روش تطابق |       |                |       |
|--------------|-----------|-------|----------------|-------|
|              | LightGlue |       | SIFT + Siamese |       |
|              | NCM       | MP    | NCM            | MP    |
| (۱۹)         | ۸۵۲       | ۰.۹۹۶ | ۱۳             | ۰.۷۶۴ |
| (۲۰)         | ۰         | ۰     | ۱۵             | ۰.۸۸۲ |

مقایسه شده و نتایج آن برای دو جفت تصویر در شکل‌های (۱۹)، (۲۰) نشان داده شده‌است. روش LightGlue با وجود عملکرد بسیار خوب برای بسیاری از جفت تصاویر (مانند شکل (۱۹))، زمانی که مواردی مانند تغییر مقیاس شدید در تصاویر وجود دارد (مانند شکل (۲۰))، عملکرد ضعیف‌تری نسبت به روش پیشنهادی دارد و هیچ نقاط متناظر صحیحی نداشته است.



(شکل-۱۸): نتایج پس از ایجاد بیست درجه چرخش و تغییر مقیاس به نسبت دو

(Figure-18): Results after applying 20 degrees of rotation and scale difference with a ratio of 2

(جدول-۲): نتایج به‌دست آمده برای شکل‌های ۱۵ تا ۱۸

(Table-2): The results obtained for figures (15) to (18)

| شماره<br>شکل | روش تطابق |      |                |      |
|--------------|-----------|------|----------------|------|
|              | Yang      |      | SIFT + Siamese |      |
|              | NCM       | MP   | NCM            | MP   |
| (۱۵)         | ۱۴۰       | ۱    | ۱۷۱            | ۰.۹۵ |
| (۱۶)         | ۹۰        | ۰.۴۰ | ۱۱۳            | ۰.۹۰ |
| (۱۷)         | ۵۲        | ۰.۲۶ | ۶۰             | ۰.۸۱ |
| (۱۸)         | ۳۷        | ۰.۱۶ | ۴۵             | ۰.۷۶ |

جدول (۳) نتایج حاصل از مقایسه روش پیشنهادی و روش LightGlue برای جفت تصاویر نشان‌داده شده در شکل‌های (۱۹) و (۲۰) ارائه شده‌است. تفاوت اصلی روش پیشنهادی در این است که با تنظیم پارامترهای الگوریتم SIFT، نقاط کلیدی با کیفیت‌تری انتخاب می‌شوند و سپس تکه‌های تصویر حول این نقاط با لحاظ مقیاس و جهت استخراج می‌شوند. برخلاف LightGlue، در روش پیشنهادی ویژگی‌های محلی پایدار استخراج شده و به‌وسیله شبکه دوقلو ارزیابی می‌شوند، که باعث مقاومت بالاتر در برابر تغییر مقیاس و چرخش شده‌است؛ همچنین ارزیابی عملکرد هر کدام از روش‌ها بر روی تمامی ۳۵ جفت تصویر جمع‌آوری شده صورت گرفت که میانگین دقت تطابق برای الگوریتم پیشنهادی، LightGlue، Yang و الگوریتم SIFT به ترتیب برابر با ۰/۸۴۹، ۰/۸۲۳، ۰/۶۲۱ و ۰/۲۶۸ حاصل شد که عملکرد مناسب روش پیشنهادی را تأیید می‌کند.

در این پژوهش به تطبیق تصاویر مرئی ماهواره‌ای با استفاده از ویژگی‌های محلی و عمیق پرداخته شده است. تصاویر مورد استفاده علاوه بر تغییرات شدید و غیرخطی در پس‌زمینه، شامل چالش‌هایی مانند تغییرات شدید و غیرخطی در شدت نور، استفاده از سنجنده‌های مختلف برای دریافت تصویر و وجود عواملی مانند تفاوت مقیاس، چرخش و زاویه دید متفاوت است. از الگوریتم SIFT برای استخراج ویژگی‌های محلی استفاده شده است. این الگوریتم برای تطابق تصاویر سنجنش‌ازدور با ویژگی‌های گفته‌شده دارای دو مشکل اساسی است که شامل محدودیت در تعداد تطابق صحیح و وجود تعداد زیاد تطابق اشتباه است. با ایجاد تغییراتی در این الگوریتم، تعداد تطابق‌های صحیح را افزایش و برای حذف تطابق‌های اشتباه و افزایش دقت الگوریتم، یک شبکه کانولوشنی از نوع دوقلو آموزش داده شده است.

با جمع‌آوری یک مجموعه از تصاویر سنجنش‌ازدور عملکرد الگوریتم پیشنهادی بررسی شد که این الگوریتم توانست با دقت قابل قبولی نقاط متناظر میان تصاویر را شناسایی کند. نتایج به‌دست‌آمده از الگوریتم پیشنهادی نشان‌دهنده عملکرد مناسب روش پیشنهادی به لحاظ تعداد تطابق‌های صحیح و دقت تطابق‌ها است. برای ادامه این مسیر، پیشنهاد می‌شود در پژوهش‌های آینده پارامترهای الگوریتم SIFT به‌صورت خودکار و وقتی تعیین شوند تا نیاز به تنظیم دستی کاهش یابد؛ همچنین، با توجه به ماهیت دومرحله‌ای روش پیشنهادی و حجم بالای نقاط کلیدی استخراج‌شده، زمان پردازش بالا یکی از چالش‌هاست که باید در مطالعات بعدی با استفاده از روش‌های بهینه‌سازی یا شبکه‌های سریع‌تر کاهش یابد؛ علاوه بر این، در این پژوهش از شبکه عصبی عمیق دوقلو برای بررسی میزان شباهت بین تکه‌های تصویری استفاده شده است. در کارهای آتی می‌توان عملکرد سایر مدل‌های یادگیری عمیق را نیز برای این منظور مورد بررسی و ارزیابی قرار داد.

## 6-References

## ۶-مراجع

- [6] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SAR-SIFT: a SIFT-like algorithm for SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 453–466, 2014.
- [7] S. Wang, H. You, and K. Fu, "BFSIFT: A novel method to find feature matches for SAR image registration," *IEEE Geoscience and remote sensing letters*, vol. 9, no. 4, pp. 649–653, 2011.
- [8] R. Song and J. Szymanski, "Well-distributed SIFT features," *Electronics letters*, vol. 45, no. 6, pp. 308–310, 2009.
- [9] L. Juan and O. Gwun, "A comparison of sift, pca-sift and surf," *International Journal of Image Processing (IJIP)*, vol. 3, no. 4, pp. 143–152, 2009.
- [10] A. Sedaghat, M. Mokhtarzade, and H. Ebadi, "Uniform Robust Scale-Invariant Feature Matching for Optical Remote Sensing Images," *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, VOL. 49, NO. 11, no. January 2011.
- [۱۱] صداقت، امین، عبادی، حمید، مختارزاده، مهدی، "بهبود الگوریتم SIFT برای تطبیق تصاویر ماهواره‌ای"، *مجله سنجنش از دور و سامانه‌های اطلاعات جغرافیایی ایران*، دوره ۲، شماره ۴، زمستان ۱۳۸۹.
- [11] A. Sedaghat, H. Ebadi, and M. Mokhtarzade, "Improving the SIFT algorithm in order to match satellite images," *Iranian Journal of Remote Sensing & GIS (GISG)*, vol. 2, no. 4, 2011.
- [۱۲] حسین‌نژاد، زهرا، نصری، مهدی، "موزایک تصاویر طبیعی بر اساس حذف نقاط کلیدی زائد در الگوریتم SIFT و الگوریتم RANSAC تطبیقی"، *پردازش علائم و داده‌ها*، دوره ۱۸، شماره ۲، صفحات ۱۴۷–۱۶۲، ۱۴۰۰.
- [12] Z. Hossein-Nejad and M. Nasri, "Natural image mosaicing based on redundant keypoint elimination method in SIFT algorithm and adaptive RANSAC method," *Signal and Data Processing*, vol. 18, no. 2, pp. 147–162, 2021.
- [13] Z. Yang, T. Dan, and Y. Yang, "Multi-temporal Remote Sensing Image Registration Using Deep Convolutional Features," *IEEE Access*, vol. PP, no. c, p. 1, 2018.
- [14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [15] H. He, M. Chen, T. Chen, and D. Li, "Matching of Remote Sensing Images with Complex Background Variations via Siamese Convolutional Neural Network," pp. 1–23, 2018.
- [16] Y. Dong et al., "Local Deep Descriptor for Remote Sensing Image Feature Matching," *Remote Sensing*, pp. 1–21, 2019.
- [17] F. Ye, Y. Su, H. Xiao, X. Zhao, and W. Min, "Remote sensing image registration using convolutional neural network features," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 232–236, 2018.
- [18] W. Ma et al., "Remote sensing image registration with modified SIFT and enhanced feature matching," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 1, pp. 3–7, 2016.
- [19] W. Ma, J. Zhang, Y. Wu, L. Jiao, H. Zhu, and W. Zhao, "A Novel Two-Step Registration Method for Remote Sensing Images Based on Deep and Local Features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4834–4843, 2019.
- [1] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image Matching from Handcrafted to Deep Features: A Survey," *International Journal of Computer Vision*, 2020.
- [2] H. P. Moravec, "Techniques towards automatic visual obstacle avoidance," 1977.
- [3] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey vision conference*, 1988, pp. 147–152.
- [4] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [5] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European conference on computer vision*, 2006, pp. 404–417.

2018-2018 IEEE International Geoscience and Remote Sensing Symposium, 2018, pp. 2115–2118.

- [34] Satellite Imaging Corporation, "Remote sensing and satellite imaging services." [Online]. Available: <https://www.satimagingcorp.com/>. [Accessed: May 15, 2025].
- [35] A. Vedaldi and B. Fulkerson, "VLFeat - An open and portable library of computer vision algorithms," pp. 1–4, 2010.



**احمد رضا زارعی**، متولد ۱۳۷۳ و دانش‌آموخته رشته مهندسی برق در مقطع کارشناسی‌ارشد از دانشگاه اصفهان در سال ۱۳۹۹ است. زمینه‌های

پژوهشی مورد علاقه وی شامل موضوعاتی مانند بینایی کامپیوتر، هوش مصنوعی و پردازش سیگنال است. نشانی رایانامه ایشان عبارت است از:

amdrza.zarei@gamil.com



**بهمن معلم**، متولد ۱۳۴۹ در تهران، دانش‌آموخته کارشناسی و کارشناسی‌ارشد مهندسی الکترونیک، به ترتیب در سال ۱۳۷۰ و ۱۳۷۴ از دانشگاه صنعتی اصفهان

و دانشگاه صنعتی امیرکبیر، است. نامبرده در سال ۱۳۸۲ موفق به دریافت درجه دکترا در رشته مهندسی برق از دانشگاه صنعتی امیرکبیر شد. وی از سال ۱۳۸۲ به عنوان عضو هیئت علمی در دانشگاه اصفهان، مشغول به خدمت است و هم‌اکنون استاد گروه مهندسی برق دانشکده فنی و مهندسی دانشگاه اصفهان است. زمینه‌های علمی مورد علاقه ایشان شامل موضوعاتی مانند پردازش تصاویر، توسعه شبکه‌های عصبی و پردازش هوشمند سیگنال است. اجرای بیش از سی پروژه کاربردی، تألیف بیش از سیصد عنوان مقاله و در مجلات و کنفرانس‌های معتبر بین‌المللی و ملی و نیز تألیف چهار عنوان کتاب در فعالیت‌های پژوهشی ایشان دیده می‌شود. ایشان در سال‌های ۱۳۸۸، ۱۳۹۰ و ۱۳۹۸ به عنوان پژوهشگر برتر استان اصفهان در حوزه مهندسی انتخاب و در سال ۱۳۹۱ موفق به کسب نشان پژوهشگر برتر در دومین جشنواره بین‌المللی تحقیق و توسعه ایران شدند. ایشان در سال‌های ۱۴۰۱، ۱۴۰۲ و ۱۴۰۳ به عنوان داوطلب دانشگاه اصفهان در جشنواره استاد نمونه کشوری به وزارت علوم، تحقیقات و فن‌آوری، معرفی شدند.

نشانی رایانامه ایشان عبارت است از:

p\_moallem@eng.ui.ac.ir

- [20] L. H. Hughes, M. Schmitt, L. Mou, Y. Wang, and X. X. Zhu, "Identifying corresponding patches in SAR and optical images with a pseudo-siamese CNN," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 784–788, 2018.
- [21] B. Li, J. Zhang, B. Liu, Y. Xiang, and Y. Zhang, "An improved algorithm with SuperPoint+ SuperGlue network for UAV remote sensing image registration," in *Proc. IGARSS – IEEE Int. Geosci. Remote Sens. Symp.*, pp. 9975–9978, 2024.
- [22] D. Quan, Z. Wang, C. Lv, S. Wang, Y. Li, B. Ren, J. Chanussot, and L. Jiao, "LM-Net: A lightweight matching network for remote sensing image matching and registration," *IEEE Trans. Geosci. Remote Sens.*, 2024.
- [23] P. Lindenberger, P. Sarlin, and M. Pollefeys, "LightGlue: Local feature matching at light speed," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2023, pp. 17627–17638.
- [24] W. Zhang, T. Li, Y. Zhang, G. Pei, X. Jiang, and Y. Yao, "LTFormer: A light-weight transformer-based self-supervised matching network for heterogeneous remote sensing images," *Inf. Fusion*, vol. 109, pp. 102425, 2024.
- [25] S. Ji, C. Zeng, Y. Zhang, and Y. Duan, "An evaluation of conventional and deep learning-based image-matching methods on diverse datasets," *Photogramm. Rec.*, vol. 38, no. 182, pp. 137–159, 2023.
- [۲۶] جهانی، سید علی، محبی، کیوان، زمانی بروجنی، فرساد، "بهبود تشخیص صحنه در سنجش از راه دور با استفاده از یادگیری عمیق و انتخابگر ویژگی"، پردازش علائم و داده‌ها، دوره ۲۱، شماره ۳، صفحات ۸۴–۶۹، ۱۴۰۳.
- [26] S. A. Jahani, K. Mohebbi, and F. Z. Boroujeni, "Improving Scene Recognition in Remote Sensing Using Deep Learning and Feature Selector," *Signal and Data Processing*, vol. 21, no. 3, pp. 69–84, 2024.
- [27] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, "Signature verification using a siamese time delay neural network," in *Advances in neural information processing systems*, 1994, pp. 737–744.
- [28] D. Chicco, "Siamese Neural Networks: An Overview," *Artificial Neural Networks*, pp. 73–94, 2020.
- [29] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *arXiv preprint arXiv:1502.03167*, 2015.
- [30] M. Brown, G. Hua, and S. Winder, "Discriminative Learning of Local Image Descriptors," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 1, pp. 43–57, 2011.
- [31] C. Wu, L. Zhang, and L. Zhang, "A scene change detection framework for multi-temporal very high resolution remote sensing images," *Signal Processing*, vol. 124, pp. 184–197, 2016.
- [32] C. Wu, L. Zhang, and B. Du, "Kernel slow feature analysis for scene change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 4, pp. 2367–2384, 2017.
- [33] R. C. Daudt, B. Le Saux, A. Boulch, and Y. Gousseau, "Urban change detection for multispectral earth observation using convolutional neural networks," in *IGARSS*



