

## بررسی مقایسه‌ای معیارهای متداول ارزیابی دقت و معیارهای مغایرت طبقه‌بندی تصاویر سنجش

### از دور

مهران دادجو<sup>۱</sup>، سید باقر فاطمی نصرآبادی<sup>۲\*</sup>

۱. دانشجوی کارشناسی ارشد سنجش از دور، گروه مهندسی نقشه برداری، دانشکده مهندسی عمران و

حمل و نقل، دانشگاه اصفهان

۲. استادیار، گروه مهندسی نقشه برداری، دانشکده مهندسی عمران و حمل و نقل، دانشگاه اصفهان

\*نویسنده عهده دار مکاتبات: اصفهان، خیابان هزارجریب، دانشگاه اصفهان، دانشکده مهندسی عمران و حمل و نقل، گروه

مهندسی نقشه برداری. کد پستی: ۷۳۴۴۱-۸۱۷۴۶، تلفن: ۰۹۱۳۱۰۹۱۰۶۴ و ۰۳۱۳۷۹۳۵۳۳۲

پست الکترونیک: sb.fatemi@eng.ui.ac.ir

### چکیده

ارزیابی دقت نتایج حاصل از طبقه‌بندی تصاویر ماهواره‌ای بسیار حائز اهمیت است. تاکنون پارامترهای متعددی برای بیان دقت طبقه‌بندی تصاویر مورد استفاده قرار گرفته‌اند. دو نمونه از متداول‌ترین پارامترها، ضریب کاپا<sup>۱</sup> و دقت کلی<sup>۲</sup> می‌باشند. بعضی از محققین ایراداتی به این پارامترهای متداول وارد کرده و پارامترهای جدیدی ارائه نموده‌اند و معتقدند این پارامترها، معایب پارامترهای متداول را ندارند. در این تحقیق رابطه دو مورد از این پارامترهای جدید ارزیابی دقت طبقه‌بندی با پارامترهای متداول مورد بررسی قرار گرفته است. این دو پارامتر مغایرت کمی<sup>۳</sup> و مغایرت تخصیصی<sup>۴</sup> نام گذاری شده‌اند که میزان عدم شباهت را گزارش می‌نمایند. به منظور بررسی رفتار این دو پارامتر طبقه‌بندی نظارت شده روی ۵۷ تصویر شامل سه دسته از تصاویر با قدرت تفکیک متوسط، قدرت تفکیک بالا و قدرت تفکیک خیلی بالا انجام شد. سپس ضریب کاپا و دقت کلی به عنوان پارامترهای متداول، و مغایرت کمی و مغایرت تخصیصی به عنوان پارامترهای جدید برای هر تصویر طبقه‌بندی شده محاسبه شده و همبستگی مقادیر به دست آمده با یکدیگر بررسی گردید. نتایج به دست آمده در این تحقیق نشان می‌دهند که میزان همبستگی بین این دو دسته پارامتر بالا ولی در جهت منفی است؛ به عبارت دیگر با افزایش مقادیر یک دسته، مقادیر دسته دیگر کاهش می‌یابد. به این ترتیب پارامترهای مغایرت اطلاعات جدیدی را درباره نتایج طبقه‌بندی به کاربر ارائه نمی‌دهند و تنها در صورتی که خطای طبقه بندی خواسته شود، می‌توان از پارامترهای مغایرت نیز در کنار آن‌ها استفاده نمود. کلیدواژه‌ها: ارزیابی دقت، دقت کلی، طبقه‌بندی، معیارهای متداول، معیارهای جدید.

۱ Kappa Coefficient

۲ Overall Accuracy

۳ Quantity Disagreement

۴ Allocation Disagreement

## ۱. مقدمه

طبقه‌بندی در سنجش از دور، فرآیند شناسایی کلاس‌های طیفی موجود در داده تصویری می‌باشد که هر کدام از این کلاس‌های طیفی متناظر با یک نوع پوشش مشخص زمینی است. به عبارتی دیگر، هدف از طبقه‌بندی برچسب‌دهی پیکسل‌ها برای اختصاص آن‌ها به کلاس‌های طیفی مشخص می‌باشد (Richards, ۲۰۱۳).

طبقه‌بندی همواره به عنوان مهم‌ترین فرآیند تولید محصولات جغرافیایی پرکاربرد (نظیر نقشه‌های پوششی و کاربری، نقشه سطح پوشش‌های گیاهی، نقشه تغییرات پارامترهای محیطی مختلف) از تصاویر اخذ شده توسط سنجنده، مورد توجه محققین بوده است (Wilkinson, ۲۰۰۵).

Congalton و Green (۲۰۰۸) معتقدند فارغ از این که از کدام یک از روش‌های طبقه‌بندی استفاده می‌شود، ارزیابی دقت نقشه‌های حاصل از طبقه‌بندی به دلایل زیر الزامی است:

- برآورد میزان نزدیکی دقت نتایج حاصل به دقت مطلوب
- نیاز به افزایش کیفیت نتایج حاصل به کمک شناسایی و تصحیح منابع خطا
- نیاز به مقایسه تکنیک‌ها، الگوریتم‌ها، تحلیل‌ها و تفسیرهای مختلف جهت یافتن بهترین روش
- لزوم مشخص بودن دقت نتایج حاصل جهت استفاده از آن‌ها در فرآیندهای تصمیم‌گیری (مانند GIS)

بنابراین آخرین مرحله از روند طبقه‌بندی تصاویر ماهواره‌ای، ارزیابی دقت می‌باشد. با استفاده از روش‌های متداول ارزیابی دقت می‌توان با محاسبه پارامترهایی که از ماتریس ابهام<sup>۱</sup> استخراج می‌شوند، دقت نتایج را مشخص کرد (Hashemian, Abkar, & Fatemi, ۲۰۰۴). ماتریس ابهام یک ماتریس مربعی به ابعاد تعداد کلاس‌ها می‌باشد که حاصل مقایسه واحد به واحد داده‌های مرجع با نتایج حاصل از طبقه‌بندی است. در این ماتریس نتیجه مقایسه برچسب پیکسل‌های داده مرجع (در ستون‌ها) با برچسب پیکسل‌های داده طبقه‌بندی شده (در سطرها) ارائه می‌شود (Congalton & Green, ۲۰۰۸).

مرسوم‌ترین پارامترهایی که از ماتریس ابهام به دست می‌آیند عبارت‌اند از: دقت کلی و ضریب کاپا. دقت کلی میانگینی از صحت طبقه‌بندی است که نسبت پیکسل‌های صحیح طبقه‌بندی شده به جمع کل پیکسل‌های معلوم را نشان می‌دهد (Sui, Tian, & Xu, ۲۰۱۴). ضریب کاپا دقت طبقه‌بندی را نسبت به یک طبقه‌بندی کاملاً تصادفی محاسبه می‌کند (Foody, ۲۰۰۲). با وجود متداول بودن کاربرد پارامترهای ضریب کاپا و دقت کلی، برخی از محققین ایراداتی نسبت به این پارامترها وارد کرده و معیارهای جدیدی برای ارزیابی دقت طبقه‌بندی ارائه نموده‌اند.

---

<sup>۱</sup> Confusion Matrix

یکی از این ایرادات، ناتوان بودن پارامترهای متداول در ارزیابی دقت طبقه بندی در تصاویر با قدرت تفکیک مکانی خیلی بالا<sup>۱</sup> بیان شده است. در این تصاویر خصوصیات هندسی عوارض با سطح بالایی از جزئیات ثبت می‌شود و از آنجایی که پارامترهای کاپا و دقت کلی بر مبنای پیکسل می‌باشند، در برخی موارد نمی‌توانند دقت هندسی عوارض استخراج شده را به خوبی ارزیابی کنند (Zhou, Wu, & Li, ۲۰۱۱). بنابراین برای ارزیابی دقت طبقه بندی تصاویر با قدرت تفکیک مکانی بالا پیشنهاد شده است از معیارهایی که بر مبنای عارضه<sup>۲</sup> عمل می‌کنند، استفاده شود (Persello & Bruzzone, ۲۰۱۰).

یکی دیگر از ایرادات وارد شده، مربوط به ارزیابی دقت طبقه بندی کننده‌های نرم<sup>۳</sup> است. طبقه‌بندی کننده‌های نرم به هر پیکسل یک کلاس قطعی نسبت نمی‌دهند. لذا پارامترهای متداول دقت نمی‌توانند این طبقه بندی کننده‌ها را به درستی ارزیابی کنند (Zhang & Foody, ۱۹۹۸). بنابراین تعریف پارامترهای جدید و سازگار با خروجی طبقه بندی کننده‌های نرم، ضروری است (Marpu, Wijaya, & Gloaguen, ۲۰۰۸; Strahler et al., ۲۰۰۶). در این راستا راه‌حل‌های متعددی ارائه شده است. از جمله این راه‌حل‌ها می‌توان به توسعه پارامترهای جانبی مستخرج از ماتریس ابهام نظیر دقت کاربر و دقت تولید کننده (Emami, ۱۹۸۶; Story & Congalton, ۲۰۰۵)، محاسبه پارامتر کاپا از طرق مختلف و بر مبنای محاسبات آماری متعدد (Marpu et al., ۲۰۰۸)، توسعه روش‌های ارزیابی دقت شیء‌مبنا (Yi & Yan, Mas, Maathuis, Xiangmin, & Van Dijk, ۲۰۰۶) و استفاده از مفهوم فازی جهت تشکیل یک ماتریس ابهام متناسب با نتایج طبقه‌بندی کننده‌های نرم (El-Aziz, ۲۰۱۲)، بهره‌گیری از فواصل اقلیدسی<sup>۴</sup> و نرم L<sub>۱</sub> (Foody & Arora, ۱۹۹۶)، محاسبه ضرایب همبستگی (MASELLI, ۲۰۰۴) و توسعه پارامترهای جدید برای ارزیابی دقت و خطا (Rodolfi, & Conese, ۱۹۹۶; Gu, Sun, Pan, & Fan, ۲۰۱۲; Li & Guo, ۲۰۱۴) اشاره نمود.

یکی از جدی‌ترین انتقادات نسبت به پارامترهای متداول، به ویژه ضرایب کاپا، توسط Pontius Jr و Millones (۲۰۱۱) بیان شده است. این محققان با بیان ایرادات ضریب کاپا، این پارامتر را برای بیان ارزیابی دقت طبقه بندی نامناسب دانسته و به جای آن استفاده از پارامترهای جدیدی را توصیه کرده‌اند. معایب ضرایب کاپا از دیدگاه این دو نویسنده عبارت‌اند از: محاسبات پیچیده و درک سخت کاپا، مبهم شدن کاپا در صورتی که مخرج آن صفر یا عدد بسیار کوچکی شود و انجام محاسبات بر اساس یک طبقه‌بندی تصادفی. همچنین از آنجا که ضریب کاپا میزان شباهت را گزارش می‌کند، چندان برای ارزیابی دقت مناسب نیست (Pontius Jr & Millones, ۲۰۱۱). این محققین معتقدند برای تفسیر نتایج طبقه‌بندی استفاده از دو پارامتر مغایرت، کارآمدتر از یک پارامتر شباهت (کاپا) می‌باشد. زیرا پارامترهای مغایرت به روشنی میزان عدم مطابقت بین نتایج طبقه بندی و

<sup>۱</sup> Very High Resolution (VHR)

<sup>۲</sup> Object-based

<sup>۳</sup> Soft Classifiers

<sup>۴</sup> Euclidean Distance

داده‌های مرجع واقعیت زمینی را بر اساس عناصر ماتریس ابهام محاسبه می‌کنند. به همین دلیل با مقایسه نسبی پارامترهای مغایرت، می‌توان منابع خطا را شناسایی کرد؛ در حالی که کاپا صرفاً میزان مشابهت را با توجه به یک طبقه‌بندی تصادفی محاسبه می‌کند (Pontius Jr & Millones, ۲۰۱۱).

Pontius Jr و Millones برای پرهیز از معایب پارامترهای ارزیابی دقت متداول، دو پارامتر جدید به نام‌های مغایرت کمی و مغایرت تخصیصی معرفی کرده‌اند که این پارامترها بر خلاف ضریب کاپا و دقت کلی میزان عدم شباهت را گزارش می‌نمایند. پارامتر مغایرت کمی بیان‌گر میزان عدم شباهت هر پیکسل در داده مرجع با پیکسل متناظر آن در نقشه حاصل از طبقه‌بندی می‌باشد. پارامتر مغایرت تخصیصی بیان‌گر میزان عدم شباهت هر پیکسل از جهت انطباق حوزه مکانی کلاس‌ها می‌باشد (Pontius Jr & Millones, ۲۰۱۱). این محققین معتقدند این دو پارامتر معایب پارامترهای متداول را ندارند و توصیه اکید دارند که استفاده از ضریب کاپا متوقف شده و به جای آن‌ها از این پارامترهای مغایرت برای ارزیابی دقت نتایج حاصل از طبقه‌بندی استفاده شود. بعضی محققین نیز انتقاداتی نسبت به پارامترهای مغایرت کمی و اختصاصی داشته‌اند و سعی نموده‌اند به نحوی ایرادات این دو پارامتر را برطرف کنند. از جمله این محققین می‌توان Matthijs J. Warrens (۲۰۱۵) را نام برد. وی معتقد است پارامترهای مغایرت کمی و اختصاصی که در مقاله Pontius Jr & Millones (۲۰۱۱) فرمول‌بندی شده‌اند صرفاً مقادیر مطلق را می‌دهند، در حالی که در بسیاری از پژوهش‌ها مقادیر نسبی معنادارتر هستند. بنابراین آقای Warrens مقادیر نسبی پارامترهای مغایرت کمی و اختصاصی را معرفی می‌کند (Warrens, ۵۱۱۵).

پس از معرفی پارامترهای مغایرت توسط Pontius Jr و Millones، این دو پارامتر در تعدادی از پژوهش‌های بعدی نیز مورد استفاده قرار گرفته است. Brian Pickard و همکاران (۲۰۱۷) با مطالعه آماری روی چهار پارامتر ارزیابی دقت از جمله ضریب کاپا و پارامترهای مغایرت کمی و مغایرت اختصاصی، نتیجه‌گیری نموده‌اند که برای پوشش‌های مختلف زمین میزان کارایی هر پارامتر متفاوت است، بنابراین نباید یک قضاوت کلی و یکسان در مورد کارآمدی یک پارامتر ارزیابی دقت در همه مناطق انجام داد (Pickard, Gray, & Meentemeyer, ۲۰۱۷).

این تحقیق به بررسی دو پارامتر مغایرت کمی و مغایرت تخصیصی، رفتار آن‌ها و ارتباط آن‌ها با پارامترهای متداول ارزیابی دقت می‌پردازد. اهمیت بررسی این دو پارامتر جدید از این روست که اولاً پیشنهاد دهندگان آن‌ها، خود در کارهای قبلی بر استفاده از پارامتر کاپا تاکید داشته‌اند و حتی نوعی تعریف جدید و روش محاسبه جدید برای آن ارائه کرده‌اند (Pontius, ۲۰۰۰). بنابراین این چرخش از نظر قبلی می‌تواند قابل توجه و بررسی باشد. علاوه بر این تا به حال هیچ بررسی علمی برای ارزیابی این دو پارامتر و همچنین ارتباط آن‌ها با پارامترهای متداول دقت وجود ندارد. از دیگر وجوه این پارامترهای جدید تعریف معیار ارزیابی بر اساس مغایرت است که در آن از وجه مکانی پیکسل‌ها نیز استفاده شده است. لذا در این مطالعه بر آنیم تا این دو پارامتر معرفی شده

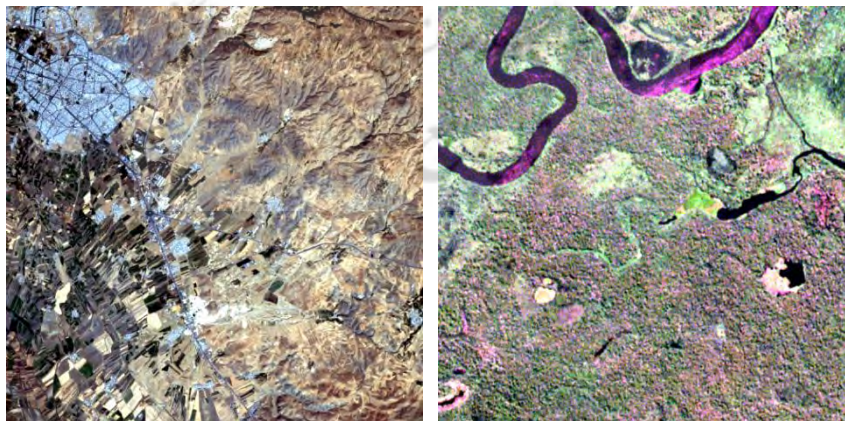
را از جنبه‌های مختلف بررسی نماییم. در نهایت نتایج تحقیق حاضر نشان می‌دهند که کارایی این دو پارامتر نسبت به پارامترهای قبلی تا چه میزان است و آیا می‌توان به عنوان پارامترهای جدید و جدی برای ارزیابی دقت به آن‌ها نظر داشت.

## ۲. موارد و روش‌ها

### ۲-۱- داده‌های مورد استفاده

در این تحقیق از تصاویر سه سنجنده با قدرت تفکیک مکانی متفاوت استفاده شده است که عبارت‌اند از: سنجنده OLI ماهواره Landsat ۸ (قدرت تفکیک متوسط) با ۱۲ باند طیفی، ماهواره Rapid eye (قدرت تفکیک بالا) با ۵ باند طیفی و ماهواره Worldview ۳ (قدرت تفکیک بسیار بالا<sup>۱</sup>) با ۸ باند طیفی. مجموعه تصاویر مورد استفاده شامل ۱۶ تصویر Landsat ۸، ۲۱ تصویر RapidEye و ۲۰ تصویر Worldview ۳ می‌باشد که ابعاد هر تصویر ۵۰۰\*۵۰۰ پیکسل در نظر گرفته شده است. در انتخاب تصاویر سعی بر آن بوده است که پوشش و کاربری مناطق مورد مطالعه متنوع باشد. عمده این پوشش‌ها شامل مناطق مسکونی، مناطق پوشش گیاهی، مناطق جنگلی، خاک، و آب می‌باشد. به ازای هر کدام از تصاویر یک نقشه مرجع واقعیت زمینی<sup>۲</sup> نیز به کمک تفسیر بصری تصاویر و تصاویر گوگل ارث<sup>۳</sup> جهت ارزیابی دقت تهیه گردید. سه نمونه از تصاویر به کار برده شده در شکل ۱ نشان داده شده است.

با استفاده از داده مرجع شبیه سازی شده صحت نتایج حاصل مورد بررسی قرار می‌گیرد. داده مرجع، یک ماتریس به ابعاد ۲۰۰\*۲۰۰ و حاوی ۷ کلاس می‌باشد. هر کلاس با یک عدد (از ۱ تا ۷) نشان داده می‌شود.



ب

الف

<sup>۱</sup> Very High Resolution

<sup>۲</sup> Ground Truth

<sup>۳</sup> Google Earth



ج

شکل ۱. سه نمونه از تصاویر مورد استفاده با ترکیب رنگی RGB الف) تصویر ماهواره RapidEye ب) تصویر سنجنده OLI ماهواره Landsat ۸ ج) تصویر ماهواره Worldview ۳

## ۲-۲- پارامترهای متداول ارزیابی دقت

پارامترهای متداول مورد استفاده در این تحقیق، دقت کلی و کاپا که در حال حاضر پرستفاده‌ترین معیارهای ارزیابی دقت طبقه‌بندی تصاویر ماهواره ای می‌باشند (Zhou et al., ۲۰۱۱). دقت کلی طبق رابطه (۱) از ماتریس ابهام محاسبه می‌شود.

$$O.A. = \frac{\sum_{i=1}^k n_{ii}}{n} \quad \text{رابطه (۱)}$$

که در آن  $k$  تعداد کلاس‌ها،  $n$  تعداد کل پیکسل‌های معلوم در داده مرجع،  $n_{ij}$  اعضای قطری ماتریس ابهام و  $O.A.$  دقت کلی طبقه بندی می‌باشد (Congalton & Green, ۲۰۰۸).

یکی از شناخته‌شده‌ترین برآوردهای کاپا با استفاده از عناصر ماتریس ابهام طبق رابطه (۲) محاسبه می‌شود (Congalton & Green, ۲۰۰۸).

$$K = \frac{n \sum_{i=1}^k n_{ii} - \sum_{i=1}^k n_{i+} n_{+i}}{n^2 - \sum_{i=1}^k n_{i+} n_{+i}} \quad \text{رابطه (۲)}$$

که در آن  $n$  تعداد کل پیکسل‌های معلوم،  $k$  تعداد کلاس‌ها،  $n_{i+}$  مجموع عناصر سطر  $i$  ام و  $n_{+i}$  مجموع عناصر ستون  $i$  ام می‌باشند.

## ۲-۳- پارامترهای مغایرت

همانطور که ذکر شد در این تحقیق رابطه دو معیار دقت کلی و ضریب کاپا با دو معیار مغایرت بررسی می‌گردند. Pontius Jr و Millones روابط محاسبه پارامترهای مغایرت کمی و تخصیصی را ارائه کرده‌اند (Pontius Jr & Millones, ۲۰۱۱). مغایرت کمی برای هر کلاس، بیانگر تفاضل تعداد نمونه‌هایی است که در نقشه حاصل از طبقه‌بندی و داده مرجع زمینی به یک کلاس

تعلق دارند (Pontius Jr & Millones, ۲۰۱۱). به عنوان مثال، چنانچه پس از طبقه‌بندی تعداد پیکسل‌هایی که برچسب پوشش گیاهی گرفته‌اند  $m$  و تعداد پیکسل‌های پوشش گیاهی در داده مرجع واقعیت زمینی  $n$  باشد، مغایرت کمی  $m-n$  خواهد شد. مغایرت کمی برای هر کلاس جداگانه محاسبه می‌گردد. مغایرت تخصیصی برای هر کلاس بیانگر مغایرت مکانی نمونه‌هایی است که در نتیجه طبقه‌بندی و داده مرجع واقعیت زمینی به یک کلاس تعلق دارند (Pontius Jr & Millones, ۲۰۱۱). به عنوان مثال، چنانچه پس از طبقه‌بندی  $m$  پیکسل برچسب پوشش گیاهی خورده باشند در صورتی که در داده مرجع واقعیت زمینی جزو پیکسل‌های کلاس دیگری باشند، مغایرت تخصیصی  $2m$  خواهد شد. بنابراین مغایرت تخصیصی همواره یک عدد زوج است زیرا برای یک زوج پیکسل با برچسب متفاوت محاسبه می‌شود (Pontius Jr & Millones, ۲۰۱۱). برای محاسبه این پارامترها، ابتدا طبق رابطه (۳)، یک ماتریس جدید به نام ماتریس جمعیت<sup>۱</sup> بر اساس ماتریس ابهام تشکیل می‌گردد. عناصر این ماتریس به صورت زیر محاسبه می‌گردند.

$$P_{ij} = \left[ \frac{n_{ij}}{\sum_{j=1}^J n_{ij}} \right] * \left[ \frac{N_i}{\sum_{i=1}^J N_i} \right] \quad \text{رابطه (۳)}$$

که در آن  $J$  تعداد کلاس‌ها،  $N_i$  تعداد پیکسل‌ها در کلاس  $i$  ام و  $n_{ij}$  بیانگر عدد متناظر سطر  $i$  ام و ستون  $j$  ام در ماتریس ابهام می‌باشد.

بر اساس ماتریس جمعیت، پارامترهای مغایرت کمی برای هر کلاس و پارامتر مغایرت کمی<sup>۲</sup> کلی به ترتیب با رابطه‌های (۴) و (۵) محاسبه می‌شوند.

$$q_g = \left| \left( \sum_{i=1}^J P_{ig} \right) - \left( \sum_{j=1}^J P_{gi} \right) \right| \quad \text{رابطه (۴)}$$

$$Q = \frac{\sum_{g=1}^G q_g}{2} \quad \text{رابطه (۵)}$$

پارامتر  $J$  برابر با تعداد کلاس‌ها و پارامتر  $g$  بیانگر یک کلاس به خصوص می‌باشد. پارامترهای مغایرت تخصیصی برای هر کلاس مطابق رابطه ۶ محاسبه می‌شوند. برای این که برآوردی کلی از میزان مغایرت تخصیصی طبقه‌بندی داشته باشیم پارامتر مغایرت تخصیصی<sup>۳</sup> کلی مطابق رابطه (۷) محاسبه می‌شود.

$$a_g = 2 \min \left[ \left( \sum_{i=1}^J P_{ig} \right) - P_{gg}, \left( \sum_{j=1}^J P_{gj} \right) - P_{gg} \right] \quad \text{رابطه (۶)}$$

$$A = \frac{\sum_{g=1}^G a_g}{2} \quad \text{رابطه (۷)}$$

طبق رابطه (۸)، مجموع مغایرت کمی و مغایرت تخصیصی، مغایرت کلی<sup>۴</sup> را به دست می‌دهد.

<sup>۱</sup> Population Matrix

<sup>۲</sup> Quantity Disagreement

<sup>۳</sup> Allocation Disagreement

<sup>۴</sup> Total Disagreement

$$T = Q + A$$

این پارامتر می‌تواند مانند پارامتر دقت کلی نوعی برآورد کلی از میزان مغایرت موجود در نتیجه طبقه‌بندی به دست بدهد.

#### ۴-۲- طبقه بندی و ارزیابی دقت

به منظور طبقه‌بندی تصاویر مورد آزمایش از روش نظارت شده بیشترین شباهت<sup>۱</sup> بهره گرفته شد (Mathers, ۱۹۹۹). با توجه به شرایط احتمالات اولیه برای همه کلاس‌ها مساوی در نظر گرفته شد. در طبقه بندی هر کدام از تصاویر، بسته به منطقه و عوارض موجود در هر تصویر، تعداد کلاس‌های انتخاب شده متفاوت است. در تصاویر Landsat ۸ از ۳ تا ۶ کلاس، در تصاویر RapidEye از ۳ تا ۵ کلاس و در تصاویر Worldview ۳ از ۵ تا ۷ کلاس در نظر گرفته شد. در فرایند طبقه‌بندی از همه باندهای موجود تصاویر بهره گرفته شد. با استفاده از داده‌های مرجع تهیه شده مرحله ارزیابی دقت برای تک تک تصاویر انجام پذیرفت. تعداد پیکسل‌های مرجع به کار برده شده در ارزیابی دقت، برای تصاویر تقریباً بین ۳۰۰ تا ۵۰۰ پیکسل (بسته به تعداد کلاس‌ها) می‌باشد. سپس پارامترهای کاپا، دقت کلی، مغایرت کمی و مغایرت تخصیصی محاسبه شد و در مرحله آخر ضریب همبستگی<sup>۲</sup> بین این پارامترها به صورت دو به دو محاسبه گردید.

### ۳. تحلیل نتایج

مقادیر پنج پارامتر ارزیابی طبقه‌بندی شامل کاپا، دقت کلی، مغایرت کمی، مغایرت تخصیصی و مغایرت کلی برای هر تصویر بر اساس قدرت تفکیک مکانی تصاویر، در جدول شماره ۱ آورده شده است. بر اساس نتایج پارامترهای ارزیابی دقت طبقه‌بندی تصاویر مورد آزمایش، کمترین مقدار کاپا و دقت کلی به ترتیب برابر ۳۸,۷۳ و ۵۲,۹۵ درصد (تصویر شماره ۶، ماهواره لندست ۸) می‌باشد. همچنین بیشترین مقدار مغایرت کمی، مغایرت تخصیصی و مغایرت کلی به ترتیب برابر ۲۰,۵۹، ۲۶,۶۶ و ۴۷,۲۶ درصد (تصویر شماره ۶، ماهواره لندست ۸) می‌باشد. در ارزیابی طبقه‌بندی تصاویر، تصویر شماره ۶ دارای کمترین مقدار پارامترهای مطابقت (کاپا و دقت کلی) و بیشترین مقدار پارامترهای مغایرت می‌باشد. این تصویر دارای ۵ کلاس طیفی شامل سه کلاس خاک، یک کلاس پوشش گیاهی و یک کلاس آب است. به دلیل شباهت زیاد طیفی عوارض، توپوگرافی منطقه و وجود سایه در تصویر همبستگی طیفی بین کلاس‌ها افزایش یافته؛ در نتیجه دقت طبقه‌بندی در این تصویر کاهش چشم‌گیری داشته است. این موضوع در مورد تصویر شماره ۳۵ به صورت معکوس صادق است. برای این تصویر در بین همه ۵۷ مورد، بیشترین مقدار کاپا و بیشترین مقدار دقت کلی (به ترتیب ۹۸,۲۱ درصد و ۹۸,۸۴ درصد) با کمترین مقدار پارامترهای مغایرت کمی، تخصیصی و کلی (به ترتیب ۱,۱۰ درصد و ۰ درصد و ۱,۱۰ درصد) دیده می‌شود. این تصویر دارای ۳ کلاس طیفی شامل دو کلاس پوشش

<sup>۱</sup> Maximum Likelihood

<sup>۲</sup> Correlation



گیاهی و یک کلاس خاک است. دلیل تعریف تعداد کلاس کم برای طبقه‌بندی این تصویر، تنوع پوششی کم و تمایز طیفی مناسب بین عوارض می‌باشد. همین امر باعث کاهش وابستگی طیفی بین کلاس‌ها و در نتیجه افزایش دقت طبقه‌بندی شده است. انحراف معیار مقادیر کاپا، دقت کلی، مغایرت کمی، مغایرت تخصیصی و مغایرت کلی به ترتیب برابر ۱۳،۱۵، ۱۰،۳۵، ۴،۱۷، ۴،۳۱ و ۸،۷۸ می‌باشد. روال رفتاری معکوس پارامترهای مغایرت با پارامترهای متداول ارزیابی دقت موضوعی قابل پیش‌بینی با توجه به تفاوت تعاریف و نحوه محاسبه این دو دسته پارامتر می‌باشد.

جدول ۱. مقادیر پارامترهای پنج‌گانه ارزیابی طبقه‌بندی برای هر تصویر

ماهواره	شماره تصویر	Kappa (%)	Overall Accuracy (%)	Quantity Disagreement (%)	Allocation Disagreement (%)	Total Disagreement (%)
Landsat	۱	۹۳،۶۸	۹۵،۵۱	۳،۶۶	۰،۵۱	۴،۱۸
	۲	۹۵،۲۵	۹۶،۵۰	۳،۱۲	۰	۳،۱۲
	۳	۹۷،۷۳	۹۸،۱۳	۱،۱۳	۰،۷۰	۱،۸۳
	۴	۹۰،۲۰	۹۲،۵۷	۴،۸۱	۰،۱۴	۴،۹۶
	۵	۷۳،۴۰	۷۹،۹۱	۱۱،۷۳	۰	۱۱،۷۳
	۶	۳۸،۷۳	۵۲،۹۵	۲۰،۵۹	۲۶،۶۶	۴۷،۲۶
	۷	۹۲،۶۳	۹۶،۶۵	۲،۶۵	۰،۲۵	۲،۹۰
	۸	۹۱،۱۶	۹۵،۹۹	۳،۳۳	۰،۰۷	۳،۴۰
	۹	۹۵،۳۲	۹۶،۸۲	۲،۷۱	۰	۲،۷۱
	۱۰	۹۲،۰۸	۹۵،۸۱	۳،۲۳	۰،۸۷	۴،۱۰
	۱۱	۸۹،۶۶	۹۴،۲۵	۱،۵۵	۴،۱۴	۵،۷۰
	۱۲	۹۶،۹۳	۹۸،۰۷	۱،۲۵	۰،۶۱	۱،۸۶
	۱۳	۹۵،۰۷	۹۷،۴۱	۲،۴۷	۰	۲،۴۷
	۱۴	۸۰،۰۶	۸۸،۲۳	۹،۶۴	۰،۶۶	۱۰،۳۱
	۱۵	۹۲،۲۹	۹۴،۹۰	۴،۵۰	۰	۴،۵۰
	۱۶	۷۳،۶۶	۸۳،۹۷	۱۱،۹۲	۰	۱۱،۹۲
RapidEye	۱۷	۹۷،۷۴	۹۸،۷۲	۱،۱۹	۰	۱،۱۹
	۱۸	۹۲،۱۴	۹۴،۳۵	۳،۵۵	۱،۳۵	۴،۹۱
	۱۹	۹۳،۸۱	۹۶،۰۴	۲،۸۵	۰،۷۸	۳،۶۳
	۲۰	۹۳،۹۵	۹۶،۵۲	۲،۹۶	۰	۲،۹۶
	۲۱	۹۴،۷۴	۹۶،۴۹	۲،۵۹	۰،۶۸	۳،۲۸
	۲۲	۸۱،۰۰	۸۸،۶۱	۹،۱۰	۰	۹،۱۰
	۲۳	۸۸،۷۹	۹۱،۶۸	۶،۲۲	۰،۹۲	۷،۱۴
	۲۴	۵۷،۳۷	۶۸،۲۴	۱۴،۶۲	۴،۴۲	۱۹،۰۵
	۲۵	۹۰،۳۱	۹۳،۴۰	۲،۱۷	۰،۶۸	۲،۸۶
	۲۶	۹۴،۳۹	۹۷،۴۴	۲،۳۲	۰	۲،۳۲
	۲۷	۹۴،۷۵	۹۷،۶۴	۱،۶۲	۰،۷۳	۲،۳۵
	۲۸	۸۹،۷۴	۹۳،۴۷	۴،۸۲	۰	۴،۸۲
	۲۹	۸۵،۹۲	۹۰،۷۴	۶،۹۷	۰	۶،۹۷
	۳۰	۹۴،۷۴	۹۶،۱۱	۳،۴۸	۰	۳،۴۸
	۳۱	۹۲،۳۷	۹۴،۹۶	۴،۲۶	۰	۴،۲۶
	۳۲	۸۹،۵۷	۹۳،۰۲	۵،۶۲	۰	۵،۶۲

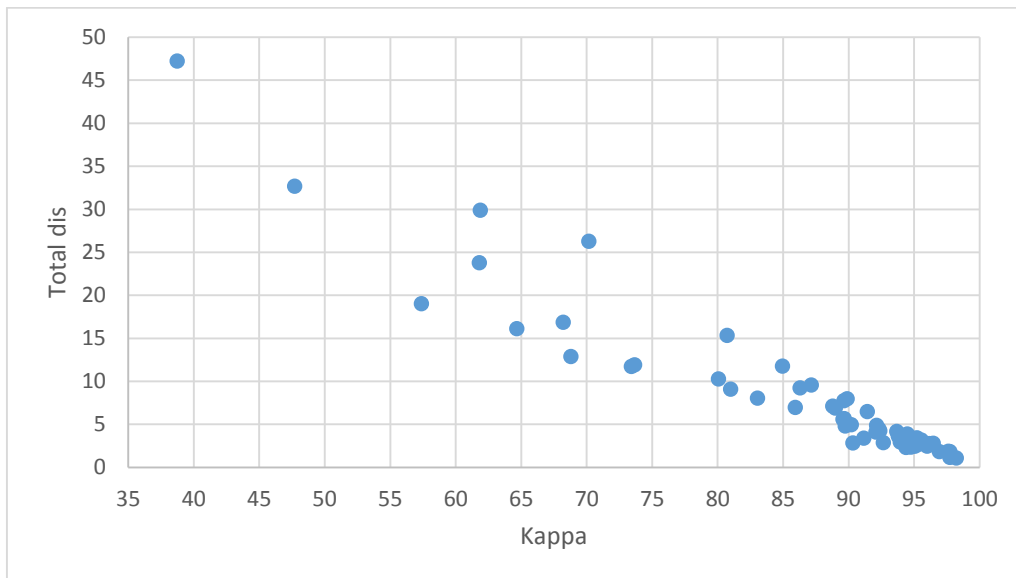
	۳۳	۹۵,۵۴	۹۷,۰۶	۲,۲۳	۰,۵۵	۲,۷۹
	۳۴	۶۴,۶۶	۷۶,۶۵	۱۶,۱۲	۰	۱۶,۱۲
	۳۵	۹۸,۲۱	۹۸,۸۴	۱,۱۰	۰	۱,۱۰
	۳۶	۹۵,۹۹	۹۷,۴۳	۲,۲۵	۰,۲۲	۲,۴۷
	۳۷	۶۸,۷۸	۷۹,۲۹	۱۲,۹۱	۰	۱۲,۹۱
	۳۸	۸۴,۹۵	۸۷,۷۹	۳,۴۱	۸,۳۵	۱۱,۷۶
	۳۹	۸۰,۷۲	۸۴,۲۳	۵,۳۳	۱۰,۰۲	۱۵,۳۶
	۴۰	۸۶,۲۸	۸۸,۸۲	۶,۹۴	۲,۳۳	۹,۲۷
	۴۱	۸۹,۸۹	۹۱,۶۱	۳,۲۸	۴,۷۰	۷,۹۹
	۴۲	۸۳,۰۳	۸۷,۳۶	۸,۰۸	۰	۸,۰
	۴۳	۹۶,۱۸	۹۷,۰۱	۲,۳۵	۰,۴۲	۲,۷۷
	۴۴	۹۷,۶۲	۹۸,۰۵	۱,۳۶	۰,۵۴	۱,۹۰
	۴۵	۸۹,۶۴	۹۱,۴۲	۵,۶۲	۲,۱۳	۷,۷۵
	۴۶	۹۴,۴۷	۹۵,۴۱	۳,۸۸	۰	۳,۸۸
	۴۷	۶۱,۸۰	۶۹,۵۴	۱۴,۵۴	۹,۲۶	۲۳,۸۰
Worldview <sup>۳</sup>	۴۸	۹۵,۵۴	۹۶,۲۱	۳,۱۸	۰	۳,۱۸
	۴۹	۸۷,۱۶	۸۹,۹۱	۴,۷۲	۴,۸۷	۹,۶۰
	۵۰	۶۸,۱۹	۷۳,۴۱	۷,۹۲	۸,۹۷	۱۶,۸۹
	۵۱	۴۷,۶۹	۵۶,۶۹	۸,۱۶	۲۴,۵۲	۳۲,۶۹
	۵۲	۹۶,۴۵	۹۷,۱۰	۱,۹۰	۰,۹۰	۲,۸۱
	۵۳	۹۵,۲۲	۹۶,۱۷	۲,۴۱	۱,۰۴	۳,۴۶
	۵۴	۶۱,۸۷	۶۹,۹۳	۴,۸۶	۲۵,۰۵	۲۹,۹۲
	۵۵	۸۸,۹۹	۹۱,۰۱	۶,۹۱	۰	۶,۹۱
	۵۶	۷۰,۱۷	۷۵,۷۱	۶,۳۴	۱۹,۹۵	۲۶,۳۰
	۵۷	۹۱,۴۱	۹۳,۰۰	۳,۸۴	۲,۶۴	۶,۴۹

مشاهدات کلی در جدول ۱ نشان می‌دهند که قدرت تفکیک چندان در تغییر نتایج مؤثر نیست و در همه موارد نتایج ضعیف و قوی دیده می‌شود. در واقع دقت نهایی طبقه‌بندی تصاویر به تعریف کلاس‌ها، دقت عملیات تمرینی و دقت طبقه‌بندی کننده بستگی دارد؛ که در این تحقیق سعی شده است شرایط برای همه تصاویر تقریباً مشابه باشد. برای مقایسه دو دسته معیارهای دقت متداول و معیارهای مغایرت جدید، میزان همبستگی بین آن‌ها محاسبه شد که نتایج آن در جدول ۲ ارائه شده است. ضریب همبستگی برای همه موارد آزمایش شده، مقداری منفی می‌باشد. منفی بودن ضریب همبستگی موضوعی قابل پیش بینی است که به خاطر تفاوت ذاتی دو دسته معیار می‌باشد. معیارهای کاپا و دقت کلی از جنس دقت هستند و میزان تطابق با داده‌های مرجع را بیان می‌کنند؛ ولی دو پارامتر جدید از جنس مغایرت و عدم تطابق با داده‌های مرجع می‌باشند.

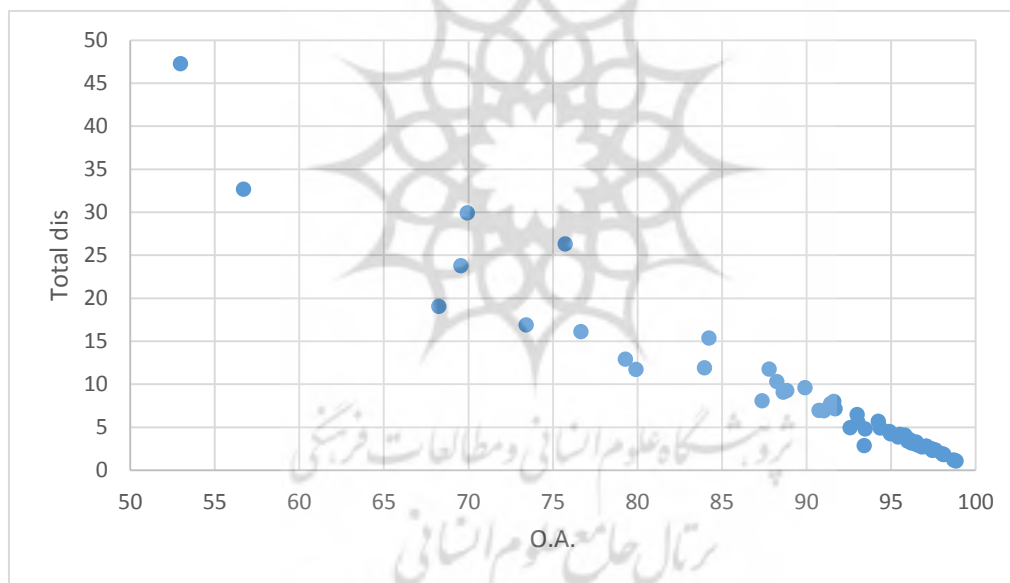
جدول ۲. مقادیر همبستگی بین معیارهای متداول و جدید ارزیابی طبقه‌بندی تصاویر

ماهواره	Kappa & Allocation Disagreement	Kappa & Quantity Disagreement	O.A. & Allocation Disagreement	O.A. & Quantity Disagreement	Kappa & Total Disagreement	O.A. & Total Disagreement
All Images	-۰,۸۲	-۰,۸۶	-۰,۸۰	-۰,۷۵	-۰,۹۵	-۰,۹۶
Landsat ۸	-۰,۸۴	-۰,۹۶	-۰,۸۶	-۰,۹۵	-۰,۹۷	-۰,۹۸
RapidEye	-۰,۵۰	-۰,۹۷	-۰,۵۶	-۰,۹۵	-۰,۹۸	-۰,۹۸
Worldview ۳	-۰,۸۷	-۰,۷۰	-۰,۸۷	-۰,۶۹	-۰,۹۶	-۰,۹۶

نتایج آنالیز همبستگی به دست آمده برای کل تصاویر، نشان می‌دهد که در بسیاری موارد همبستگی بالایی بین پارامترهای متداول و پارامترهای جدید وجود دارد. همبستگی بالا بین این دو دسته پارامتر ارزیابی دقت، به این معنی است که از لحاظ رفتاری بسیار مشابه یکدیگر هستند. به عبارت دیگر، با تغییر مقدار یکی از پارامترهای دقت متداول مقدار پارامترهای مغایرت نیز تقریباً به همان تناسب تغییر می‌کند. بنابراین می‌توان نتیجه گرفت دو پارامتر جدید معرفی شده توسط Pontius Jr و Millones، کاملاً مجزا از کاپا و دقت کلی نیستند و تفاوت رفتاری جدی ندارند. این موضوع به خصوص برای پارامتر مغایرت کلی برای همه تصاویر با ضریب همبستگی ۰,۹۶ پررنگ‌تر است. پارامتر مغایرت کلی برای تصاویر به تفکیک سنجنده نیز همبستگی بالایی با پارامترهای متداول دقت نشان داده است. در جدول ۲ در حالی که همه تصاویر در آنالیز همبستگی شرکت کرده‌اند (سطر اول) کمترین ضریب همبستگی بین دقت کلی و مغایرت کمی با مقدار ۰,۷۵- و بیشترین ضریب همبستگی بین ضریب کاپا و مغایرت کمی با مقدار ۰,۸۶- اتفاق افتاده است. با توجه به نحوه محاسبه ضرایب کاپا و دقت کلی، و همچنین پارامتر مغایرت کلی، ضریب همبستگی بالایی بین این معیارها وجود دارد (به ترتیب ۰,۹۵- و ۰,۹۶-). شکل‌های ۲ و ۳ نمودار پراکندگی دو معیار متداول نسبت به معیار مغایرت کلی برای همه تصاویر مورد آزمایش را نشان می‌دهند. همبستگی منفی بالایی بین این دو معیار کاملاً مشاهده می‌گردد.



شکل ۲. پراکندگی مقادیر پارامترهای کاپا (%) و مغایرت کلی (%). وجود یک همبستگی منفی بالا بین این دو معیار کاملاً مشاهده می‌گردد.



شکل ۳. پراکندگی مقادیر پارامترهای دقت کلی (%) و مغایرت کلی (%). وجود یک همبستگی منفی بالا بین این دو معیار کاملاً مشاهده می‌گردد.

اگر بخواهیم در جدول ۲ بر اساس تصاویر سنجنده‌های مختلف مورد آزمایش، مقادیر همبستگی را تحلیل کنیم اولین مورد قابل مشاهده، مشابهت رفتاری دو پارامتر دقت کلی و ضریب کاپا می‌باشد. به این معنی که برای هر سه سنجنده مورد مقایسه، هر جا ضریب کاپا با مقادیر مغایرت کمی و تخصیصی همبستگی بالایی داشته است، دقت کلی نیز چنین رفتاری را از خود نشان داده است و بالعکس، یعنی با کم شدن میزان همبستگی ضریب کاپا، مقدار همبستگی دقت کلی نیز کم شده است. در میان سنجنده‌های مورد آزمایش، ضرایب کاپای محاسبه شده برای سنجنده RapidEye کمترین همبستگی را با معیار مغایرت تخصیصی

داشته است (۵، ۰-). با توجه جدول شماره ۱ متوجه می‌شویم که این موضوع مربوط به وجود مقادیر صفر بیشتر در ستون مغایرت تخصیصی برای این سنجنده نسبت به دو سنجنده دیگر است. همانطور که از معادله شماره ۶ مشخص است محاسبه این پارامتر وابسته به مقادیر صفر در ماتریس ابهام است. لذا در صورتی که برخی کلاس‌ها در ماتریس ابهام هیچگونه تداخلی با یکدیگر نداشته باشند آنگاه به سلول مربوط به آن کلاس‌ها مقدار صفر تعلق می‌گیرد و در نتیجه مقدار مغایرت تخصیصی آن‌ها نیز صفر می‌شود. در صورتی که مقدار کاپا یا دقت کلی اینگونه نخواهند بود. نتیجه عملی این موضوع، پایین آمدن مقدار ضریب همبستگی دو معیار متداول ارزیابی دقت با مقدار مغایرت تخصیصی است؛ که برای سنجنده RapidEye دقیقاً این اتفاق افتاده است.

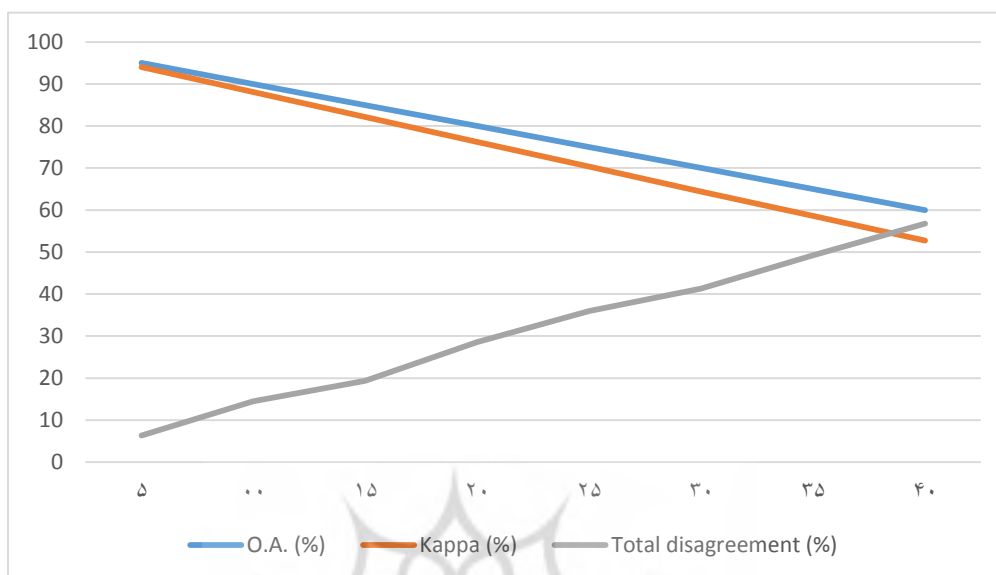
مقدار همبستگی مغایرت کلی با دو پارامتر متداول برای هر سه سنجنده بالا می‌باشد (حداقل ۰،۹۶-). در مقابل یک اتفاق جالب در جدول شماره ۲ مشاهده می‌گردد. داده‌ها برای دو سنجنده OLI Landsat و RapidEye نشان می‌دهند که مغایرت تخصیصی نسبت به مغایرت کمی، همبستگی کمتری با دو پارامتر دقت کلی و کاپا نشان داده است. در صورتی که در تصاویر مربوط به سنجنده WorldView<sup>۳</sup> نتایج دقیقاً برعکس می‌باشد و میزان همبستگی برای مغایرت تخصیصی بیشتر از مقدار همبستگی مغایرت کمی با دو پارامتر کاپا و دقت کلی می‌باشد. در تصاویر با قدرت تفکیک بالای WorldView<sup>۳</sup> تعریف کلاس‌های بیشتری ممکن است و به همین خاطر تداخل کلاس‌ها نیز امکان بیشتری دارد. لذا توزیع خطاها در ماتریس ابهام یکنواخت‌تر است و به همین دلیل مغایرت تخصیصی دقیق‌تر محاسبه می‌گردد. لذا همبستگی این پارامتر نیز بالاتر از مغایرت کمی شده است. در این تحقیق جهت اعتبارسنجی نتایج به دست آمده، از یک داده مرجع شبیه‌سازی شده نیز استفاده شده است. داده مرجع استفاده شده، یک ماتریس به ابعاد ۲۰۰\*۲۰۰ پیکسل و حاوی ۷ کلاس می‌باشد. به این صورت عملاً یک نقشه مرجع واقعیت زمینی با ۷ کلاس تولید شده است. برای بررسی رابطه پارامترهای جدید و متداول ارزیابی طبقه‌بندی در تکرارهای منظم، به صورت تصادفی درصدی از داده مرجع تغییر داده شد و با استفاده از داده مرجع اصلی، پارامترهای کاپا، دقت کلی و مغایرت کلی، محاسبه گردید. جدول شماره ۳ مقادیر این پارامترها را متناسب با درصد پیکسل‌های دارای برچسب اشتباه نشان می‌دهد.

جدول شماره ۳. درصد پیکسل‌های با برچسب تغییر داده شده در نقشه مرجع به همراه مقادیر پارامترهای متداول و مغایرت کلی مربوطه

درصد تغییرات/ پارامتر	۵	۱۰	۱۵	۲۰	۲۵	۳۰	۳۵	۴۰
Overall Accuracy (%)	۹۵،۰۰	۹۰،۰۰	۸۵،۰۰	۸۰،۰۰	۷۵،۰۰	۷۰،۰۰	۶۵،۰۰	۶۰،۰۰
Kappa (%)	۹۴،۰۳	۸۵،۸۰	۸۲،۱۴	۷۶،۲۲	۷۰،۳۲	۶۴،۴۳	۵۸،۵۷	۵۲،۷۰
Total Disagreement (%)	۶،۳۷	۱۴،۵۰	۱۹،۳۶	۲۸،۵۷	۳۶،۰۱	۴۱،۳۲	۴۹،۲۲	۵۶،۸۰

شکل ۴ نتیجه حاصل از این اعتبارسنجی را نشان می‌دهد. محور افقی این نمودار، درصد تغییرات داده مرجع و محور عمودی مقادیر مربوط به سه پارامتر دقت کلی، کاپا و مغایرت کلی را نشان می‌دهد. در این شکل، علاوه بر تبعیت کامل دو پارامتر متداول

از یکدیگر، مشاهده می‌شود که تغییرات کاپا و دقت کلی در تناسب با تغییرات پارامتر مغایرت کلی می‌باشد و همبستگی بالایی بین این پارامترهای ارزیابی دقت و مغایرت نتایج طبقه‌بندی تصاویر مشاهده می‌گردد.



شکل ۴. نمودار تغییرات پارامترهای دقت کلی، کاپا و مغایرت کلی با تغییرات داده مرجع شبیه‌سازی شده. محور افقی درصد تغییرات داده مرجع و محور عمودی مربوط مقادیر پارامترهای مذکور می‌باشد.

#### ۴. نتیجه گیری

ارزیابی دقت نتایج حاصل از طبقه بندی تصاویر ماهواره ای، مرحله ای ضروری برای تشخیص ارزش طبقه بندی انجام شده می‌باشد. برخی محققین معتقدند پارامترهای متداول ارزیابی دقت طبقه بندی مانند کاپا و دقت کلی، علی رغم کاربردهای فراوانی که در کارهای تحقیقاتی دارند، دارای معایبی هستند؛ از این رو توسط محققین مختلف پارامترهای جدید ارزیابی دقت بیان شده است. در این تحقیق به بررسی دو نمونه از این پارامترهای جدید به نام مغایرت کمی و مغایرت تخصیصی پرداخته شد. به این منظور نتایج طبقه‌بندی نظارت شده روی مجموعه‌ای از تصاویر توسط پارامترهای متداول و جدید ارزیابی دقت گردیده و در نهایت بین مقادیر چهارگانه پارامترهای مذکور، میزان همبستگی محاسبه شد. در این تحقیق نتایج حاصل از طبقه‌بندی و همچنین اعتبارسنجی نتایج با استفاده از داده شبیه سازی شده نشان می‌دهد که بر خلاف ادعاهای ارائه کنندگان پارامترهای مغایرت کمی و مغایرت تخصیصی، این پارامترهای جدید مستقل از پارامترهای متداول نیستند. تنها تفاوت بین پارامترهای معرفی شده توسط Pontius Jr و Millones با پارامتر کاپا، در نحوه گزارش دقت طبقه بندی بر اساس مغایرت می‌باشد. نتایج تحقیق نشان داد که تغییر قدرت تفکیک تصاویر اثر جدی روی نتایج رفتاری این دو دسته پارامتر ندارد. تنها در مواردی که ماتریس ابهام در اثر عدم توزیع یکنواخت خطا دارای عناصر صفر زیاد باشد میزان پارامتر مغایرت تخصیصی همبستگی کمتری با پارامترهای متداول دارد. به عنوان یک نتیجه‌گیری کلی می‌توان عنوان نمود که پارامترهای کاپا و دقت کلی به منظور ارزیابی

دقت طبقه بندی تصاویر ماهواره ای علیرغم ایراداتی که به آن‌ها وارد است همچنان معتبر می‌باشند و در صورتی که خطای طبقه بندی خواسته شود، می‌توان از پارامترهای مغایرت نیز در کنار آن‌ها استفاده نمود.

## منابع

- Congalton, R. G., & Green, K. (2008). *Assessing the accuracy of remotely sensed data: principles and practices*: CRC press.
- El-Aziz, M. A. E.-A. (2004). Evaluation of soft classifiers for remote sensing data .
- Emami, H. (2005). Introducing correctness coefficient as an accuracy measure for sub pixel classification results.
- Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote sensing of environment*, 80(1), 185-201 .
- Foody, G. M., & Arora, M. K. (1996). Incorporating mixed pixels in the training, allocation and testing stages of supervised classifications. *Pattern Recognition Letters*, 17(13), 1389-1398 .
- Gu, J., Sun, G., Pan, Y., & Fan, D. (2012). *Accuracy assessment based on the distribution of the classified errors on classified map: Simulation analysis*. Paper presented at the Agro-Geoinformatics (Agro-Geoinformatics), 2012 First International Conference on.
- Hashemian, M., Abkar, A., & Fatemi, S. (2004). *Study of sampling methods for accuracy assessment of classified remotely sensed data*. Paper presented at the International congress for photogrammetry and remote sensing.
- Li, W., & Guo, Q. (2014). A new accuracy assessment method for one-class remote sensing classification. *IEEE Transactions on Geoscience and Remote Sensing* .4632-4621 ,(8)52 ,
- Marpu, P. R., Wijaya, A., & Gloaguen, R. (2008). *Soft classification and assessment of kalman filter neural network for complex landcover of tropical rainforests*. Paper presented at the Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International.
- MASELLI, F., Rodolfi, A., & Conese, C. (1996). Fuzzy classification of spatially degraded Thematic Mapper data for the estimation of sub-pixel components. *International Journal of Remote Sensing*, 17(3), 537-551 .
- Mathers, P. (1999). *Computer Processing of Remotely-Sensed Images*: John Wiley & Sons.
- Persello, C., & Bruzzone, L. (2010). A novel protocol for accuracy assessment in classification of very high resolution images. *IEEE Transactions on Geoscience and Remote Sensing*-1232 ,(3)48 , .1244
- Pickard, B., Gray, J., & Meentemeyer, R. (2017). Comparing Quantity, Allocation and Configuration Accuracy of Multiple Land Change Models. *Land*, 6(3), 52 .
- Pontius Jr, R. G., & Millones, M. (2011). Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32(15), 4407-4429 .
- Pontius, R. G. (2000). Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering and remote sensing*, 66(8), 1011-1016 .
- Richards, J. A. (2013). *Remote sensing digital image analysis*: Springer.
- Story, M., & Congalton, R. G. (1986). Accuracy assessment: a user's perspective. *Photogrammetric Engineering and remote sensing*, 52(3), 397-399 .
- Strahler, A. H., Boschetti, L., Foody, G. M., Friedl, M. A., Hansen, M. C., Herold, M., . . . Woodcock, C. E. (2006). Global land cover validation: Recommendations for evaluation and accuracy assessment of global land cover maps. *European Communities, Luxembourg*, 51(4) .(
- Sui, C., Tian, Y., & Xu, Y. (2014). *An Unsupervised Band Selection Method Based on Overall Accuracy Prediction*. Paper presented at the Pattern Recognition (ICPR), 2014 22nd International Conference on.

- Warrens, M. J. (2015). Relative quantity and allocation disagreement measures for category-level accuracy assessment. *International Journal of Remote Sensing*, 36(23), 5959-5969 .
- Wilkinson, G. G. (2005). Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3), 433-440 .
- Yan, G., Mas, J. F., Maathuis, B., Xiangmin, Z., & Van Dijk, P. (2006). Comparison of pixel-based and object-oriented image classification approaches—a case study in a coal fire area, Wuda, Inner Mongolia, China. *International Journal of Remote Sensing*, 27(18), 4039-4055 .
- Yi, L., & Zhang, G. (2012). *Object-oriented remote sensing imagery classification accuracy assessment based on confusion matrix*. Paper presented at the Geoinformatics (GEOINFORMATICS), 2012 20th International Conference on.
- Zhang, J., & Foody, G. (1998). A fuzzy classification of sub-urban land cover from remotely sensed imagery. *International Journal of Remote Sensing*, 19(14), 2721-2738 .
- Zhou, G., Wu, B & ,Li, M. (2011). *Improved accuracy assessment indices for object-based high resolution remotely sensed imagery classification*. Paper presented at the Image Analysis and Signal Processing (IASP), 2011 International Conference on.





## **A Comparative study of the traditional accuracy assessment and disagreement measures of the classification of remote sensing imagery**

Mehran Dadjoo<sup>۱</sup>, Sayyed Bagher Fatemi Nasrabadi<sup>۲\*</sup>

۱. Student master of Remote sensing, Department of Geomatics Engineering, Faculty of Civil Engineering and Transportation, University of Isfahan
۲. Assistant Professor, Department of Geomatics Engineering, Faculty of Civil Engineering and Transportation, University of Isfahan

\*Correspondence Address: Hezar Jarib Avenue, University of Isfahan, Faculty of Civil Engineering and Transportation, Department of Geomatics Engineering, ۸۱۷۴۶-۷۳۴۴۱, Isfahan, Iran,

Tel: +۹۸ ۳۱ ۳۷۹۳۰۳۳۲, Fax: +۹۸ ۳۱ ۳۷۹۳۲۲۳۱, Email: [sb.fatemi@eng.ui.ac.ir](mailto:sb.fatemi@eng.ui.ac.ir), ORCID: ۱۱۱۱-۰۰۰۲-۱۷۹۹-۷۱۶۶

### **Abstract**

Evaluation of the image classification results is very important in the remote sensing projects. So far, many indices have been presented to assess the accuracy of image classification, though Kappa coefficient and Overall accuracy are the most famous ones. Some researchers have criticized these two parameters, and have presented new parameters for evaluation of the classification results. In this paper, the relation between two new accuracy assessment parameters (presented by Pontius & Millones) and traditional accuracy assessment parameters (Overall accuracy and kappa coefficient) is studied. These two new parameters are called “Quantity disagreement” and “Allocation disagreement” which report disagreement between ground truth and classification data. In order to apply the comparative study on the traditional and new disagreement measures, supervised maximum likelihood classification was applied on ۵۷ satellite images with different spatial resolutions. Then, Kappa and Overall accuracy as traditional accuracy parameters and Quantity disagreement and Allocation disagreement as new measures were computed for each classified image and then the correlation coefficients of the both measures were calculated. The results show a high correlation between new parameters and traditional ones in negative direction irrespective the spatial resolution. In this way, the disagreement do not provide new information about the classification results to the user, and only if there is any request for classification error, the new disagreement parameters can be used along with the traditional ones.

**Key words:** Accuracy Assessment, Classification, Disagreement, Overall accuracy, Traditional Parameters.