



## مقایسه توانایی طبقه‌بندی‌کننده‌های ماشین‌بردار پشتیبان، شبکه عصبی و حداکثر احتمال در بهبود نتایج الگوریتم‌های فیلترینگ داده‌های لیدار

علی‌اکبر متکان<sup>۱</sup>، محمد حاجب\*<sup>۲</sup>

۱. دانشیار گروه سنجش از دور و GIS، دانشگاه شهید بهشتی

۲. مربی گروه سنجش از دور و GIS، دانشگاه شهید بهشتی

تاریخ پذیرش مقاله: ۱۳۹۱/۸/۲۲

تاریخ دریافت مقاله: ۱۳۹۰/۱۰/۷

### چکیده

بررسی‌های اخیر نشان داده است که سیستم لیدار در برداشت سریع و دقیق اطلاعات سه‌بعدی از عوارض، در مناطق شهری توانایی بالایی دارد. از مهم‌ترین پردازش‌هایی که روی داده‌های لیدار صورت می‌گیرد، فیلترینگ آن‌هاست که عبارت است از تفکیک نقاط مربوط به عوارض ارتفاعی (ساختمان‌ها و درختان) از نقاط زمینی. تاکنون الگوریتم‌های فیلترینگ فراوانی طراحی شده است، اما هر یک از آن‌ها معایب و نواقصی دارند. مشکل اساسی این الگوریتم‌ها، ناتوانی‌شان در حذف ساختمان و عوارض بزرگ است که ناشی از عملکرد ناحیه‌ای آن‌هاست. در این تحقیق تلاش شده است تا با بهره‌گیری از قابلیت‌های تکنیک‌های طبقه‌بندی، راه‌حلی برای این مشکل ارائه شود. در این پژوهش ابتدا داده‌های لیدار به‌وسیله الگوریتم شیب‌مبنا که - از شناخته‌شده‌ترین روش‌های فیلترینگ به‌شمار می‌آید- فیلتر شدند. در ادامه، با استفاده از سه روش طبقه‌بندی ماشین‌بردار پشتیبان، شبکه عصبی مصنوعی، و حداکثر احتمال، داده‌های لیدار به پنج کلاس راه آسفالت، درختان، ساختمان، سیمان، و چمن طبقه‌بندی گردیدند. در نهایت، نقاط ساختمان‌های بزرگ که به‌وسیله الگوریتم شیب‌مبنا فیلتر نشده بودند، با نتایج حاصل از روش‌های طبقه‌بندی حذف گردیدند. ارزیابی‌ها نشان می‌دهند که روش حداکثر احتمال نتایج ضعیفی را ارائه می‌کند، اما روش‌های ماشین‌بردار پشتیبان و شبکه عصبی مصنوعی نتایج نزدیک به هم و بسیار خوبی را عرضه کرده‌اند. به‌طور کلی استفاده از این تکنیک‌های طبقه‌بندی برای بهبود نتایج الگوریتم‌های فیلترینگ، باعث افزایش ناچیزی در خطای نوع اول می‌گردد ولی سبب کاهش شدیدی در خطای نوع دوم و خطای مجموع می‌شود. از آنجا که در فرایند فیلترینگ داده‌های لیدار، اهمیت خطاهای نوع دوم و مجموع بیش از خطای نوع اول است، می‌توان ادعا کرد که انجام این پردازش تکمیلی نتایج بسیار مفیدی را دربرداشته است. ارزیابی کمی دقت نتایج نشان می‌دهد که خروجی الگوریتم با آستانه شیب  $20^\circ$  پس از بهبود داده شدن با استفاده از خروجی طبقه‌بندی‌کننده شبکه عصبی مصنوعی، بهترین نتیجه را به دست داده است که در آن خطای نوع اول از  $4/98$  درصد به  $5/04$  درصد افزایش ولی خطاهای نوع دوم و مجموع به‌ترتیب از  $9/043$  درصد و  $7/03$  درصد به  $4/49$  درصد و  $4/76$  درصد کاهش یافته‌اند.

**کلیدواژه‌ها:** لیدار، فیلترینگ، الگوریتم فیلترینگ شیب‌مبنا، طبقه‌بندی‌کننده حداکثر احتمال، ماشین‌های بردار پشتیبان، شبکه‌های عصبی مصنوعی.

\* نویسنده مکاتبه‌کننده: تهران، اوین، دانشگاه شهید بهشتی، گروه سنجش از دور و GIS. تلفن: ۰۹۱۲۲۰۱۵۹۳۱

## ۱- مقدمه

کار می‌برند (Axelsson, ۱۹۹۹; Axelsson, ۲۰۰۰). اکثر الگوریتم‌های فوق فقط از داده‌های فاصله و آن هم بازگشت آخر لیدار استفاده می‌کنند، در صورتی که سیستم‌های عملیاتی لیدار می‌توانند چندین بازگشت پالس‌های لیزر و نیز اندازه‌گیری قدرت پالس برگشتی را ثبت کنند. بدیهی است که استفاده هر چه بیشتر از امکاناتی که با پیشرفت فناوری برداشت اطلاعات به وسیله سیستم لیدار در اختیار قرار می‌گیرد، به بالا رفتن دقت و بهبود نتایج قبلی می‌انجامد. بنابراین، به طراحی الگوریتم‌های فیلترینگ که از بازگشت‌های اول و آخر لیزر و نیز داده‌های شدت استفاده کنند، نیاز است. برای این منظور، گروهی دیگر از الگوریتم‌های فیلترینگ ارائه شده‌اند که از روش‌های طبقه‌بندی برای فیلترینگ داده‌های لیدار استفاده می‌کنند.

صلاح و ترانیدر (Salah & Trinder, ۲۰۱۰) از تکنیک طبقه‌بندی ماشین‌های بردار پشتیبان<sup>۶</sup> بهره گرفته‌اند و علاوه بر داده‌های فاصله لیدار، از داده‌های شدت و نیز عکس رنگی هوایی منطقه و همچنین لایه‌هایی که از آنالیز بافت تصاویر فوق به دست آمده‌اند، استفاده کرده‌اند. بلیغ و دیگران (Baligh et al., ۲۰۱۱) از تکنیک طبقه‌بندی شبکه‌های عصبی مصنوعی<sup>۷</sup> استفاده کرده‌اند. در این پژوهش، ابتدا داده‌های آموزشی در دو کلاس عارضه و زمین، به روشی نیمه خودکار و با استفاده از معیارهایی براساس شیب و نیز اختلاف بین بازگشت اول و آخر شناسایی می‌شوند. در آخر، نقاط زمین لخت و عارضه در کل منطقه با

سیستم لیدار<sup>۱</sup> قادر به نمونه‌برداری از سطح زمین در مدت زمان کوتاه با دقتی مناسب و در قالب سه‌بعدی است، به همین دلیل در سال‌های اخیر استفاده از آن به سرعت گسترش یافته است. با این حال به دلیل حجم زیاد داده‌ها، پردازش آنها به صورت اتوماتیک همچنان چالشی بزرگ قلمداد می‌شود. یکی از مهم‌ترین پردازش‌های داده‌های لیدار، فیلترینگ داده‌ها یا همان جداسازی نقاط متعلق به زمین لخت<sup>۲</sup> از نقاط متعلق به عوارض است. آگاهی از زمین لخت برای بسیاری از طرح‌های عمرانی و برنامه‌ریزی‌ها ضروری است. از آنجا که استخراج زمین لخت از داده‌های لیدار به صورت دستی، زمان‌بر و پرهزینه است، تلاش‌های زیادی برای طراحی الگوریتم‌های فیلترینگ خودکار داده‌های لیدار انجام شده است. این الگوریتم‌ها را براساس خصوصیات‌شان به گروه‌هایی دسته‌بندی می‌کنند. دسته‌ای از آنها الگوریتم‌هایی هستند که براساس قطعه‌بندی<sup>۳</sup> یا خوشه‌بندی<sup>۴</sup> ابرنقاط لیدار کار می‌کنند (Roggero, ۲۰۰۱; Brovelli, et al., ۲۰۰۲; Filin, ۲۰۰۲; Jacobsen & Lohmann, ۲۰۰۳; Tovari & Pfeifer, ۲۰۰۵; Sithole & Vosselman, ۲۰۰۵; Rabbania, et al., ۲۰۰۶). دسته‌ای دیگر، از مدل‌های شکل فعال<sup>۵</sup> برای فیلترکردن داده‌های لیدار استفاده می‌کنند (Elmqvist, ۲۰۰۲; Elmqvist, et al., ۲۰۰۱; Elmqvist, ۲۰۰۲). گروهی از الگوریتم‌های فیلترینگ از مفاهیم مورفولوژی ریاضی برای حذف نقاط عوارض ارتفاعی بهره می‌گیرند (Zhang, et al., ۲۰۰۳; Arefi & Hahn, ۲۰۰۵; Chen, et al., ۲۰۰۷; Kobler, et al., ۲۰۰۹; Meng et al., ۲۰۰۷). دسته دیگری از الگوریتم‌های فیلترینگ از روش اسکن کردن جهت‌دار ابرنقاط و محاسبه شیب یا اختلاف ارتفاع در طول خط اسکن استفاده می‌کنند (Vosselman, ۲۰۰۰; Sithole, ۲۰۰۵; Shan & Sampath, ۲۰۰۵; Meng, ۲۰۰۹b; Meng, et al., ۲۰۰۵). گروهی از الگوریتم‌ها نیز روش‌های درون‌یابی یا TIN‌های متراکم‌شونده را به

۱. LIDAR (Light Detection and Ranging)

۲. Bare Earth

۳. Segmentation

۴. Clustering

۵. Shape active

۶. Support Vector Machine (SVM)

۷. Artificial Neural Network (ANN)

است که پایه چندی از الگوریتم‌های دیگر است و دیگر این‌که واسلمن و ماس (Vosselmann & mass, ۲۰۰۱) به ناتوانی الگوریتم شیب‌مبنا در حذف ساختمان‌های بزرگ اشاره کرده‌اند. آنها (۲۰۰۱) عنوان کردند که اگر ابعاد همسایگی در نظر گرفته‌شده، کوچک‌تر از نصف ابعاد یک ساختمان باشد، آن ساختمان به‌وسیله الگوریتم شیب‌مبنا فیلتر نمی‌شود.

بخش دوم این پژوهش، به معرفی داده‌های استفاده‌شده و همچنین، تشریح الگوریتم پیشنهادشده، اختصاص دارد. بخش سوم به ارائه نتایج اعمال الگوریتم روی داده‌ها پرداخته است. تحلیل نتایج در بخش چهارم آورده شده است، و بخش پنجم به نتیجه‌گیری می‌پردازد.

## ۲- مواد و روش‌ها

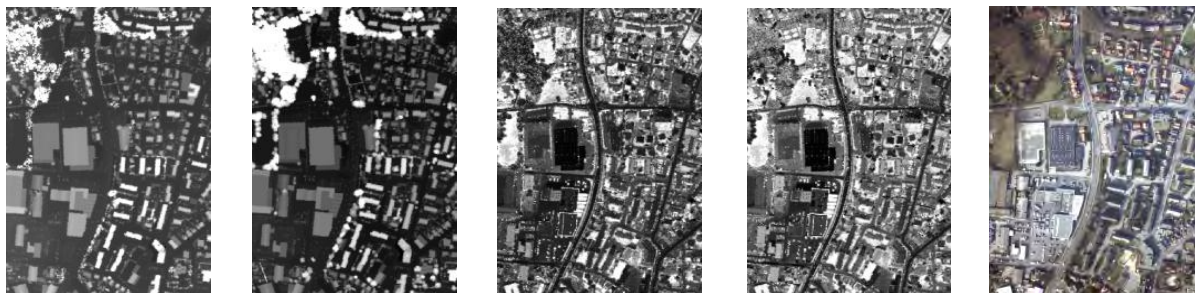
### ۲-۱- مشخصات داده‌های استفاده‌شده

داده‌های لیدار استفاده‌شده در این تحقیق شامل داده‌های بازگشت اول و آخر فاصله و شدت است. داده‌ها ساختار شبکه‌ای منظم دارند و فاصله بین نقاط شبکه ۱ متر است. تصاویر مربوط به داده‌های لیدار به ابعاد ۶۹۷×۴۷۲ پیکسل هستند. تصویری هوایی با قدرت تفکیک بالا (۰/۲ متر) نیز از منطقه وجود دارد. این داده‌ها مربوط به شهر Castrop-Rauxel در غرب کشور آلمان است و شرکت Toposys آنها را تهیه کرده است.

استفاده از شبکه عصبی طبقه‌بندی می‌شوند.

با وجود الگوریتم‌های فراوان فیلترینگ داده‌های لیدار، اکثر این الگوریتم‌ها فقط در مناظر ساده نتایج خوبی به دست می‌دهند و در مواجهه با مناظر شهری پیچیده مشکلاتی دارند. از این رو، نیاز به رفع نواقص روش‌های قبلی و طراحی الگوریتم‌های جدید به شدت احساس می‌شود. سیتول و واسلمن مقایسه‌ای بین برخی از الگوریتم‌ها صورت داده‌اند (Sithole & Vosselmann, ۲۰۰۳). در این گزارش عوارض بزرگ در منطقه به‌عنوان یکی از مشکلات پیش روی الگوریتم‌های فیلترینگ مطرح شده و علت آن عملکرد محلی الگوریتم‌ها ذکر شده است. الگوریتم‌های فیلترینگ برای در نظر گرفتن اثر ناهمواری‌های زمین منطقه در مناطق با توپوگرافی پیچیده، مجبورند در تصمیم‌گیری برای عارضه یا زمینی بودن یک نقطه، به صورت محلی و در یک همسایگی عمل کنند. بدین ترتیب این الگوریتم‌ها هنگام برخورد با ساختمان‌های بزرگ‌تر از این همسایگی دچار مشکل می‌شوند.

در این تحقیق سعی شده است تا برای حل مشکل ساختمان‌های بزرگ روشی پیشنهاد شود. بدین منظور الگوریتم فیلترینگ شیب‌مبنا (Vosselmann, ۲۰۰۰) به‌عنوان الگوریتم پایه انتخاب و تلاش شده است تا با بهره‌گیری از روش‌های طبقه‌بندی، ساختمان‌های بزرگی که الگوریتم شیب‌مبنا قادر به حذف آن‌ها نبوده است، فیلتر شوند. دلایل انتخاب الگوریتم شیب‌مبنا به‌عنوان الگوریتم پایه این است که این الگوریتم از مؤثرترین و شناخته‌شده‌ترین الگوریتم‌های فیلترینگ



ه

د

ج

ب

الف

شکل ۱. داده‌های منطقه مطالعه شده؛ الف) تصویر هوایی ب) بازگشت اول شدت ج) بازگشت آخر شدت د) بازگشت اول فاصله  
ه) بازگشت آخر فاصله

همان‌طور که در شکل ۱ مشاهده می‌شود، تغییرات فراوانی در بافت منطقه وجود دارد. در سمت راست تصویر، مناطق مسکونی با ساختمان‌های تقریباً کوچک و در سمت چپ، مناطق صنعتی با ساختمان‌های بزرگ دیده می‌شود. در گوشه سمت چپ بالای تصویر، پوشش گیاهی نسبتاً متراکم وجود دارد و درختان بزرگ و کوچک دیگری نیز به‌طور پراکنده در نقاطی از منطقه مشاهده می‌شوند. بدین ترتیب منطقه پیچیدگی لازم از نظر انواع ساختمان‌ها و عوارض با ابعاد گوناگون را برای ارزیابی روش پیشنهادی دارد.

## ۲-۲-۲ روش انجام کار

در این تحقیق، پس از حذف نقاط اشتباه، داده‌های بازگشت آخر فاصله لیدار با استفاده از الگوریتم شیب‌مبنا، فیلتر شدند. در ادامه، داده‌های بازگشت اول و آخر فاصله و شدت لیدار با بهره‌گرفتن از سه روش طبقه‌بندی ماشین‌بردار پشتیبان، شبکه عصبی مصنوعی، و حداکثر احتمال<sup>۱</sup>، به پنج کلاس راه آسفالت، درختان، ساختمان، سیمان، و چمن طبقه‌بندی گردیدند. سپس، نقاط ساختمان‌های بزرگ که به‌وسیله الگوریتم شیب‌مبنا فیلتر نشده بودند به‌وسیله نتایج حاصل از روش‌های طبقه‌بندی و با بهره‌گیری از عملگرهای منطق بولین، حذف گردیدند. در ادامه درباره هر یک از مراحل انجام این تحقیق توضیحاتی ارائه خواهد شد.

## ۲-۲-۱ پیش‌پردازش

در این قسمت با اعمال دو نوع قید روی داده‌ها، نقاط اشتباه، شناسایی و حذف می‌شوند. با توجه به تعریف بازگشت اول و آخر، مقدار تفاضل مؤلفه Z بازگشت اول از بازگشت آخر نباید کمتر از آستانه‌ای نزدیک به صفر باشد. بنابراین، از رابطه (۱) برای تشخیص نقاط اشتباه

استفاده می‌شود.  
(1)  $(\text{First Pulse} - \text{Last Pulse}) > T$

نکته مهم این است که با اعمال این قید فقط نقاطی که مقدار مؤلفه Z بازگشت آخر آن‌ها از مقدار مؤلفه Z بازگشت اول‌شان بیشتر است حذف می‌شوند. بنابراین، اگر در داده‌های بازگشت آخر نقاط با ارتفاع بسیار کم، و یا در داده‌های بازگشت اول نقاط با ارتفاع بسیار زیاد وجود داشته باشد، با اعمال این قید، این نقاط از مجموعه داده‌ها حذف نمی‌شوند. از این رو، ضروری است که از قید دیگری برای حذف این نقاط استفاده شود. برای این منظور، قید محدوده ارتفاعی نقاط به‌کار گرفته می‌شود. با توجه به حداکثر و حداقل ارتفاع منطقه، ارتفاع هیچ نقطه‌ای نباید بیشتر از مقدار ماکزیمم بازگشت اول و یا کمتر از مقدار مینیمم بازگشت آخر باشد. نقاطی که این شرایط را ندارند، به توپوگرافی منطقه تعلق ندارند و باید از داده‌ها حذف شوند.

## ۲-۲-۲ فیلترینگ داده‌های لیدار

از مهم‌ترین پردازش‌هایی که روی داده‌های لیدار صورت می‌پذیرد، فیلترینگ آنهاست. منظور از عملیات فیلترینگ، شناسایی نقاط عوارض غیرزمینی و حذف آنها از مجموعه داده‌هاست. بنابراین حاصل این فرایند، نقاط متعلق به سطح زمین لخت است. اهمیت فرایند فیلترینگ داده‌های لیدار به اندازه‌ای است که از آن به پردازش اولیه داده‌ها تعبیر می‌شود؛ زیرا در اغلب پردازش‌هایی که روی داده‌های لیدار صورت می‌گیرد، از داده‌های فیلترشده به‌عنوان وسیله‌ای برای رسیدن به هدف نهایی استفاده می‌شود.

۱. Maximum Likelihood (MLH)

این روش با این فرض که توزیع آماری هر کلاس در هر باند نرمال است، احتمال تعلق یک پیکسل به کلاس‌ها محاسبه می‌شود و در نهایت، پیکسل به کلاسی تعلق می‌گیرد که بیشترین احتمال را دارد.

### شبکه‌های عصبی مصنوعی

شبکه‌های عصبی مصنوعی از سال ۱۹۸۵ که نخستین شبکه پرسپترون یک لایه را رزنبلات برای طبقه‌بندی الگوها ایجاد کرد، مطالعه شده‌اند. شبکه‌های عصبی مصنوعی مدل ساده‌شده‌ای از مغز انسان هستند. هر شبکه عصبی ساختار ریاضی‌ای غیرخطی است که توانایی نشان دادن فرایندها و ترکیبات دلخواه غیرخطی برای ارتباط بین ورودی‌ها و خروجی‌های هر سیستمی را دارد. این شبکه با داده‌های موجود در طول دوره فرایند یادگیری، آموزش می‌بیند و برای پیش‌بینی در آینده استفاده می‌شود.

### ماشین بردار پشتیبان

در سال ۱۹۹۸ محقق روسی به نام ولادیمیر واپنیک<sup>۱</sup> به جامعه کاربران یادگیری ماشینی<sup>۲</sup> را پیشنهاد کرد. ماشین بردار پشتیبان برای تفکیک کلاس‌ها در داده‌های آموزشی ابررویه‌ای<sup>۳</sup> را روی آن‌ها برآزش می‌دهد. نخستین مرحله برای تعیین ابررویه، پیدا کردن نزدیک‌ترین نمونه‌های آموزشی دو کلاس است. این نمونه‌های آموزشی، بردار پشتیبان<sup>۴</sup> نام دارند. همان‌طور که در شکل ۲ مشاهده می‌شود، برای جداکردن کلاس‌ها ابررویه‌های گوناگونی وجود دارد ولی فقط یکی از آن‌ها بهینه است و آن ابررویه‌ای است که بیشینه حاشیه را بین دو کلاس ایجاد می‌کند. از جمله

الگوریتم فیلترینگ شیب‌مبنا (Vosselman, ۲۰۰۰) از مؤثرترین الگوریتم‌هایی است که تاکنون ارائه شده و پایه بسیاری از الگوریتم‌های دیگر است. در این الگوریتم فرض بر این است که گرادیان در زمین لخت محدود است. این فیلتر بر مبنای مفاهیم مورفولوژی ریاضی استوار است و ساختار هندسی محلی زمین لخت را با استفاده از المان ساختاری تخمین می‌زند. المان ساختاری فرضیه‌ای است درباره بیشترین اختلاف ارتفاع بین هر دو نقطه بر زمین لخت با توجه به فاصله بین آن‌ها. المان ساختاری از نظر مسطحاتی روی یک نقطه مرکزیت پیدا می‌کند و سپس تا زمانی که نقطه را لمس کند، بالا می‌آید. بعد از آن، اگر هیچ نقطه همسایه‌ای زیر المان ساختاری وجود نداشته باشد، نقطه به‌عنوان زمین لخت پذیرفته می‌شود و در غیر این صورت، به‌عنوان عارضه در نظر گرفته می‌شود. اگر  $N$  مجموعه کل داده‌های لیدار و  $BE$  مجموعه نقاط زمینی باشد، تابع همسایگی به‌صورت رابطه (۲) تعریف می‌شود.

رابطه (۲)

$$BE = \begin{cases} 0 & \forall v_i \in N \exists \Delta h(v_i, v_j) > \Delta h_{\max}(d(v_i, v_j)) \\ 1 & \text{else} \end{cases}$$

در اینجا  $\Delta h(v_i, v_j)$  اختلاف ارتفاع بین نقطه  $v_i$  و  $v_j$  که در همسایگی آن قرار دارد و  $\Delta h_{\max}(d(v_i, v_j))$  بیشترین اختلاف ارتفاع مورد انتظار بین دو نقطه از زمین لخت با فاصله مسطحاتی  $d$  است. المان ساختاری، تابع همسایگی برای مقادیر  $d$  است. الگوریتم با به‌کارگیری المان ساختاری در هر نقطه از ابرنقاط کار می‌کند.

### ۲-۲-۳- طبقه‌بندی داده‌های لیدار

#### طبقه‌بندی‌کننده حداکثر احتمال

طبقه‌بندی‌کننده حداکثر احتمال متداول‌ترین روش نظارت‌شده است که در سنجش از دور به کار گرفته شده و یکی از طبقه‌بندی‌کننده‌های آماری است که به توزیع نرمال چندبعدی داده‌ها در هر کلاس بستگی دارد. در

۱. Vladimir Vapnik

۲. Machine Learning

۳. Hyperplane

۴. Support Vector

### ۳-۱- پیش پردازش داده‌ها

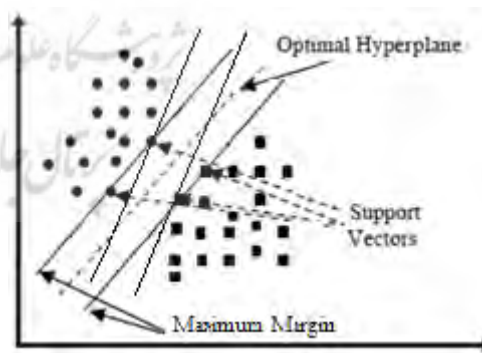
برای حذف نقاط اشتباه، از دو قید مقدار تفاضل مؤلفه Z بازگشت اول از بازگشت آخر و محدوده ارتفاعی نقاط استفاده شده است. با توجه به دقت سیستم‌های لیدار امروزی، انتخاب مقادیر  $T < 10 \text{ cm}$  معیار مناسبی برای حد آستانه (T) نیست. با توجه به دقت داده‌های بررسی شده (15cm)، با در نظر داشتن خطای درون‌یابی فرایند تبدیل این نقاط به داده‌های شبکه‌ای منظم مقدار  $T = 30 \text{ cm}$  (دوبرابر دقت داده‌ها) به عنوان مقدار آستانه انتخاب گردید. با در نظر گرفتن این حد آستانه، تعداد 1742 نقطه به عنوان نقاط اشتباه تشخیص داده شدند. با توجه به مقادیر حداقل و حداکثر ارتفاع پذیرفتنی منطقه که در Header file داده‌ها موجود است، تعداد 59 نقطه دیگر به عنوان نقاط اشتباه تشخیص داده شد. پس از شناسایی و حذف نقاط اشتباه، مقادیر درون‌یابی شده آن‌ها به روش IDW<sup>2</sup> جایگزین گردید. دلیل انتخاب روش درون‌یابی IDW این است که این روش جزء روش‌های دقیق<sup>3</sup> به شمار می‌آید و از این رو با اعمال آن، با تغییرات ناخواسته مقادیر در نقاط معلوم، مواجه نخواهیم شد.

### ۳-۲- اجرای الگوریتم فیلترینگ شیب‌مبنا

از آنجا که هدف از انجام دادن فرایند فیلترینگ، دسترسی به اطلاعات زمین لخت است، داده‌های بازگشت آخر لیدار که حاوی اطلاعات بیشتری از سطح زمین هستند، مناسب‌ترند. بنابراین در این تحقیق، الگوریتم شیب‌مبنا روی داده‌های بازگشت آخر لیدار پیاده‌سازی شده است. الگوریتم فیلترینگ شیب‌مبنا دو پارامتر ورودی شعاع همسایگی و آستانه شیب دارد. در انتخاب شعاع همسایگی باید دقت کرد، زیرا

قابلیت‌های ماشین بردار پشتیبان، توانایی غلبه بر مشکل توزیع غیرخطی داده‌های آموزشی است. در این حالت با استفاده از توابع کرنل، داده‌ها به فضایی با بعد بزرگ‌تر انتقال می‌یابند که در آن تفکیک‌پذیری بهتر انجام می‌شود و ابررویه جداکننده در آن فضا تعیین می‌شود. عمومی‌ترین کرنل‌ها، کرنل‌های خطی، چندجمله‌ای، تانژانت هایپربولیک، و تابع پایه‌شعاعی (RBF)<sup>1</sup> هستند. در کاربردهای سنجش از دور، کرنل RBF نتایج بهتری را ارائه کرده است (Van der Linden et al., 2009). ماشین‌بردار پشتیبان در اصل برای طبقه‌بندی‌های دوکلاسه ارائه شده است، ولی از آن برای طبقه‌بندی چندکلاسه نیز استفاده می‌کنند.

از مزایای ماشین‌بردار پشتیبان، توانایی حل مسائل طبقه‌بندی پیچیده با تعداد لایه‌های زیاد و نمونه‌های آموزشی کم است که آن را به عنوان گزینه‌ای مناسب برای کاربردهای سنجش از دور معرفی کرده است. برای کسب اطلاعات بیشتر در مورد ماشین‌های بردار پشتیبان به منبع بورگس (Burgess, 1998) مراجعه کنید. شکل ۲ عملکرد ماشین بردار پشتیبان را در تعیین ابررویه بهینه و بردارهای پشتیبان برای دو کلاس فرضی نشان می‌دهد.



شکل ۲. ماشین‌بردار پشتیبان در حالت جدایی‌پذیر

### ۳- نتایج

در این قسمت نتایج پیاده‌سازی الگوریتم روی داده‌ها در هر مرحله ارائه می‌گردد.

۱. Radial Basis Function
۲. Inverse Distance Weighting
۳. Exact

### ۳-۳- طبقه‌بندی داده‌های لیدار

با توجه به موضوع این تحقیق که حل مشکل ساختمان‌های بزرگ‌تر الگوریتم‌های فیلترینگ است، داده‌های فاصله و شدت لیدار با استفاده از سه روش طبقه‌بندی ماشین بردار پشتیبان، شبکه عصبی مصنوعی، و حداکثر احتمال، به پنج کلاس راه آسفالت، درختان، ساختمان، چمن، و سیمان طبقه‌بندی گردیدند. برای این منظور تعداد ۶۰۰ نقطه به‌عنوان نمونه‌های آموزشی، به‌طور تصادفی و با توزیع همگن در منطقه انتخاب شدند که از بین آنها تعداد ۱۱۰ نمونه برای کلاس راه آسفالت، ۱۰۰ نمونه برای کلاس درختان، ۱۱۰ نمونه برای کلاس ساختمان، ۱۰۰ نمونه برای کلاس چمن و ۱۸۰ نمونه برای کلاس سیمان در نظر گرفته شد. همچنین تعداد ۵۳۰ نقطه به‌عنوان داده‌های تست انتخاب شدند. برای مقایسه عملکرد سه تکنیک طبقه‌بندی استفاده‌شده در شرایط یکسان، این داده‌های آموزشی و تست برای هر سه روش به کار برده شدند.

به‌منظور اجرای طبقه‌بندی به‌روش حداکثر احتمال، آستانه‌ای برای احتمال در نظر گرفته نشد و بدین ترتیب هر یک از پیکسل‌ها به یکی از کلاس‌ها اختصاص داده شدند و پیکسلی به‌صورت طبقه‌بندی‌نشده باقی نماند. نتیجه طبقه‌بندی داده‌های شدت و فاصله لیدار به‌روش حداکثر احتمال در شکل ۳-ب نشان داده شده است. ماتریس خطای حاصل از طبقه‌بندی به‌روش حداکثر احتمال در جدول ۱ آمده است.

همسایگی‌های کوچک منجر می‌شود که نقاط موجود روی ساختمان‌های بزرگ به‌عنوان نقاط زمینی در نظر گرفته شوند و همسایگی‌های بزرگ نیز باعث نرم‌شدن بیش از حد زمین می‌شود. از نظر تئوری، شعاع همسایگی باید اندکی از بزرگ‌ترین ساختمان موجود در منطقه بزرگ‌تر باشد تا هنگامی که همسایگی روی آن ساختمان قرار می‌گیرد، نقاطی از زمین نیز در همسایگی واقع شوند. در عمل چون ممکن است ابعاد ساختمان‌ها ناهمگن باشند، انتخاب شعاع همسایگی براساس ابعاد بزرگ‌ترین ساختمان اشتباه است و موجب از دست رفتن اطلاعات زیادی از سطح زمین در سایر نواحی می‌شود. در این تحقیق با در نظر گرفتن خصوصیات منطقه، شعاع همسایگی ۲۵ متر در نظر گرفته شده است. انتخاب آستانه شیب مناسب نیز مهم است، چرا که مقادیر کم این پارامتر سبب قرار گرفتن نقاط زمین در کلاس عوارض و حذف آن‌ها می‌شود، و مقادیر بزرگ آن باعث می‌شود تعداد زیادی از نقاط عوارض به اشتباه در کلاس زمین قرار گیرند. به‌منظور انتخاب آستانه شیب بهینه در این تحقیق، فرایند فیلترسازی داده‌ها با ۵ آستانه شیب  $20^{\circ}$ ،  $24^{\circ}$ ،  $27^{\circ}$ ،  $30^{\circ}$  و  $33^{\circ}$  انجام شد. نتیجه فیلترینگ داده‌ها با استفاده از الگوریتم شیب‌مبنا در شکل ۳-الف نشان داده شده است. همان‌طور که در این شکل مشاهده می‌شود، نقاط زیادی روی سقف ساختمان‌های بزرگ به‌اشتباه در کلاس زمین قرار گرفته و فیلتر نشده‌اند. الگوریتم فیلترینگ شیب‌مبنا در محیط نرم‌افزار MATLAB برنامه‌نویسی شده است.

جدول ۱. ماتریس خطای طبقه‌بندی به روش حداکثر احتمال

		داده‌های مرجع					
		آسفالت	درخت	ساختمان	چمن	سیمان	مجموع
روش طبقه‌بندی	آسفالت	۸۹	۰	۰	۰	۱۶	۱۰۵
	درخت	۲	۸۹	۳	۰	۵	۹۹
	ساختمان	۱	۱	۹۷	۰	۲	۱۰۱

چمن	۰	۰	۰	۸۹	۰	۸۹
سیمان	۸	۰	۰	۱	۱۲۷	۱۳۶
مجموع	۱۰۰	۹۰	۱۰۰	۹۰	۱۵۰	۵۳۰
ضریب کاپا: ۰/۹۰۷۳			دقت کلی: ۹۰/۶۴۱۵ (درصد)			

همه داده‌های استفاده‌شده و جلوگیری از غلبه کردن داده‌های با دامنه بزرگ‌تر بر داده‌های با دامنه کوچک‌تر است. برای انجام طبقه‌بندی چندکلاسه در این تحقیق از استراتژی یکی در مقابل یکی استفاده شده است. در این تحقیق از کرنل RBF استفاده شده است. به‌منظور تعیین مقادیر بهینه پارامتر عرض کرنل RBF (g) و پارامتر جریمه (C)، روش جست‌وجوی شبکه‌ای<sup>۴</sup> با استفاده از ۵ بار ارزیابی<sup>۵</sup> به کار گرفته شده است. با اجرای روش جست‌وجوی شبکه‌ای، مقادیر  $g=1/0.8$  و  $C=1311/0.0$  به‌عنوان مقادیر بهینه نهایی به‌دست آمدند. نتیجه طبقه‌بندی داده‌های شدت و فاصله لیدار به‌روش شبکه عصبی در شکل ۳-د نشان داده شده است. ماتریس خطای حاصل از طبقه‌بندی به‌روش شبکه عصبی در جدول ۳ آمده است.

مهم‌ترین موضوع در اجرای طبقه‌بندی شبکه عصبی، طراحی ساختار شبکه است. در این تحقیق با انجام سعی و خطا بهترین ساختار شبکه بر پایه معیارهای دقت کلی و ضریب کاپا مشخص شد. بر این اساس، شبکه‌ای با یک لایه پنهان و تعداد تکرار<sup>۱</sup> ۳۰۰۰ و نرخ یادگیری<sup>۲</sup> ۰/۳ به‌عنوان بهترین شبکه در نظر گرفته شد. نتیجه طبقه‌بندی داده‌های شدت و فاصله لیدار به روش شبکه عصبی در شکل ۳-ج نشان داده شده است. ماتریس خطای حاصل از طبقه‌بندی به روش شبکه عصبی در جدول ۲ آمده است. برای انجام طبقه‌بندی به‌روش ماشین بردار پشتیبان، ابتدا داده‌های شدت و فاصله لیدار به دامنه (۰/۱) تغییر مقیاس داده شدند. هدف از این کار، یکسان شدن دامنه

جدول ۲. ماتریس خطای طبقه‌بندی به‌روش شبکه عصبی

		داده‌های مرجع					
		آسفالت	درخت	ساختمان	چمن	سیمان	مجموع
خروجی طبقه‌بندی	آسفالت	۸۸	۰	۰	۰	۱۵	۱۰۳
	درخت	۱	۸۸	۱	۰	۰	۹۰
	ساختمان	۰	۲	۹۹	۰	۰	۱۰۱
	چمن	۰	۰	۰	۹۰	۰	۹۰
	سیمان	۱۱	۰	۰	۰	۱۳۵	۱۳۶
	مجموع	۱۰۰	۹۰	۱۰۰	۹۰	۱۵۰	۵۳۰
ضریب کاپا: ۰/۹۲۸۵			دقت کلی: ۹۴/۳۳۹۶ (درصد)				

۱. Iteration
۲. Learning Rate
۳. Penalty parameter
۴. Grid-search
۵. ۵-fold cross-validation



مقایسه توانایی طبقه‌بندی‌کننده‌های ماشین بردار پشتیبان، شبکه عصبی و حداکثر احتمال در بهبود ...

جدول ۳. ماتریس خطای طبقه‌بندی به‌روش ماشین بردار پشتیبان

		داده‌های مرجع					
		آسفالت	درخت	ساختمان	چمن	سیمان	مجموع
خروجی طبقه‌بندی	آسفالت	۹۱	۰	۰	۰	۱۵	۱۰۶
	درخت	۰	۸۷	۱	۰	۰	۸۸
	ساختمان	۰	۳	۹۹	۰	۰	۱۰۲
	چمن	۰	۰	۰	۹۰	۰	۹۰
	سیمان	۹	۰	۰	۰	۱۳۵	۱۴۴
	مجموع	۱۰۰	۹۰	۱۰۰	۹۰	۱۵۰	۵۳۰

ضریب کاپا: ۰/۹۳۳۳

دقت کلی: ۹۴/۷۱۷۰ (درصد)

فیلترینگ در نظر گرفته شدند. داده‌های فیلترشده (آستانه شیب  $20^{\circ}$ ) پس از استفاده از نتایج طبقه‌بندی به روش‌های حداکثر احتمال شبکه عصبی مصنوعی، و ماشین بردار پشتیبان، به ترتیب در شکل‌های ۳-هـ، ۳-و، و ۳-ز آمده‌اند.

پس از گرفتن خروجی الگوریتم شیب‌مبنا و خروجی تکنیک‌های طبقه‌بندی، با استفاده از عملگرهای منطقه بولین، نقاطی که در داده‌های فیلترشده به‌وسیله الگوریتم شیب‌مبنا در کلاس زمین لخت قرار دارند و در تصاویر طبقه‌بندی‌شده در کلاس ساختمان قرار ندارند، به‌عنوان خروجی نهایی فرایند



(ج) طبقه‌بندی به روش ANN



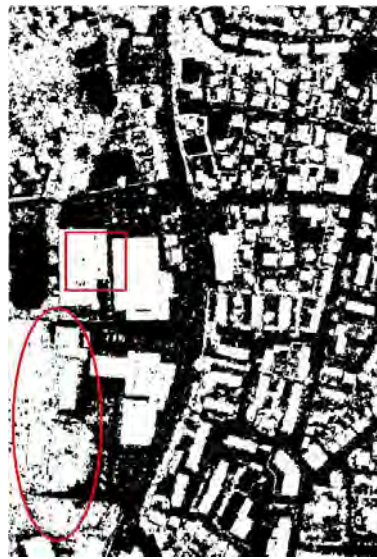
(ب) طبقه‌بندی به روش MLH



(الف) خروجی الگوریتم شیب‌مبنا



(و) خروجی نهایی پس از بهبود نتایج به وسیله ANN



(ه) خروجی نهایی پس از بهبود نتایج به وسیله MLH



(د) طبقه‌بندی به روش SVM



(ح) داده‌های مرجع



(ز) خروجی نهایی پس از بهبود نتایج به وسیله SVM

شکل ۳. نتایج فیلترینگ داده‌ها

### ۳-۴- ارزیابی دقت

به منظور انجام ارزیابی دقت، ابتدا داده‌های مرجع به صورت دستی و با بهره‌گیری از داده‌های شدت و فاصله لیدار و نیز تصویر هوایی رنگی با قدرت تفکیک بالای منطقه تولید شدند. شکل ۳-ح داده‌های مرجع را نشان می‌دهد.

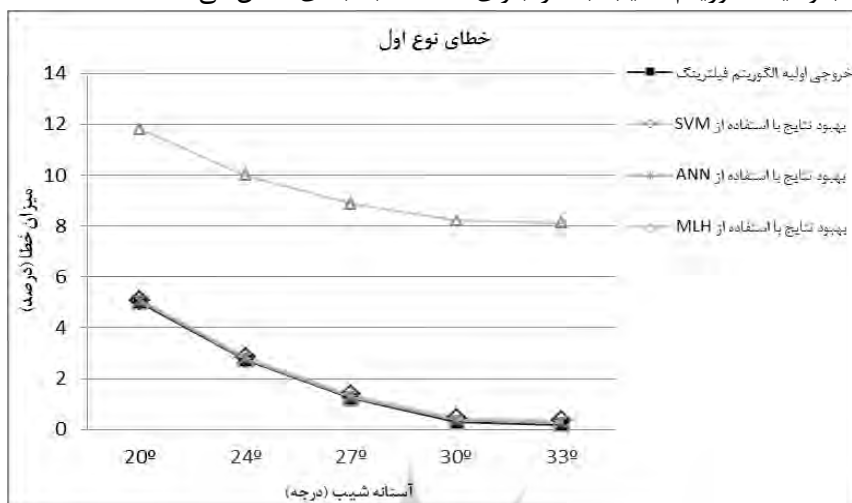
ارزیابی دقت براساس مقایسه نتایج حاصل از الگوریتم پیشنهادی با داده‌های مرجع صورت پذیرفت. برای این منظور میزان خطاهای نوع اول و نوع دوم و

همچنین خطای مجموع محاسبه شدند. خطای نوع اول عبارت است از نسبت تعداد نقاط زمینی که به اشتباه به عنوان زمین طبقه‌بندی شده‌اند به کل تعداد نقاط زمینی. خطای نوع دوم عبارت است از نسبت تعداد نقاط عارضه که به اشتباه به عنوان زمین طبقه‌بندی شده‌اند به تعداد نقاط عارضه. خطای مجموع، خطای کلی است که هر دو خطای نوع اول و دوم را در نظر دارد و عبارت است از نسبت تعداد کل طبقه‌بندی‌های اشتباه به تعداد کل نقاط.

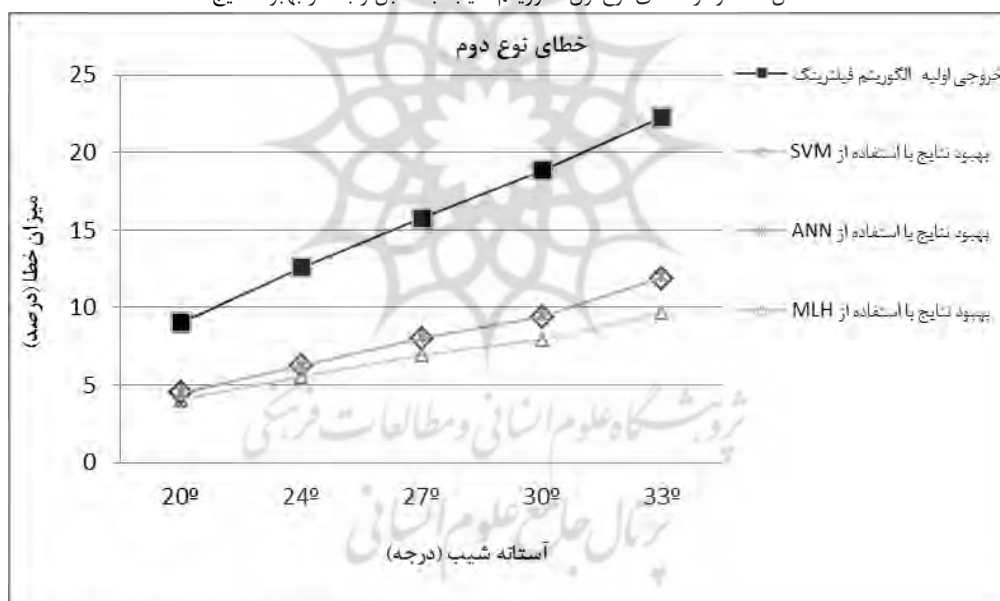
مقایسه توانایی طبقه‌بندی‌کننده‌های ماشین بردار پشتیبان، شبکه عصبی و حداکثر احتمال در بهبود ...

آستانه شیب‌های  $20^\circ$ ،  $24^\circ$ ،  $27^\circ$ ،  $30^\circ$  و  $33^\circ$  در حالت‌های قبل و بعد از بهبود نتایج با تکنیک‌های طبقه‌بندی، نشان می‌دهند.

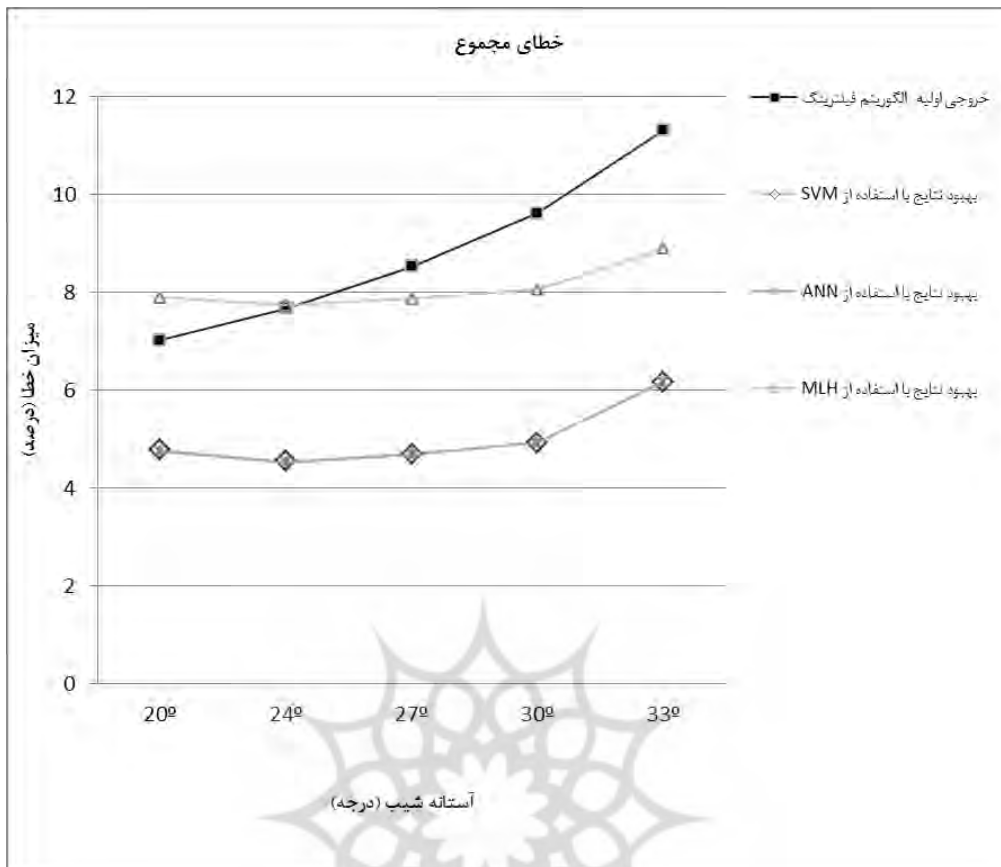
شکل‌های ۴ و ۵ به ترتیب میزان خطای نوع اول، خطای نوع دوم، و خطای مجموع برای خروجی فرایند فیلترینگ به وسیله الگوریتم شیب‌مبنا را برای



شکل ۴. نمودار خطای نوع اول الگوریتم شیب‌مبنا، قبل و بعد از بهبود نتایج



شکل ۵. نمودار خطای نوع دوم الگوریتم شیب‌مبنا، قبل و بعد از بهبود نتایج



شکل ۶. نمودار خطای مجموع الگوریتم شیب‌مبنا، قبل و بعد از بهبود نتایج

#### ۴- تحلیل نتایج

برطبق نتایج ارائه شده در نمودار شکل ۶، خروجی الگوریتم شیب‌مبنا با آستانه شیب ۲۴° که نتایج آن با طبقه‌بندی شبکه عصبی بهبود داده شده، دارای کمترین خطای مجموع معادل با ۴/۵۳ درصد است. البته، باید در نظر داشت که برای تعیین بهترین خروجی، علاوه بر پارامتر خطای مجموع، لازم است تا به توازن بین خطای نوع اول و خطای نوع دوم نیز توجه شود. مقدار خطای نوع اول و خطای نوع دوم، برای آستانه شیب ۲۴° پس از بهبود با طبقه‌بندی شبکه عصبی، به ترتیب ۲/۷۹ و ۶/۲۵ درصد است. همان‌طور که ملاحظه می‌شود اختلاف زیادی بین این دو مقدار وجود دارد. حال آنکه مقدار خطای نوع اول و نوع دوم، برای آستانه شیب ۲۰° پس از بهبود به وسیله طبقه‌بندی شبکه عصبی، به ترتیب ۵/۰۴ و ۴/۴۹ درصد

است که توازن بسیار بهتری در مقایسه با آستانه شیب ۲۴° دارد. علاوه بر این، در آستانه شیب ۲۰° خطای نوع دوم از خطای نوع اول کمتر شده است. با توجه به اهمیت بیشتر کم کردن خطای نوع دوم، می‌توان ادعا کرد که بهترین خروجی، خروجی الگوریتم شیب‌مبنا با آستانه شیب ۲۰° پس از بهبود با طبقه‌بندی شبکه عصبی است، زیرا علاوه بر این که خطای نوع دوم و خطای مجموع در آن کم است، بهترین توازن بین خطای نوع اول و خطای نوع دوم را نیز دارد.

کسب کمترین خطای مجموع در آستانه شیب ۲۴° نشان‌دهنده مناسب بودن بازه انتخاب شده برای آنالیز حساسیت پارامتر آستانه شیب است، زیرا با انحراف از این مقدار آستانه شیب - چه به سمت مقادیر کمتر و چه به سمت مقادیر بیشتر - خطای مجموع افزایش خواهد یافت.

نمودار شکل ۴ مشاهده می‌شود، پس از استفاده از نتایج طبقه‌بندی با روش‌های شبکه عصبی و ماشین بردار پشتیبان، خطای نوع اول به مقدار بسیار ناچیزی افزایش یافته است. در عوض بر طبق نمودارهای شکل‌های ۵ و ۶، خطای نوع دوم و خطای مجموع به میزان شایان توجهی کم شده‌اند. از آنجا که تأثیر منفی خطای نوع دوم در روش‌های تولید DTM از داده‌های فیلتر شده بسیار بیشتر از خطای نوع اول است، می‌توان گفت که کاهش خطای نوع دوم به بهای افزایش خطای نوع اول، حتی در حالت برابر هم ارزشمند است. این امر مؤید تأثیر بسیار مطلوب استفاده از روش‌های طبقه‌بندی شبکه عصبی و ماشین بردار پشتیبان بر نتیجه فرایند فیلترینگ است که در ازای افزایشی بسیار اندک در خطای نوع اول، کاهش شدیدی در خطای نوع دوم و نیز خطای مجموع اتفاق می‌افتد.

#### ۵- نتیجه‌گیری

در این تحقیق سعی شده است با استفاده از روش‌های طبقه‌بندی، نتایج الگوریتم‌های فیلترینگ بهبود داده شود. در بسیاری از الگوریتم‌های فیلترینگ از جمله الگوریتم شیب‌منا، فقط اطلاعات هندسی موجود در داده‌های فاصله‌لیدار استفاده می‌شوند و به اطلاعات شبه‌طیفی داده‌های شدت توجهی نمی‌شود. از آنجا که عملکرد این الگوریتم به صورت محلی و با در نظر گرفتن شعاع همسایگی است، مشکلاتی نظیر قرار گرفتن ساختمان‌های بزرگ در کلاس زمین لخت وجود دارد، حال آن‌که در روش‌های طبقه‌بندی، از هر دو اطلاعات هندسی و شبه‌طیفی لیدار می‌توان بهره برد و به دلیل عملکرد سراسری این روش، مشکل عوارض بزرگ مطرح نیست. این روش‌ها نیز، معایبی مانند عملکرد ضعیف در مواجهه با تشابهات طیفی (مثلاً ساختمان‌ها و جاده‌ها در داده‌های شدت و ساختمان‌ها و درختان در داده‌های فاصله و از این دست) دارند. در روش ارائه شده در این تحقیق با ترکیب خروجی‌های این دو روش، معایب هر دو برطرف شده و نتایج به دست آمده بسیار

با توجه به مقادیر ضریب کاپا و دقت کلی ارائه شده در جدول‌های ۱ تا ۳ می‌توان گفت که عملکرد روش‌های شبکه عصبی و ماشین بردار پشتیبان در طبقه‌بندی داده‌های لیدار بهتر از روش حداکثر احتمال بوده است. این ضعف عملکرد طبقه‌بندی‌کننده حداکثر احتمال، به فرایند استفاده از آن در بهبود نتایج الگوریتم فیلترینگ نیز منتقل و موجب شده است تا نتایج ضعیفی در این زمینه به دست آید. همان‌طور که در نمودار شکل ۴ مشاهده می‌شود، استفاده از طبقه‌بندی حداکثر احتمال، خطای نوع اول را در حد نامعقولی افزایش داده است. دلیل این امر حذف تعداد زیادی از نقاط زمینی است که به وسیله الگوریتم فیلترینگ شیب‌منا به درستی در کلاس زمین قرار گرفته‌اند اما به علت طبقه‌بندی اشتباه این نقاط با روش حداکثر احتمال، در کلاس ساختمان قرار گرفته و فیلتر شده‌اند. حذف این نقاط به وضوح در بیضی مشخص شده در شکل ۳- ه مشاهده می‌شود. اشکال دیگر استفاده از روش طبقه‌بندی حداکثر احتمال در بهبود نتایج الگوریتم فیلترینگ، طبقه‌بندی اشتباه پیکسل‌های کلاس ساختمان در کلاس درختان است. با این‌که استفاده از این روش کمترین خطای نوع دوم را ارائه کرده است (شکل ۵)، اما به دلیل محل قرارگیری این خطاها، نتایج حاصل پذیرفتنی نیست. همان‌طور که در مربع مشخص شده در شکل ۳- ه مشاهده می‌شود در داده‌های فیلتر شده نهایی، نقاطی روی ساختمان‌های بزرگ وجود دارند که بعد از بهبود با نتایج طبقه‌بندی حداکثر احتمال نیز حذف نشده‌اند. نمودار خطای مجموع ارائه شده در شکل ۶ نیز حاکی از نتایج ضعیف استفاده از طبقه‌بندی‌کننده حداکثر احتمال است.

برخلاف روش حداکثر احتمال، روش‌های شبکه عصبی و ماشین بردار پشتیبان، نتایج بسیار راضی‌کننده‌ای را ارائه کرده‌اند. نتایج حاصل از به کارگیری این دو روش بسیار به هم نزدیک است ولی نتایج روش شبکه عصبی به میزان بسیار اندکی بهتر از روش ماشین بردار پشتیبان است. همان‌طور که در

- Models, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXIII (Pt. B4/1), PP. 110-117.**
- Baligh, J.A., Valadan Zoej, M.J., Mohammadzadeh, M., Sadeghian, S., ۲۰۱۱, **A Novel Filtering Algorithm for Bare-earth Extraction from Airborne Laser Scanning Data Using an Artificial Neural Network**, IEEE Journal of Selected Topics In Applied Earth Observations And Remote Sensing, Vol. ۴, No. ۴, PP. ۸۳۶-۸۴۳.
- Brovelli, M., Cannata, M., Longoni, U., ۲۰۰۲, **Managing and Processing Lidar Data Within GRASS, Proceedings of the Open Source GIS-GRASS Users conference**, ۱۱-۱۳ September, Trento, Italy, ۲۹ pages.
- Burges, C. J. C., ۱۹۹۸, **A Tutorial on Support Vector Machines for Pattern Recognition**, Data Mining and knowledge Discovery, Vol. ۲, PP. ۱۲۱-۱۶۷.
- Chen, Q., Gong, P., Baldocchi, D.D., Xie, G., ۲۰۰۷, **Filtering Airborne Laser Scanning Data With Morphological Methods**, Photogramm, Eng, Remote Sens, Vol. ۷۳, PP. ۱۷۵-۱۸۵.
- Elmqvist, M., ۲۰۰۱, **Ground Estimation of Laser Radar Data Using Active Shape Models**, Proceedings of OEEPE workshop on Airborne Laser Scanning and
- راضی کننده است. در این راستا عملکرد سه روش طبقه‌بندی ماشین بردار پشتیبان، شبکه عصبی مصنوعی، و حداکثر احتمال با هم مقایسه شدند، که نتایج نشان‌دهنده کارایی بسیار خوب روش‌های شبکه عصبی مصنوعی و ماشین‌بردار پشتیبان است. برطبق نتایج این تحقیق، روش حداکثر احتمال برای طبقه‌بندی داده‌های لیدار به‌منظور بهبود نتایج الگوریتم‌های فیلترینگ مناسب شناخته نشد.
- ۶- منابع
- Arefi, H., Hahn, H., ۲۰۰۵, **A Morphological Reconstruction Algorithm for Separating Off-Terrain Points from Terrain Points in Laser Scanning Data**, in Proc. ISPRS Workshop Laser Scanning, Enschede, The Netherlands, Sep.
- Axelsson, P., ۱۹۹۹, **Processing of Laser Scanner Data - Algorithms and Applications**, ISPRS/JPRS, Vol. ۵۴, PP. ۱۳۸-۱۴۷.
- Axelsson, P., ۲۰۰۰, **DEM Generation From Laser Scanner Data Using Adaptive TIN** interferometric SAR for Detailed Digital Elevation Models, March ۱-۳, Stockholm, Sweden, Official Publication No. ۴۰. CD-ROM., ۸ pages.
- Elmqvist, M., ۲۰۰۲, **Ground Surface Estimation from Airborne Laser Scanner Data Using Active Shape Models**, In Proc. ISPRS Commission III Symposium, Photogrammetric and Computer Vision,



- Graz, Austria, PP. ۱۱۴-۱۱۸.
- Elmqvist, M., Jungert, E., Lantz, F., Persson, A., Soderman, U., ۲۰۰۱, **Terrain Modelling and Analysis Using Laser Scanner Data**, In International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences XXXIV (Pt. ۳/W۴), PP. ۲۱۹-۲۲۷.
- Filin, S., ۲۰۰۲, **Surface Clustering from Airborne Laser Scanning Data**, Int.Arch. Photogramm, Remote Sens. Spatial Inf, Sci. XXXIV, PP. ۱۱۹-۱۲۴.
- Jacobsen, K., Lohmann, P., ۲۰۰۳, **Segmented Filtering of Laser Scanner DSMs**, In Proc. ISPRS Working Group III/۳ Workshop, ۳-D Reconstruction From Airborne Laserscanner and InSAR Data, Dresden, Germany.
- Kobler, A., Pfeifer, N., Ogrinc, P., Todorovski, L., Ostir, K., Dzeroski, S., ۲۰۰۷, **Repetitive Interpolation: A Robust Algorithm for DTM Generation from Aerial Laser Scanner Data in Forested Terrain**, **Remote Sens. Environ**, vol. ۱۰۸, PP. ۹-۲۳.
- Meng, X., ۲۰۰۵, **A Slope-and Elevation-based Filter to Remove Non-ground Measurements from Airborne LIDAR Data**, In Proc. ISPRS WG III/۳, III/۴, V/۳ Workshop Laser scanning ۲۰۰۵, The Netherlands, PP. ۲۳-۲۳.
- Meng, X., Wang, L., Currit, N., ۲۰۰۹, **Morphology-Based Building Detection From Airborne LIDAR Data**, Photogramm, Eng, Remote Sens., vol. ۷۵, PP. ۴۲۷-۴۴۲.
- Meng, X., Wang, L., Silvan-Cardenas, J.L., Currit, N., ۲۰۰۹, **A Multidirectional Ground Filtering Algorithm for airborne LIDAR**, J. Photogrammetry and Remote Sensing (ISPRS), Vol. ۶۴, No. ۱, PP. ۱۱۷-۱۲۴.
- Rabbania, T., van den Heuvelb, F.A., Vosselman, G., ۲۰۰۶, **Segmentation of Point Clouds, Using Smoothness Constraint**, In Int, Arch, Photogramm, Remote Sens, Spatial Inf, Sci, Vol. ۳۶, PP. ۲۴۸-۲۵۳.
- Roggero, M., ۲۰۰۱, **Airborne Laser Scanning: Clustering in Raw Data**, IAPRS, Vol. ۳۴ (۳W/۴, WG IV/۳), October ۲۲-۲۴, Annapolis (MD), USA, PP. ۲۲۷-۲۳۲.
- Salah, M., Trinder, J., ۲۰۱۰, **Support Vector Machines Based Filtering of Lidar Data: A Grid Based Method**, FIG Congress ۲۰۱۰, Sydney, Australia, ۱۱-۱۶ April.
- Shan, J., Sampath, A., ۲۰۰۵, **Urban DEM Generation from Raw LIDAR data: A Labeling Algorithm and its Performance**,

- Photogramm, Eng, Remote Sens, Vol. ۷۱,  
No. ۲, PP. ۲۱۷-۲۲۶.
- Sithole, G., ۲۰۰۵, **Segmentation and Classification of Airborne Laser Scanner Data**, University of Delft, Ph.D. thesis.
- Sithole, G., Vosselman, G., ۲۰۰۵, **Filtering of Airborne Laser Scanner Data Based on Segmented Point Clouds**, In Proc. ISPRS Workshop Laser Scanning ۲۰۰۵, Enschede, The Netherlands, PP. ۶۶-۷۱.
- Sithole, G., Vosselman, G., ۲۰۰۳, **Comparison of Filtering Algorithms**, in IAPRS ۳۴(۳/W۱۳), PP. ۷۱-۷۸.
- Tovari, D., Pfeifer, N., ۲۰۰۵, **Segmentation Based Robust Interpolation-A New Approach to Laser Filtering**, In Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci, PP. ۷۹-۸۴.
- Van der Linden, S., Rabe, A., Okujeni, A., Hostert, P., ۲۰۰۹, **Image SVM Classification**, Application Manual: image SVM version ۲,۰, Humboldt-Universität zu Berlin, Germany.
- Vosselman, G., ۲۰۰۰, **Slope Based Filtering of Laser Altimetry Data**, IAPRS, WG III/۳, Amsterdam, The Netherlands, Vol. ۳۳ (B۳), PP. ۹۳۵-۹۴۲.
- Vosselman, G. Maas, H., ۲۰۰۱, **Adjustment and Filtering of Raw Laser Altimetry Data**, Proceedings of OEEPE workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Elevation Models, paper ۵ (۱۱ pages), Royal Institute of Technology Department of Geodesy and Photogrammetry ۱۰۰ Stockholm, Sweden.
- Zhang, K., Chen, S., Whitman, D., Shyu, M., ۲۰۰۳, **A Progressive Morphological Filter for Removing Nonground Measurements from Airborne LiDAR Data**, IEEE Trans, Geosci, Remote Sens, Vol. ۴۱, PP. ۸۷۲-۸۸۲.