

روشی تلفیقی برای شناسایی ساختمان از تصویر SAR با قدرت تفکیک بالا در مناطق شهری

مریم تیموری^{۱*}، مهدی مختارزاده^۲، محمدجواد ولدان زوج^۳

^۱کارشناس ارشد سنجش از دور - دانشکده مهندسی نقشه برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی
mteimouri@mail.kntu.ac.ir

^۲دانشیار دانشکده مهندسی نقشه برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی
m_mokhtarzade@kntu.ac.ir

^۳استاد دانشکده مهندسی نقشه برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی
valadanzouj@kntu.ac.ir

(تاریخ دریافت فروردین ۱۳۹۵، تاریخ تصویب شهریور ۱۳۹۵)

چکیده

به منظور مدیریت هرچه بهتر مناطق شهری و نیمه شهری وجود اطلاعات مکانی دقیق و به هنگام، از عوارض و پدیده های طبیعی و مصنوعی مورد نیاز می باشد. در این راستا استخراج اطلاعات مکانی ساختمان ها، به عنوان یکی از عوارض غالب بر محیط شهری مورد توجه محققان سنجش از دور و فتوگرامتری می باشد. به همین منظور، دستیابی به الگوریتمی مناسب جهت شناسایی و استخراج ساختمان ها از تصاویر هوایی و ماهواره ای بسیار مورد اهمیت است. از این رو در این مقاله تلاش شده است، روشی جهت بهبود شناسایی ساختمان از تصویر شدت TerraSAR_X ارائه شود. در روش پیشنهادی، با به کارگیری تلفیق طبقه بندی کننده ها و بهره گیری از اطلاعات همسایگی هر پیکسل، باعث بهبود نتایج شناسایی ساختمان و کاهش نویز در این تصاویر شده است. بنابراین در این روش، ابتدا به استخراج و انتخاب ویژگی بهینه از تصویر TerraSAR_X پرداخته شده است. سپس به مقایسه و ارزیابی توانایی طبقه بندی کننده های مختلف که شامل شبکه عصبی، ماشین بردار پشتیبان، بیشترین شباهت و نزدیک ترین همسایگی در شناسایی ساختمان پرداخته شده است. سپس به منظور بهبود نتایج، اقدام به تلفیق نتایج در سطح تصمیم گیری توسط پنجره ای متحرک شده است. نتایج حاصل از روش پیشنهادی با دقت کلی و دقت شناسایی ساختمان به ترتیب ۸۶،۴۱٪ و ۷۳،۰۸٪ مؤید توانایی این روش می باشد. همچنین استراتژی تلفیق پیشنهادی منجر به حداقل ۵٪ بهتر شدن نتایج دقت شناسایی ساختمان ها نسبت به استفاده مجزا از هر طبقه بندی کننده شده است.

واژگان کلیدی: ساختمان، SAR، تلفیق، شبکه عصبی، بیشترین شباهت، ماشین بردار پشتیبان، نزدیک ترین همسایگی

۱- مقدمه

در چند دهه اخیر رویکرد اتوماتیک‌سازی استخراج ساختمان از تصاویر هوایی و ماهواره‌ای موضوع بسیاری از تحقیقات بوده است. برخی از الگوریتم‌ها در این زمینه جهت شناسایی و استخراج ساختمان برای منابع مختلف داده توسعه یافته است. اما وجود عواملی از قبیل سایه، انسداد، ناهمگونی سقف ساختمان‌ها و شباهت‌های طیفی عوارض شهری، شناسایی و استخراج ساختمان را نسبتاً پیچیده می‌کند. در این راستا محققان تکنیک‌هایی برای مقابله با مشکلات استخراج و بازسازی ساختمان و ارائه نتایج مؤثرتر با داده‌های متفاوت پیشنهاد داده‌اند.

با ظهور نسل جدیدی از ماهواره‌های راداری از قبیل CosmoSkyMed, TerraSAR-X و غیره با قدرت تفکیک مکانی بالا و همچنین برداشت متفاوت اطلاعات زمینی (خواص فیزیکی اشیاء) توسط این ماهواره‌ها و امکان تصویربرداری در طول شبانه‌روز و در تمامی شرایط آب‌وهوایی نقش بسزایی در زمینه‌ی استخراج عارضه ساختمان ایفا می‌نمایند. الگوریتم‌هایی نیز با استفاده از این تصاویر به شناسایی و استخراج ساختمان پرداخته‌اند. بنابراین در این بخش به بررسی اجمالی برخی از مطالعات انجام‌شده در این زمینه پرداخته شده است.

Gouinaud و همکارش در سال ۱۹۹۶، با استفاده از تصاویر ERS-1 با قدرت تفکیک ۳۰ متر بر مبنای روش آماری fmax-filter به شناسایی مناطق شهری پرداختند. آن‌ها در این روش به بررسی قانون باز پراکنش از مناطق شهری به عنوان نتایج حاصل از اثرات ترکیبی از سه سطح زبر، صاف، سطوح هادی پرداختند. در سال ۱۹۹۸، Tupin و همکاران با استفاده از یک آشکارساز ویژگی‌هایی از قبیل خطوط روشن از تصاویر SAR که ردپاهایی از بخش‌های ساختمان بودند را استخراج کردند [۱]. همچنین در آن سال، Chellappa به منظور شناسایی خطوط روشن ساختمان‌ها از الگوریتم نرخ هشدار کاذب ثابت (CFAR) و تبدیل هاف بهره برد. سپس بخش‌بندی بیش‌ترین احتمال را برای شناسایی سایه ساختمان‌ها بکار برد و در مجموع این اطلاعات را برای شناسایی ساختمان استفاده کرد [۲]. Dell'Acqua و همکارش در سال ۲۰۰۳، روشی با استفاده از ویژگی‌های بافت و طبقه‌بندی

کننده شبکه عصبی- فازی بر تصاویر ERS-1 به منظور تفکیک به سه زیرمجموعه مرکز شهر، مناطق مسکونی و مناطق حومه ارائه دادند [۳]. Stasolla و همکارش در سال ۲۰۰۸، با کمک ویژگی‌های محلی همبستگی مکانی و روش‌های مورفولوژیکی به تفکیک مناطق ساخته شده از پس‌زمینه پرداختند. ضعف اصلی روش پیشنهادی وابستگی الگوریتم ارائه شده به انتخاب حد آستانه و نیاز روش پیشنهادی به داشتن منطقه‌ای با پراکندگی‌های قوی بود [۴]. Wang و همکاران، در سال ۲۰۰۸، روشی بر مبنای اطلاعات ناحیه مینا و لبه مینا با استفاده از داده پلاریمتری با حد تفکیک بالا ارائه دادند [۵].

He و همکارش در سال ۲۰۰۹ با استفاده از داده هوایی پلاریمتری SAR با قدرت تفکیک ۱,۴۹ متر در رنج و ۰,۷۴۸ متر در آزیموت به مدل‌سازی شبکه بیزین از تعاملات شواهد سطح، ساختمان و زاویه به شناسایی ساختمان پرداختند [۶]. در سال ۲۰۱۰، Matsuoka و همکارش با معرفی یک تابع احتمال به تعیین میزان آسیب وارده به ساختمان با استفاده از تصاویر JERS-1 پرداختند [۷]. در این راستا آن‌ها با تجزیه و تحلیل تفکیک رگرسیون با استفاده از ضریب همبستگی و مقادیر اختلاف از تصاویر قبل و بعد از زلزله، این تابع را تعریف کردند. آن‌ها نشان دادند که میزان آسیب وارده بر ساختمان می‌تواند از تصاویر JERS-1 از طریق ادغام با تابع fragility برای آسیب وارده از شدت زمین‌لرزه برآورد شود. Wegner و همکاران در سال ۲۰۱۱ به ارائه روشی با استفاده از ترکیب ارتوفتو (قدرت تفکیک ۰,۳۱ متر) و داده InSAR (قدرت تفکیک ۰,۳۸ متر در رنج و ۰,۱۸ متر در آزیموت) با یک چارچوب میدان تصادفی شرطی^۲ پرداختند [۸]. در این تحقیق از ویژگی‌هایی از قبیل خطوط گوشه ساختمان، توزیع فاز InSAR در امتداد خطوط گوشه‌ای استفاده کردند. Dong و همکاران در سال ۲۰۱۱، جهت استخراج ویژگی‌های بافتی از دو توصیفگر بافت، جذر اختلاف جفت ریشه (SPRD) و *Gi استفاده کردند. سپس از طریق این دو توصیفگر به استخراج ساختمان‌ها در مناطق شهری با استفاده از تصاویر TerraSAR-X و COSMO/SkyMed اقدام کردند [۹]. در سال ۲۰۱۳، Zhao و همکارانش با استفاده از

^۲ conditional random field (CRF)

^۳ Square Root Pair Difference

^۴ Getis-Ord

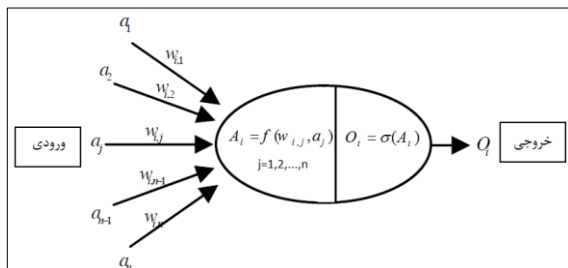
^۱ Constant False Alarm Rate

پایه‌سازی و ارزیابی نتایج و در نهایت در بخش پنجم نتیجه‌گیری ارائه شده است.

۲- مبانی تئوری تحقیق

۲-۱- شبکه عصبی مصنوعی (NNA)

شبکه‌های عصبی مصنوعی الگویی برای پردازش اطلاعات می‌باشند که با تقلید از شبکه‌های عصبی بیولوژیکی مثل مغز انسان ساخته شده‌اند. یک شبکه عصبی شامل اجزای سازنده لایه‌ها، وزن‌ها و نرون‌ها می‌باشد. رفتار شبکه نیز وابسته به ارتباط بین اعضا است. مهم‌ترین وظیفه هر نرون، دریافت اطلاعات از نرون‌های همسایه به‌عنوان ورودی، انتساب پارامتر وزن به هر یک از ورودی‌ها، تأثیر بر ورودی‌های موجود توسط تابع عملکرد داخلی نرون جهت محاسبه پاسخ نرون به داده‌های ورودی و نهایتاً ارسال پاسخ محاسبه شده به سایر نرون‌های شبکه به‌عنوان ورودی جدید می‌باشد [۱۵]. نحوی محاسبه‌ی پاسخ شبکه از ورودی‌ها طبق روابط (۱) محاسبه می‌شوند. همچنین شکل (۱) نحوی عملکرد نرون در شبکه عصبی را نشان می‌دهد.



شکل ۱- نحوی عملکرد نرون در یک شبکه‌ی عصبی

$$A_i = f(w_{i,j}, a_j) = \left(\sum_{j=1}^n w_{i,j} a_j \right) + b_i \quad (1)$$

$$O_i = \sigma(A_i)$$

در رابطه (۱)، a ، بیانگر ورودی‌ها، w معرف وزن و b بایاس می‌باشد. مهم‌ترین ویژگی شبکه عصبی را می‌توان عدم وابستگی به فرض‌های اولیه درباره‌ی توزیع داده‌های ورودی دانست. بنابراین در استفاده از شبکه عصبی برای شناسایی ساختمان، لایه ورودی شامل نرون‌هایی است که به تعداد پارامترهای ورودی طراحی شد به نحوی که هر نرون مسئول دریافت یک پارامتر و ارسال آن به دیگر

تبدیل watershed با ترکیب ویژگی‌ها و اطلاعات بافت به شناسایی ساختمان از تصویر SAR اقدام کردند [۱۰]. Liu و همکارانش در سال ۲۰۱۴، به شناسایی ارتفاع ساختمان با استفاده از تصاویر TerraSAR-X و GIS پرداختند. آن‌ها با استفاده از ویژگی و وارونگی در ساختمان‌ها، ارتفاع ساختمان‌ها را تخمین زده و میزان جابجایی ردپاهای ساختمان‌ها را با استفاده از داده GIS به دست آورده و اصلاح کردند [۱۱]. قنبری و صاحبی در سال ۲۰۱۴، به منظور بهبود کیفیت طیفی الگوریتم IHS در مناطق شهری به ترکیب تصاویر نوری و راداری پرداختند [۱۲]. آن‌ها با استفاده از تلفیق تصویر با حد تفکیک مکانی بالاتر و تصویر شدت بر اساس اطلاعات آماری و سپس تصویر تلفیقی جایگزین تصویر شدت شده و عملیات معکوس انجام شده است. تیموری و همکارانش در سال ۲۰۱۶، روشی بهینه جهت شناسایی ساختمان از تصاویر SAR و نوری را پیشنهاد دادند [۱۳]. در این مقاله، آن‌ها با استفاده از ویژگی‌های بهینه به دست آمده از هر دو منبع داده به تلفیق داده‌ها در دو سطح ویژگی و تصمیم‌گیری پرداختند. همچنین در سال ۲۰۱۶، Zhai و همکارانش با استفاده از اطلاعات بافت هم رخداد و روش cross reclassification به معرفی روشی جدید به منظور بهبود دقت شناسایی ساختمان از تصاویر PolSAR پرداختند [۱۴].

با نگاه کلی بر تحقیقات انجام شده بر شناسایی ساختمان از تصاویر راداری می‌توان نتیجه گرفت که اغلب مطالعات صورت گرفته بر این تصاویر متکی بر اندازه‌گیری ارتفاع، بررسی سایه، تشخیص پیکسل‌های روشن می‌باشد. برخی از نقاط ضعف وارده بر این مطالعات، به‌کارگیری الگوریتم‌های پیشنهادی بر مناطقی با ساختمان‌های ساده، بزرگ و مستطیلی شکل و بدون پیچیدگی شهری است. در تحقیق حاضر تلاش شده است با به‌کارگیری طبقه‌بندی کننده‌های مختلف و ارائه روشی تلفیقی در سطح تصمیم‌گیری به شناسایی ساختمان‌ها در هر دو حالت مجزا و بلوک‌های ساختمانی از یک منطقه پیچیده شهری بپردازد. همچنین هدف دیگر این مقاله، علاوه بر بهبود دقت در شناسایی ساختمان، با بکار بردن اطلاعات همسایگی هر پیکسل سعی در کاهش نویز در تصویر رادار دارد. ساختار ادامه مقاله به شرح زیر است:

بخش دوم، مروری بر مبانی تئوری تحقیق. بخش سوم، روش تحقیق توضیح داده می‌شود. بخش چهارم

۱ Artificial Neural Network

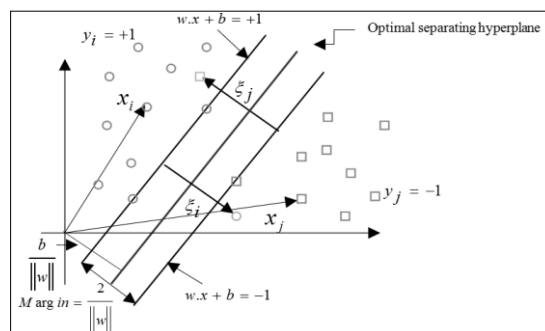
لایه‌های شبکه می‌باشد. لایه خروجی در این تحقیق، دربرگیرنده دو نرون است که وظیفه تصمیم‌گیری درباره ماهیت پیکسل ورودی را به عهده دارد.

۲-۲- ماشین بردار پشتیبان (SVM)

الگوریتم‌های ماشین بردار پشتیبان یکی از الگوریتم‌های یادگیری نظارت‌شده تشخیص الگو است که در سال ۱۹۹۵ توسط ریاضیدان روسی به نام وپنیک ارائه شد و اصول آن بر پایه نظریه یادگیری آماری است [۱۶]. مبنای اصلی این روش دسته‌بندی خطی داده‌ها با در نظر گرفتن حاشیه اطمینان بوده و اساساً یک جداکننده دودویی محسوب می‌شود که هدف اصلی آن دستیابی به ابر صفحه بهینه، به‌عنوان سطح تصمیم‌گیری به‌گونه‌ای می‌باشد که حاشیه بین دو کلاس را بیشینه کند. در صورت عدم تفکیک داده‌ها به‌صورت خطی با استفاده از کرنل‌های غیرخطی آن‌ها را فضایی با ابعاد بالاتر انتقال داده و ابر صفحه‌ای تشکیل می‌شود (شکل (۲)). فرض نمایید P داده‌ی آموزشی به‌صورت (x_i, y_i) موجود می‌باشد که در آن بردار ویژگی n بعدی و $y_i \in \{-1, 1\}$ برچسب آن می‌باشد. این ابر صفحه‌ای طبق رابطه (۲) به دست می‌آید.

$$w^T \phi(x) + b = 0 \quad (2)$$

در معادله فوق w بیانگر بردار وزن می‌باشد که به‌عنوان بردار عمود بر، ابر صفحه موردنظر و b بردار بایاس می‌باشد که یک مقدار ثابت بوده به‌طوری‌که بیانگر فاصله بین ابر صفحه‌ای و مبدأ می‌باشد و $\Phi(X)$ کرنلی است جهت انتقال داده به فضایی با ابعاد بالاتر می‌باشد.



شکل ۲- نحوه طبقه‌بندی ماشین بردار پشتیبان

همان‌طور که بیان شد هدف از این طبقه‌بندی کننده یافتن ابر صفحه‌ای است که با بیشینه نمودن حاشیه و کمینه نمودن جمع خطاهای وارد شده (روابط (۳)) به دست آید.

$$\text{Min} \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^k \xi_i \right) \quad (3)$$

$$\text{Subject to: } y_i (w\phi(x_i) + b) > 1 - \xi_i \quad i = 1, \dots, k$$

پارامتر C بیانگر پارامتر تنظیم بوده که میزان عمومیت پذیری را تنظیم می‌نماید. جهت در نظر گرفتن نویز موجود در داده و تداخل بین داده‌های آموزشی متغیر ξ مورد استفاده قرار می‌گیرد.

۲-۳- K، نزدیک‌ترین همسایه (KNN)

روش KNN، یکی از ساده‌ترین و متداول‌ترین روش‌های طبقه‌بندی تصاویر است که در آن نرخ خطای مجانبی^۳ را که در ترم‌های خطای بهینه بیز محدود می‌شود را ترکیب می‌کند. اگر Train، نمونه آموزشی از c جفت نمونه تصادفی (x_i, y_i) باشد در آن $i=1, 2, \dots, n$ و همچنین برچسب y مقادیر $\{1, 2, \dots, c\}$ را بگیرد، می‌توان آن را طبق رابطه (۴) تعریف کرد:

$$\text{Train} = \{(x_1, y_1), (x_2, y_2), \dots, (x_c, y_c)\} \quad (4)$$

$$y = \{1, 2, \dots, n\}$$

در رابطه‌ی بالا y_i کلاس x_i را در بین c کلاس محتمل مشخص می‌کند. به همین منظور جهت طبقه‌بندی، در ابتدا نزدیک‌ترین همسایگی x' از X در نمونه‌های آموزشی (رابطه (۵)) مشخص گردد [۱۷].

$$d(X, x') = \min_{x_i \in \text{Train}} d(X, x_i) \quad (5)$$

$d(X, x')$ هر فاصله‌ی در فضای ویژگی می‌تواند باشد که معمولاً روش نزدیک‌ترین همسایگی، بر اساس فواصل اقلیدسی (رابطه (۶)) تعیین می‌شوند.

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \quad (6)$$

^۲ K-Nearest Neighbors

^۳ Asymptotic Error Rate

^۱ Support Vector Machines

$$p(W_i | x) = \frac{p(x | W_i)p(W_i)}{p(x)} \quad (8)$$

$p(W_i)$ بیانگر احتمال اولیه کلاس W_i ، $p(x)$ احتمال x ، $p(W_i|x)$ احتمال تعلق x به کلاس W_i و $p(x|W_i)$ احتمال وجود کلاس W_i که x به آن تعلق داشته باشد. تابع discriminant برای طبقه‌بندی بیشترین شباهت طبق رابطه (۹) زیر به دست می‌آید.

$$g_i(x) = \ln p(w_i) - \frac{1}{2} \ln \left| \sum_i \left| -\frac{1}{2}(x - m_i)' \sum_i^{-1} (x - m_i) \right. \right| \quad (9)$$

۳- روش تحقیق

در این تحقیق، ابتدا با استفاده از طبقه‌بندی کننده‌های مختلف شامل شبکه عصبی، ماشین بردار پشتیبان، نزدیک‌ترین همسایه و بیشترین شباهت اقدام به شناسایی ساختمان شد. به همین منظور ویژگی‌های گوناگونی از تصویر رادار تولید شدند. سپس ویژگی‌های بهینه برای این مجموعه داده پس از بررسی‌های زیادی که توسط آزمون سعی و خطا بر روی ویژگی‌های تولیدشده از تصویر رادار حاصل شد. سپس این ویژگی‌های بهینه به‌عنوان داده ورودی وارد طبقه‌بندی کننده‌های مختلف شدند. سپس ضمن تجزیه و تحلیل نقاط ضعف و قوت هر یک از طبقه‌بندی کننده‌ها اقدام به تلفیق این طبقه‌بندی کننده‌ها در سطح تصمیم‌گیری شد. شکل (۳) فلوچارت روش پیشنهادی را نشان می‌دهد.

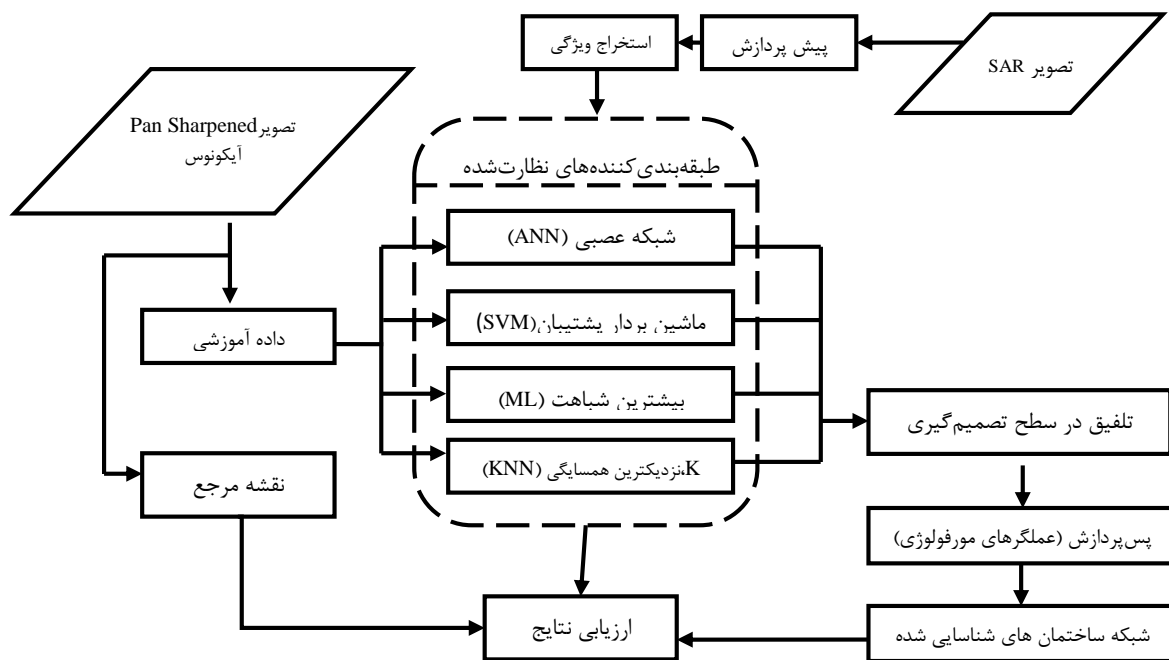
دو ویژگی که باعث محبوبیت این روش شده است، پیاده‌سازی ساده و شناخته شدن مرزهای میزان خطا می‌باشد. ولی از ضعف‌های این روش می‌توان به بار محاسباتی بالای آن و حساسیت آن به مقدار K اشاره کرد. بنابراین تعیین مقدار K در این روش نقش کلیدی را ایفا می‌کند. بدین صورت که اگر مقدار K خیلی کوچک انتخاب شود، الگوریتم به نویز حساس شده و اگر K خیلی بزرگ باشد، ممکن است در میان نزدیک‌ترین همسایه‌ها، برچسبی از کلاس‌های دیگر نیز قرار گرفته باشد.

۲-۴- طبقه‌بندی کننده بیشترین شباهت (ML)

طبقه‌بندی کننده بیشترین شباهت، یکی از متداول‌ترین روش‌های نظارت‌شده می‌باشد. روند اجرایی این طبقه‌بندی کننده آماری بدین صورت می‌باشد که بیشترین احتمال تعلق به یک کلاس موجب انتساب آن کلاس می‌شود. احتمال تعلق پیکسل x به کلاس W_i طبق رابطه (۷) تعریف می‌شود.

$$x \in W_i \text{ if } p(W_i|x) > p(W_j|x) \text{ for all } i \neq j \quad (7)$$

در صورتی که احتمال تعلق پیکسل به کلاس $p(W_i|x)$ بزرگ‌تر از احتمال تعلق پیکسل به کلاس دیگر باشد، نحوی محاسبه احتمالات از قانون بیز، طبق رابطه (۸) تعریف می‌شود.



شکل ۳- فلوچارت روش پیشنهادی

۴- پیاده سازی و ارزیابی نتایج

۴-۱- مشخصات داده‌های مورد استفاده و پیش‌پردازش

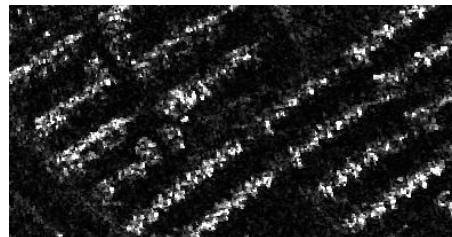
منطقه‌ی مورد مطالعه، یک منطقه‌ی شهری متشکل از تک ساختمان‌ها و بلوک‌های ساختمانی با بام‌پوش‌های متفاوت (آسفالت، ایزوگام، بتن) از شهر شیراز می‌باشد. ابعاد منطقه مورد مطالعه ۲۰۵×۳۹۵ پیکسل است. همچنین تصاویر مورد استفاده در این تحقیق (شکل ۴)، توسط ماهواره‌ی TerraSAR_x در مد Spotlight است که در سال ۲۰۱۱ از منطقه مذکور اخذ شده است. این تصویر شامل دو پلاریزاسیون HH و VV با قدرت تفکیک یک متر در جهت range و azimuth می‌باشد. مشخصات مربوط به این تصویر در جدول (۱) آورده شده است.

جدول ۱- مشخصات تصویر استفاده شده در این تحقیق

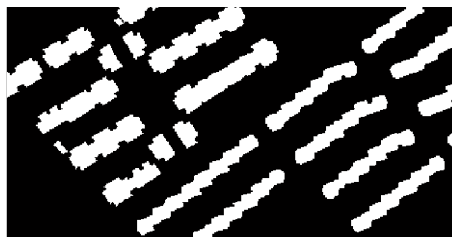
TerraSAR-x	Polarization	Equatorial crossing	IFOV at nadir	incident angle	heading angle
	HH-VV	Ascending node:18:00+0.25h	1m in spotlight mode	5.2E+0.1	3.5E+02



(ب)



(الف)



(ت)



(پ)

شکل ۴- الف) تصویر SAR در پلاریزاسیون HH (ب) تصویر SAR در پلاریزاسیون VV (پ) تصویر pan sharpened مربوط به تصویر آیکونوس (ت) تصویر باینری تولید شده جهت ارزیابی دقت نتایج

زیادی میان توان تفکیک مکانی تصویر در جهت حرکت سنجنده و عمود بر آن وجود ندارد. به همین دلیل فرآیند multi look با ۱ در ۱ پیکسل انجام شد. سپس تصویر از هندسه‌ی مایل به هندسه‌ی زمینی تبدیل شد. تصویر SAR

در این تحقیق به منظور پردازش‌های لازم بر داده‌ی خام TerraSAR_X از نرم‌افزار Gamma استفاده شد. در ابتدا در محیط نرم‌افزاری Gamma تصویر multi look تولید شده و به دلیل وجود توان تفکیک مکانی ۱ متر در این تصاویر، اختلاف

عصبی، SVM، ML و KNN معرفی شدند. در نهایت نتایج شناسایی ساختمان توسط هر طبقه‌بندی کننده مورد بررسی و ارزیابی قرار گرفته و نقاط ضعف و قوت هر طبقه‌بندی کننده شناسایی شد.

شبکه عصبی: در این تحقیق از شبکه عصبی به دلیل قابلیت آن در عدم وابستگی به فرضیه‌های اولیه درباره توزیع داده‌های ورودی مورد استفاده قرار گرفت. بنابراین انعطاف‌پذیری بالای آن را می‌توان یکی از دلایل انتخاب آن به عنوان یک طبقه‌بندی کننده مناسب دانست. داشتن یک شبکه عصبی بهینه وابسته به تعداد نرون‌های موجود در لایه‌های میانی، تعداد نمونه‌های آموزشی، الگوریتم آموزش شبکه، دقت مورد نظر و ساختار و ماهیت داده‌ها و تعداد خروجی‌ها می‌باشد و نمی‌توان یک رابطه کلی بین این متغیرها پیدا کرد در این راستا بیشتر مدل‌های پیشنهادی به‌طور تجربی بوده و یا به روش سعی و خطا طراحی شده‌اند. بنابراین در این مقاله، تعداد نرون‌های لایه ورودی بر حسب تعداد ویژگی‌های استخراج شده از تصویر رادار در نظر گرفته شد که شامل ۶ نرون می‌باشد. با توجه به اینکه لایه خروجی شامل پیکسل‌های ساختمان و زمینه می‌باشد، برای این منظور دو نرون در این لایه لحاظ شده است. بر طبق تحقیق صورت گرفته توسط ریچارد [۱۹]، طراحی یک لایه میانی برای اهداف طبقه‌بندی کفایت می‌کند، همچنین تعداد نرون‌های موجود در این لایه متغیر بوده و در کاربردهای مختلف متفاوت است. در تحقیق حاضر، تعداد نرون‌های متفاوتی (۵، ۱۰، ...، ۲۵) در لایه میانی در نظر گرفته شد. بهترین نتیجه‌ی به دست آمده مربوط به لایه میانی با ۱۰ نرون بوده است که معیارهای ارزیابی حاصل از پاسخ شبکه عصبی در جدول (۲) نشان داده شده است.

ماشین‌های بردار پشتیبان: جهت طبقه‌بندی با روش SVM از تابع کرنل گوسین استفاده شد. در این راستا تعیین بهینه پارامترهای گاما (g) و C حائز اهمیت است. تا به امروز جهت تعیین این پارامترها روش‌های بسیاری ارائه شده است اما هیچ‌یک توانایی کامل حل نیاز کاربران را نداشته است. بنابراین جهت طبقه‌بندی، مقادیر مختلفی (g=0.25, ..., 1.25, c=50, ..., 200) برای این دو پارامتر در نظر گرفته شده است. بهترین نتایج به دست آمده در مقادیر g=0.25, c=50 می‌باشد که در جدول (۲) آورده شده است.

K، نزدیک‌ترین همسایه: به منظور طبقه‌بندی به روش KNN، تعیین k بهینه نقش مؤثری در طبقه‌بندی تصویر

در حالت صعودی اخذ شده است. به همین دلیل شمال و جنوب تصویر با واقعیت منطبق نمی‌باشد. بنابراین بهتر است تصویر در راستای شمالی-جنوبی انعکاس داده شود تا عملیات انطباق با دقت بیشتر و آسان‌تر صورت پذیرد.

در نگاه اول، حذف نویز یکی از پیش‌پردازش‌های مهم در تصاویر SAR است. از آنجاکه تصاویر Terra-SAR در باند X تصویربرداری شده و دارای طول موج حدوداً ۳ سانتی‌متر می‌باشند، بنابراین نویز این تصاویر بالا است. همچنین به دلیل داشتن قدرت تفکیک مکانی ۱ متر، حذف نویز امکان کاهش کیفیت مکانی نتایج را می‌تواند در پی داشته باشد. به دلیل این که در تحقیق حاضر، اطلاعات مکانی نقش مهمی در طبقه‌بندی ایفا می‌کنند بنابراین روش پیشنهادی بر تصویر نویزدار و بدون نویز اعمال شد. به منظور حذف نویز از فیلتر Gamma با ابعاد ۳×۳ و ۵×۵ استفاده شد. در این راستا، طبقه‌بندی کننده‌های مختلف بر تصویر نویزدار و بدون نویز پیاده‌سازی شده و مورد مقایسه قرار گرفتند. از نتایج به دست آمده، این نتیجه‌گیری حاصل شد که این فرآیند باعث کاهش اطلاعات مکانی شده است، بنابراین از حذف نویز صرف نظر شد.

۲-۴- استخراج ویژگی‌های مناسب از تصویر SAR

همان‌طور که قبلاً ذکر شد، از تصویر شدت-TerraSAR در دو پلاریزاسیون HH و VV استفاده شده است. از این رو دامنه و ویژگی‌های بافت در هر دو پلاریزاسیون به عنوان ویژگی‌های مناسب در این تحقیق در نظر گرفته شدند. بنابراین توصیفگرهای آماری مرتبه اول [۱۸]، که بر اساس هیستوگرام تصویر به دست می‌آیند، به عنوان پارامترهای مؤثر در تفسیر تصویر رادار انتخاب شدند و با سعی و خطاهای بسیاری که بر پارامترهای مختلف بافت در ابعاد پنجره‌های مختلف صورت گرفت، میانگین و data range به عنوان پارامتر مؤثر در شناسایی ساختمان انتخاب گردیدند. این دو پارامتر از پنجره‌ای با ابعاد ۳×۳ از هر دو پلاریزاسیون HH و VV استخراج شدند. بنابراین علاوه بر دو دامنه تصویر، چهار پارامتر بافت نیز به عنوان ویژگی‌های بهینه انتخاب شدند.

۳-۴- طبقه‌بندی تصویر

در این مرحله ویژگی‌های استخراج شده از تصویر رادار به چهار طبقه‌بندی کننده‌ی مختلف که شامل شبکه

دارد. در این راستا از روش سعی و خطا جهت تعیین k بهینه استفاده شده است. لذا مقادیر متفاوتی $k=140,160,\dots,240$ برای k در نظر گرفته شد. بهترین نتیجه حاصل شده از $k=180$ می باشد که نتیجه بدست آمده در جدول (۲) آورده شده است.

بیشترین شباهت (ML): در این مقاله، جهت پیاده سازی روش بیشترین شباهت از تابع discrimination طبق رابطه (۹) استفاده شد. در این راستا به دلیل نبود اطلاعات کافی، احتمال یکسانی برای دو کلاس ساختمان و زمینه در نظر گرفته شد. نتیجه ی بدست آمده در جدول (۲) ارائه شده است.

جدول ۲- پارامترهای ارزیابی دقت از طبقه بندی کننده های مختلف

پارامترهای ارزیابی				بهترین پارامتر	طبقه بندی کننده
دقت شناسایی زمینه	دقت شناسایی ساختمان	ضریب کاپا	دقت کلی		
۸۷,۴۲	۶۸,۸۱	۵۴,۵۲	۸۳,۰۰	$n=10$	شبکه عصبی
۸۸,۵۷	۶۸,۶۶	۵۶,۲۱	۸۳,۸۳	$c=50, g=0,25$	SVM
۹۱,۷۳	۵۶,۶۴	۵۱,۳۵	۸۳,۳۹	$k=180$	KNN
۸۹,۳۲	۶۳,۴۷	۵۳,۲۰	۸۳,۱۷	-	ML

آن ها در نظر نگرفتن وابستگی مکانی پیکسل های مجاور در فرآیند طبقه بندی می باشد. بنابراین در روش پیشنهادی تلاش بر این شد علاوه بر تلفیق تصاویر، از اطلاعات همسایگی پیکسل ها بهره برده شود. بنابراین در این مرحله، به منظور تلفیق تصاویر طبقه بندی شده، توسط چهار روش شبکه ی عصبی، SVM و بیشترین شباهت و KNN از یک پنجره ی پیکسلی متحرک 3×3 استفاده شده است. در این پنجره تعداد آرای مربوط به پیکسل های ساختمان و زمینه شمارش شده و برچسب به کلاسی تعلق گرفته است که بیشترین آرا را کسب کرده است. جدول (۳) نتایج حاصل را نشان می دهد.

از بررسی و مقایسه نتایج طبقه بندی کننده ها در شناسایی ساختمان مشخص شد که طبقه بندی SVM تا حدودی نسبت به دیگر طبقه بندی کننده ها عملکرد بهتری داشته است. همچنین شبکه عصبی در شناسایی پیکسل های ساختمان و روش KNN در شناسایی پیکسل های زمینه عملکرد بهتری را از خود نشان داده اند. بنابراین می توان از توانایی هر یک از طبقه بندی کننده ها در جهت تلفیق تصاویر سود برد.

۴-۴- تلفیق در سطح تصمیم گیری

اگرچه نتایج به دست آمده از طبقه بندی کننده های مختلف در این تحقیق مناسب بوده است اما ضعف اصلی

جدول ۳- پارامترهای ارزیابی دقت روش پیشنهادی

پارامترهای ارزیابی				استراتژی
دقت شناسایی زمینه	دقت شناسایی ساختمان	ضریب کاپا	دقت کلی	
۹۰,۰۲	۶۸,۰۷	۵۸,۰۶	۸۴,۸۰	تلفیق در سطح تصمیم گیری
۹۰,۵۷	۷۳,۰۸	۶۲,۹۳	۸۶,۴۱	تلفیق در سطح تصمیم گیری + پس پردازش



شکل ۵- شبکه ساختمان های شناسایی شده از روش پیشنهادی

همان طور که از نتایج جدول (۳) ملاحظه می شود به کارگیری تلفیق تصاویر و استفاده از اطلاعات همسایگی باعث افزایش دقت در تمامی پارامترهای ارزیابی شده است. این فرآیند باعث کاهش نویز در طبقه بندی شده است. در مرحله ی بعد جهت اعمال پس پردازش از عملگرهای مورفولوژی (انسداد، گشایش) استفاده شد. همانطور که از نتایج جدول (۳) و شکل (۵) مشخص است اعمال پس پردازش باعث بهبود نتایج شده است.

۵- نتیجه گیری

و شبکه عصبی و K، نزدیک ترین همسایگی به ترتیب در شناسایی پیکسل های ساختمان و پیکسل های زمینه عملکرد بهتری داشته اند. به منظور بهره گیری از توانایی هر طبقه بندی کننده و رفع مشکلات آن ها اقدام به تلفیق نتایج این طبقه بندی کننده ها با استفاده از روش رأی گیری و بکار بردن اطلاعات همسایگی هر پیکسل شد. در نهایت با اعمال مورفولوژی (انسداد، گشایش) بر تصویر تلفیقی، شبکه ای از ساختمان های شناسایی شده حاصل شد. با مقایسه نتایج روش پیشنهادی با نتایج طبقه بندی کننده های مختلف، کارایی این روش اثبات شد و باعث افزایش دقت در پارامترهای ارزیابی شد. نتایج حاصل از روش پیشنهادی با دقت کلی، ضریب کاپا و دقت شناسایی ساختمان و دقت شناسایی زمینه به ترتیب ۸۶،۴۱٪، ۶۲،۹۳٪، ۷۳،۰۸٪ و ۹۰،۵۷٪ می باشد.

در این مقاله یک روش نوین تلفیق در سطح تصمیم گیری به منظور شناسایی ساختمان از تصویر SAR ارائه شد. روش پیشنهادی به منظور بهبود دقت نتایج طبقه بندی کننده ها اقدام به تلفیق در سطح تصمیم گیری می کند و هم زمان با آن برای بکار بردن اطلاعات همسایگی هر پیکسل از پنجره ای متحرک استفاده می کند. در این روش، ابتدا با استفاده از شش ویژگی استخراج شده از تصویر SAR به شناسایی ساختمان با استفاده از روش های ANN، SVM، ML و KNN اقدام شد و سپس نتایج به دست آمده از هر روش مقایسه و ارزیابی شدند. با بررسی های صورت گرفته مشخص شد که SVM نسبت به سه طبقه بندی کننده ML، KNN، NN نتایج بهتری داشته

مراجع

- [1] F. Tupin, H. Maitre, J.-F. Mangin et al., "Detection of linear features in SAR images: application to road network extraction," *IEEE Transactions on Geoscience and remote Sensing*, vol. 36, no. 2, pp. 434-453, 1998.
- [2] R. Chellappa, *Advanced Automatic Target Recognition*, DTIC Document, 1998.
- [3] F. Dell'Acqua, and P. Gamba, "Texture-based characterization of urban environments on satellite SAR images," *IEEE Transactions on Geoscience and remote Sensing*, vol. 41, no. 1, pp. 153-159, 2003.
- [4] M. Stasolla, and P. Gamba, "Spatial indexes for the extraction of formal and informal human settlements from high-resolution SAR images," *Selected Topics in Applied Earth Observations and Remote Sensing*, *IEEE Journal of*, vol. 1, no. 2, pp. 98-106, 2008.
- [5] Y. Wang, F. Tupin, C. Han et al., "Building detection from high resolution POLSAR data by combining region and edge information." pp. IV-153-IV-156.
- [6] W. He, and O. Hellwich, "Bayesian Building Extraction from high resolution polarimetric SAR data." pp. IV-402-IV-405.
- [7] M. Matsuoka, and N. Nojima, "Building damage estimation by integration of seismic intensity information and satellite L-band SAR imagery," *Remote Sensing*, vol. 2, no. 9, pp. 2111-2126, 2010.
- [8] J. D. Wegner, R. Hansch, A. Thiele et al., "Building detection from one orthophoto and high-resolution InSAR data using conditional random fields," *IEEE Journal of selected topics in applied Earth Observations and Remote Sensing*, vol. 4, no. 1, pp. 83-91, 2011.
- [9] Y. Dong, H. Chen, D. Yu et al., "Building extraction from high resolution SAR imagery in urban areas," *Geospatial Information Science*, vol. 14, no. 3, pp. 164-168, 2011.
- [10] L. Zhao, X. Zhou, and G. Kuang, "Building detection from urban SAR image using building characteristics and contextual information," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, pp. 1, 2013.
- [11] W. Liu, F. Yamazaki, B. Adriano et al., "Development of Building Height Data in Peru from High-Resolution SAR Imagery," *Journal of Disaster Research Vol.*, vol. 9, no. 6, pp. 1043, 2014.
- [12] Z. Ghanbari, and M. R. Sahebi, "Improved IHS algorithm for fusing high resolution satellite images of urban areas," *Journal of the Indian Society of Remote Sensing*, vol. 42, no. 4, pp. 689-699, 2014.
- [13] M. Teimouri, M. Mokhtarzade, and M. J. Valadan Zoej, "Optimal fusion of optical and SAR high-resolution images for semiautomatic building detection," *GIScience & Remote Sensing*, vol. 53, no. 1, pp. 45-62, 2016.

- [14] W. Zhai, H. Shen, C. Huang et al., "Fusion of polarimetric and texture information for urban building extraction from fully polarimetric SAR imagery," *Remote Sensing Letters*, vol. 7, no. 1, pp. 31-40, 2016.
- [15] P. M. Atkinson, and A. Tatnall, "Introduction neural networks in remote sensing," *International Journal of remote sensing*, vol. 18, no. 4, pp. 699-709, 1997.
- [16] V. Vapnik, "The nature of statistical learning theory. 2000," There is no corresponding record for this reference.
- [17] R. Barandela, and E. Gasca, "Decontamination of training samples for supervised pattern recognition methods." pp. 621-630.
- [18] H. Anys, A. Bannari, D. He et al., "Texture analysis for the mapping of urban areas using airborne MEIS-II images." pp. 231-245.
- [19] J. A. Richards, and J. Richards, *Remote sensing digital image analysis*: Springer, 1999.