

ارائه روشی نوین بر پایه الگوریتم‌های یادگیری چندکرنلی برای طبقه‌بندی محصولات کشاورزی با استفاده از سری‌های زمانی چند متغیره

سعید نیازمردی^{۱*}، عبدالرضا صفری^۲، سعید همایونی^۳

^۱ دانشجوی دکتری نقشه‌برداری، سنجش از دور - دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - پردیس

دانشکده‌های فنی - دانشگاه تهران

s.niazmardi@ut.ac.ir

^۲ دانشیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - پردیس دانشکده‌های فنی - دانشگاه تهران

asafari@ut.ac.ir

^۳ استادیار گروه جغرافیا، محیط زیست و ژئوماتیک - دانشگاه Ottawa

saeid.homayouni@uOttawa.ca

(تاریخ دریافت دی ۱۳۹۵، تاریخ تصویب اردیبهشت ۱۳۹۶)

چکیده

امروزه، تشخیص دقیق نوع محصولات کشاورزی با استفاده از طبقه‌بندی سری‌های زمانی حاصل از تصاویر سنجنده‌های مختلف سنجش از دوری، رو به افزایش است. دسته‌ای از سری‌های زمانی که با استفاده از تصاویر سنجنده‌های چندطیفی و یا فراطیفی ایجاد می‌شوند، از نوع سری‌های زمانی چندمتغیره هستند. علیرغم محتوای بالای اطلاعاتی این نوع از سری‌های زمانی، الگوریتم‌های طبقه‌بندی موجود، به دلیل چهار - بعدی بودن این نوع از داده‌ها، قادر به طبقه‌بندی آنها نیستند. در مقاله پیش رو، به منظور رفع این مشکل و طبقه‌بندی سری‌های زمانی چندمتغیره، روشی بر پایه‌ی استفاده از الگوریتم‌های یادگیری چندکرنلی ارائه شده است. در این روش، ابتدا یک کرنل با استفاده از داده‌ی اخذ شده در هر تاریخ، ساخته می‌شود و سپس این کرنل‌ها با استفاده از الگوریتم‌های یادگیری چندکرنلی در قالب کرنلی به نام کرنل ترکیبی، با یکدیگر ادغام می‌شوند. سپس از کرنل ترکیبی حاصل، برای طبقه‌بندی داده‌ها در الگوریتم‌های طبقه‌بندی مبتنی بر کرنل استفاده می‌شود. در ادامه برای ارزیابی روش پیشنهادی، از دو سری‌زمانی تصاویر سنجش از دوری مختلف استفاده شده است. هر دو سری‌زمانی شامل ده تصویر اخذ شده توسط سنجنده‌های RapidEye هستند که از منطقه‌ای کشاورزی در کشور کانادا تهیه شده‌اند. از روش الحاق داده‌ها، که در آن داده‌ی زمان‌های مختلف سری‌زمانی در قالب یک مکعب داده به هم الحاق می‌شوند، به عنوان روش معیار برای مقایسه با روش پیشنهادی استفاده شده است. کرنل ترکیبی حاصل از روش پیشنهادی و همچنین کرنل حاصل از روش الحاق داده‌ها، برای آموزش الگوریتم طبقه‌بندی ماشین‌های بردار پشتیبان بکار گرفته شده است. در این مقاله برای محاسبه کرنل ترکیبی، علاوه بر استفاده از چند الگوریتم متداول یادگیری چندکرنلی، الگوریتم یادگیری چندکرنلی تعمیم یافته نیز برای نخستین بار در حوزه طبقه‌بندی سری‌زمانی معرفی شده است. مقایسه دقت طبقه‌بندی الگوریتم ماشین‌های بردار پشتیبان در دو حالت، نشان دهنده‌ی دقت بسیار بالاتر روش پیشنهادی بود. همچنین نتایج مقایسه الگوریتم‌های مختلف یادگیری چندکرنلی در چارچوب روش پیشنهادی، نشان داد که الگوریتم یادگیری چندکرنلی تعمیم یافته علاوه بر دقت بیشتر، حساسیت کمتری نیز به نویز موجود در سری‌های زمانی دارد.

واژگان کلیدی: طبقه‌بندی سری‌زمانی، سری‌زمانی چندمتغیره، الگوریتم یادگیری چندکرنلی، الگوریتم یادگیری چندکرنلی تعمیم یافته

۱- مقدمه

تشخیص نوع محصولات کشاورزی از موضوعات حائز اهمیت برای مسئولین و تصمیم‌گیران حوزه کشاورزی دقیق است؛ چرا که نوع محصول یکی از مهمترین اطلاعات لازم برای انجام بسیاری از پردازش‌ها و مطالعات این حوزه است. برای مثال تخمین سطح زیر کشت محصولات، پیش‌بینی حجم محصول قابل برداشت و تخمین آب مورد نیاز هر محصول، همگی در گرو دانستن نوع محصول است [۱، ۲]. امروزه استفاده از طبقه‌بندی داده‌های سنجش از دوری به علت پوشش وسیع جغرافیایی و همچنین اطلاعات طیفی مناسب این داده، با هدف تشخیص نوع محصولات کشاورزی رو به افزایش است [۳، ۴]. با این وجود به دلیل تغییرات زیاد طیف بازتابی محصولات کشاورزی در طول دوره رشد آنها، نمی‌توان با استفاده از اطلاعات موجود در یک تصویر، به دقت مطلوبی برای طبقه‌بندی رسید [۵]. برای رفع این مشکل، از تصاویر ماهواره‌ای به صورت سری‌های زمانی^۱ (SITS) برای طبقه‌بندی محصولات کشاورزی استفاده می‌شود. سری‌های زمانی تصاویر ماهواره‌ای، در اصطلاح به مجموعه‌ای از تصاویر سنجش از دوری گفته می‌شود که در زمان‌های مختلف و از یک منطقه واحد اخذ شده باشند. این داده‌ها به دلیل دارا بودن داده‌ی چندین تاریخ مختلف، دربردارنده اطلاعات بسیار ارزشمندی در مورد تغییر طیف بازتابی محصولات در دوره رشد آنها هستند. این اطلاعات، که از داده‌های هیچ یک از تاریخ‌های سری‌زمانی به تنهایی قابل استخراج نیستند، می‌توانند باعث تفکیک‌پذیری بیشتر محصولات و افزایش دقت طبقه‌بندی شوند [۶].

سری‌های زمانی تصاویر ماهواره‌ای را می‌توان به دو دسته تک‌متغیره^۲ و چندمتغیره^۳ تقسیم کرد [۷]. سری‌های زمانی تک‌متغیره، به سری‌زمانی‌هایی اطلاق می‌شود که داده در هر تاریخ از آنها به صورت یک تصویر تک‌باندی قابل ارائه است [۸]. یکی از رایج‌ترین راه‌ها برای به دست آوردن این نوع از سری‌زمانی، استخراج یک شاخص گیاهی از هر تاریخ داده است. سری‌زمانی‌های حاصل از شاخص‌های گیاهی مختلف، به صورت گسترده‌ای برای پردازش‌های گوناگون در حوزه کشاورزی

دقیق از جمله طبقه‌بندی محصولات استفاده شده‌اند [۵، ۹-۱۲]. دسته‌ی دوم، سری‌زمانی‌های چندمتغیره هستند. داده‌ی هر تاریخ در این نوع از سری‌های زمانی، به صورت یک تصویر چندباندی قابل ارائه است. تمام سری‌زمانی‌هایی را که با استفاده از سنجنده‌های چندطیفی و فراطیفی اخذ می‌شوند، می‌توان در این دسته طبقه‌بندی نمود. از آنجایی که استفاده از چند باند محدود داده برای محاسبه شاخص‌های گیاهی، باعث از دست رفتن اطلاعات باندهایی می‌شود که در محاسبه شاخص استفاده نشده‌اند، سری‌های زمانی چندمتغیره محتوای اطلاعاتی بسیار بیشتری در مقایسه با سری‌های زمانی تک‌متغیره‌ی شاخص‌های گیاهی دارند.

علیرغم محتوای اطلاعاتی بالای سری‌های زمانی چندمتغیره، طبقه‌بندی آنها براحتی قابل انجام نیست. چرا که الگوریتم‌های موجود به دلیل ۴-بعدی بودن سری‌های زمانی چندمتغیره (با در نظر گرفتن زمان به عنوان بعد چهارم)، قادر به طبقه‌بندی این نوع از سری‌زمانی نیستند [۸]. به همین علت طبقه‌بندی سری‌های زمانی چندمتغیره در سنجش از دور به عنوان مسئله‌ای حل نشده باقی مانده است.

ساده‌ترین راهکار برای حل این مشکل و طبقه‌بندی سری‌های زمانی چندمتغیره، استفاده از روشی است که در این مقاله، الحاق داده‌ها نامیده شده است. در این روش، ابتدا داده‌های اخذ شده در زمان‌های مختلف به هم الحاق شده و سپس مکعب داده حاصل از این کار برای طبقه‌بندی استفاده می‌شود. این راهکار علیرغم عملی بودن، دو مشکل عمده دارد. اول اینکه تصاویر اخذ شده در تاریخ‌های مختلف سری‌های زمانی، اطلاعات مختلف و گاهاً ناهمگونی^۴ در مورد طبقه‌بندی محصولات کشاورزی دارند، که این راهکار قادر به رفع این مشکل نیست. برای مثال طیف دو محصول کشاورزی مختلف ممکن است در تصاویر ابتدا و انتهای سری‌زمانی بسیار به هم شبیه باشند و اما طیف همین دو محصول در تصاویر میانه سری‌زمانی به خوبی از هم قابل تفکیک باشند. دومین مشکل این راهکار، بالا بودن ابعاد مکعب داده حاصل از الحاق داده‌ها است، که احتمال رخداد پدیده Hughes را افزایش می‌دهد [۱۳].

بنابراین برای طبقه‌بندی سری‌های زمانی چندمتغیره، این داده‌ها باید به نحوی ارائه شوند که علاوه بر قابل

^۱ Satellite Image Time-Series (SITS)

^۲ Univariate Satellite Image Time-series

^۳ Multi-variate Satellite Image Time-series

^۴ Heterogenous

استفاده از الگوریتم‌های MKL ارائه شده است. از آنجایی که اغلب الگوریتم‌های MKL برای طبقه‌بندی SVM ارائه شده‌اند، در این بخش مبانی این الگوریتم‌ها برای حالت طبقه‌بندی دوکلاسه شرح داده می‌شود. برای استفاده از این الگوریتم‌ها در حالت طبقه‌بندی چندکلاسه می‌توان با استفاده از استراتژی‌هایی مثل یک-دربار-یک یا یک-دربار-همه، مسئله طبقه‌بندی چندکلاسه را به چندین طبقه‌بندی دوکلاسه تقسیم نمود و از روش‌های MKL برای هر یک استفاده کرد [۱۷، ۱۸].

۲-۱- الگوریتم‌های MKL

فرض کنید که مجموعه $T = \{\mathbf{x}_i, y_i\}_{i=1, \dots, n}$ شامل n نمونه‌ی آموزشی $\mathbf{x}_i \in \mathbb{R}^D, i=1, \dots, n$ به همراه برچسب کلاس‌های متناظر آنها $y_i \in \{+1, -1\}$ در دسترس باشد. با استفاده از این داده، مجموعه‌ای از M کرنل مبنا $\{K_1, K_2, \dots, K_M\}$ محاسبه شده است، که فضای کرنل هریک از آنها با $\mathcal{H}_m$ نمایش داده می‌شود. کرنل‌های مبنا می‌توانند از ویژگی‌های مختلف داده و یا انتخاب توابع کرنل مختلف و از یک داده محاسبه شوند [۱۴].

هدف الگوریتم‌های MKL تخمین ترکیب بهینه‌ای از کرنل‌های مبنا برای بیشینه‌سازی استخراج اطلاعات از داده است [۱۸]. برای دستیابی به این هدف، این الگوریتم‌ها، با استفاده از یک تابع ترکیب^۶، کرنلی به نام کرنل ترکیبی^۷ سازند [۱۴]. کرنل ترکیبی محتوای اطلاعات بسیار غنی-تری برای طبقه‌بندی داده‌ها در مقایسه با هر یک از کرنل‌های مبنا خواهد داشت، چرا که تمام اطلاعات مفید هر کرنل مبنا توسط تابع ترکیب به این کرنل انتقال یافته است. تابع ترکیب می‌تواند هر تابعی از کرنل‌های مبنا باشد مادامی که کرنل ترکیبی حاصل از آن، یک کرنل مثبت-معین باشد. انواع مختلفی از این تابع ارائه شده است، اما معمولاً تابع خطی که در معادله زیر نشان داده شده است، به عنوان تابع ترکیب استفاده می‌شود [۱۹].

$$\mathbf{K}_c = \sum_{i=1}^M d_i \mathbf{K}_i \quad (1)$$

^۵ Features

^۶ Combination function

^۷ Composite Kernel

طبقه‌بندی بودن با الگوریتم‌های طبقه‌بندی موجود، تاثیر ناهمگونی داده‌ها نیز بر روی نتایج تا حد ممکن کاهش بیابد. به این منظور در این مقاله، راهکاری برای طبقه‌بندی سری‌های زمانی چندمتغیره با استفاده از الگوریتم یادگیری چندکرنلی^۱ (MKL) ارائه شده است. الگوریتم‌های MKL گروهی از الگوریتم‌های یادگیری کرنل^۲ هستند که هدف آنها بهبود دقت و کارایی الگوریتم‌های مبتنی بر کرنل (مانند الگوریتم ماشین‌های بردار پشتیبان^۳، SVM) است. به این منظور الگوریتم‌های MKL به جای استفاده از یک کرنل در الگوریتم‌های مبتنی بر کرنل، از یک کرنل ترکیبی که حاصل ادغام گروهی از کرنل‌ها است، استفاده می‌کنند. تفاوت الگوریتم‌های MKL مختلف، روش‌های گوناگونی استفاده شده توسط آنها برای بهینه کردن ترکیب کرنل‌هاست [۱۴].

در سال‌های اخیر الگوریتم‌های MKL مختلفی در حوزه‌های یادگیری ماشین و همچنین سنجش از دور ارائه شده‌اند. یکی از موفق‌ترین این الگوریتم‌ها، الگوریتم یادگیری چندکرنلی تعمیم یافته^۴ (GMKL) است [۱۵]. این الگوریتم که اخیراً برای طبقه‌بندی داده‌های سنجش از دوری نیز استفاده شده است [۱۶]، علیرغم داشتن نتایج بهتری از سایر الگوریتم‌ها، تاکنون برای طبقه‌بندی سری‌های زمانی استفاده نشده است. بنابراین جنبه‌های نوآورانه این مقاله عبارتند از: ارائه روشی جدید بر مبنای الگوریتم‌های MKL برای طبقه‌بندی سری‌های زمانی چندمتغیره و همچنین ارزیابی الگوریتم GMKL و مقایسه آن با سایر الگوریتم‌های MKL برای طبقه‌بندی سری‌های زمانی چندمتغیره.

در ادامه این مقاله، در بخش دوم، الگوریتم یادگیری چندکرنلی، الگوریتم GMKL و همچنین روش پیشنهادی برای طبقه‌بندی سری‌های زمانی چندمتغیره با استفاده از الگوریتم GMKL معرفی می‌شوند. در بخش سوم، داده‌ها و پیاده‌سازی روش‌ها بررسی می‌گردند. نتایج در بخش چهارم و نهایتاً نتیجه‌گیری در بخش انتهایی مقاله ارائه شده اند.

۲- روش شناسی

در این بخش ابتدا مبانی الگوریتم‌های MKL همچنین الگوریتم GMKL توضیح داده خواهند شد و در انتها روش پیشنهادی برای طبقه‌بندی سری‌های زمانی چندمتغیره با

^۱ Multiple Kernel Learning (MKL)

^۲ Kernel Learning algorithms

^۳ Support Vector Machines (SVM)

^۴ Generalized Multiple Kernel Learning (GMKL)

با استفاده از این تابع ترکیب، علاوه بر ساده‌تر شدن محاسبات نتایج بهتری در مقایسه با سایر انواع غیرخطی توابع بدست خواهد آمد [۱۸]. در این تابع $d_i, i=1, \dots, M$ وزن‌های غیرمنفی متناظر با کرنل مبنای i - ام هستند. بنابراین با استفاده از این تابع ترکیب، مسئله تخمین بهینه‌ترین کرنل ترکیبی توسط MKL، تبدیل به مسئله بهینه کردن وزن کرنل‌های مبنای می‌شود. الگوریتم‌های MKL بسیار مختلفی در سال‌های اخیر ارائه شده است، که از روش‌ها و فرضیات مختلفی برای بهینه‌سازی وزن‌ها استفاده می‌کنند. در بین تمام الگوریتم‌های MKL، پراستفاده‌ترین روش برای تخمین بردار وزن‌ها $\mathbf{d} = (d_1, \dots, d_M)^T$ ، بهینه کردن خطای طبقه‌بندی تنظیم شده^۱ است. در این حالت مسئله بهینه‌سازی MKL به صورت زیر قابل بیان خواهد بود:

$$\min_{\mathbf{w}, \mathbf{b}, \mathbf{d}} \left(\frac{1}{2} \|\mathbf{w}\|_{\mathcal{H}_c}^2 + C \sum_{i=1}^n \ell(y_i, f(x_i)) \right) \quad (2)$$

Subject to: $\mathbf{d} \in \Delta$

در این رابطه $\ell(y_i, f(x_i)) = \max(0, 1 - (y_i f(x_i)))$ خطای طبقه‌بندی^۲ است. تابع طبقه‌بندی (ابرفضا جداکننده) با $f(\mathbf{x}) = \mathbf{w}^T \phi_c(\mathbf{x}) + b$ نشان داده شده است که در آن $\mathbf{w}$ بردار نرمال بر ابرفضا در فضای کرنل ترکیبی $\mathcal{H}_c$ است. پارمتر b و C به ترتیب نمایانگر بایاس و پارمتر تنظیم الگوریتم هستند. در این رابطه Δ یک مجموعه محدب است که وزن کرنل‌های مبنای از اعضای این مجموعه انتخاب می‌شود. می‌توان نشان داد که معادله (۲) به صورت فرم دوگانه^۳ زیر قابل بازنویسی است:

$$\min_{\mathbf{d}} \max_{\mathbf{a}} \left(1^T \mathbf{a} - \frac{1}{2} (\mathbf{a} \circ \mathbf{y})^T \mathbf{K}_c(\mathbf{d})(\mathbf{a} \circ \mathbf{y}) \right)$$

$$\sum_{i=1}^n \alpha_i y_i = 0 \quad (3)$$

Subject to: $0 \leq \alpha_i \leq C \quad \forall i=1, \dots, n$
 $\mathbf{d} \in \Delta$

که در آن $\mathbf{a} = (\alpha_1, \dots, \alpha_n)^T$ بردار ضربگرهای لاگرانژ است و $\mathbf{K}_c(\mathbf{d})$ کرنل ترکیبی است که از معادله (۱) بدست

می‌آید. $\mathbf{y} = (y_1, \dots, y_n)^T$ بردار برچسب‌هاست که ضرب مولفه به مولفه آن با بردار ضربگرهای لاگرانژ به صورت $(\mathbf{a} \circ \mathbf{y})$ نشان داده شده است.

تفاوت روش‌های مختلف MKL در روش استفاده شده و همچنین فرضیات آنها برای بهینه‌سازی معادلات (۲) و (۳) است. استفاده از الگوریتم‌های MKL در سال‌های اخیر مورد توجه محققان زیادی در سنجش از دور قرار گرفته است. روش‌های MKL برای ترکیب ویژگی‌های طیفی-مکانی داده‌ها در [۲۰] استفاده شده است. این الگوریتم‌ها همچنین برای ترکیب داده‌های سنجنده‌های مختلف استفاده شده‌اند [۲۱, ۲۲]. در این مقالات تعیین وزن‌های هر کرنل با استفاده از روش‌های استقرائی انجام شده است. الگوریتم SimpleMKL برای تعیین میزان اهمیت هر ویژگی برای طبقه‌بندی داده در مقاله [۲۳] استفاده شده است. در این مطالعه کرنل‌های مختلف (محاسبه شده از ویژگی‌های گوناگون داده) با این الگوریتم با هم ترکیب شده‌اند و وزن هر کرنل به عنوان میزان اهمیت آن تلقی شده است. دسته‌ی دیگری از الگوریتم‌های MKL برای داده‌های سنجش از دوری پیشنهاد شده‌اند، که در این مقاله، روش‌ها زیرفضا نامیده می‌شوند. در این روش‌ها ابتدا تمام کرنل‌های مبنای بردارهایی تبدیل می‌شوند و سپس یک ماتریس از الحاق این بردارها به هم ایجاد می‌شود. پس از این مرحله، ضرایب انتقال این ماتریس به یک زیرفضای یک بعدی از آن به عنوان ضرایب کرنل‌ها استفاده می‌شود. الگوریتم‌های مختلفی با استفاده از روش‌های گوناگون استخراج زیرفضا پیشنهاد شده است. برای نمونه الگوریتم Representative MKL (RMKL) از روش تجزیه مقادیر ویژه منفرد^۴، الگوریتم Discriminative MKL، از ضابطه فیشر و همچنین روش no-negative matrix factorization برای تعیین زیرفضا استفاده می‌کنند [۲۴-۲۶].

از بین الگوریتم‌های مختلف MKL در اینجا الگوریتم GMKL به دلیل نتایج بهتر و همچنین دارا بودن مبانی نظری قوی، با جزئیات بیشتر ارائه خواهد شد. برای مطالب بیشتر در مورد جزئیات و انواع الگوریتم‌ها و همچنین تفاوت‌های آنها، می‌توان به مقالات [۱۴, ۱۶, ۱۸] مراجعه نمود.

^۱ Regularized Classification Error

^۲ خطای طبقه‌بندی انواع مختلفی دارد، اما در اینجا از متداول‌ترین آنها که با نام Hinge Loss شناخته می‌شود، استفاده شده است.

^۳ Dual form

^۴ Singular Value Decomposition

۲-۲- الگوریتم GMKL

الگوریتم GMKL در سال ۲۰۰۹ توسط Varma و Babu ارائه شده است [۱۵]. در این الگوریتم معادله (۲)، با حذف محدب بودن مجموعه Δ و اضافه شدن یک جمله تنظیم کننده^۱ ($\Omega(\mathbf{d})$) به این معادله، به صورت زیر بازنویسی شده است:

$$\min_{\mathbf{w}, \mathbf{b}, \mathbf{d}} \left(\frac{1}{2} \|\mathbf{w}\|_{\mathcal{H}_c}^2 + C \sum_{i=1}^n \ell(y_i, f(x_i)) + \Omega(\mathbf{d}) \right) \quad (4)$$

Subject to: $\mathbf{d} \geq 0$

در این حالت به راحتی می توان نشان داد، که فرم دوگانه بهینه سازی بالا، به صورت زیر قابل محاسبه است:

$$\min_{\mathbf{d}} \max_{\mathbf{a}} \left(\mathbf{1}^T \mathbf{a} - \frac{1}{2} (\mathbf{a} \circ \mathbf{y})^T \mathbf{K}_c(\mathbf{d})(\mathbf{a} \circ \mathbf{y}) + \Omega(\mathbf{d}) \right) \quad (5)$$

$$\sum_{i=1}^n \alpha_i y_i = 0$$

subject to: $0 \leq \alpha_i \leq C \quad \forall i=1, \dots, n$

$$\mathbf{d} \geq 0$$

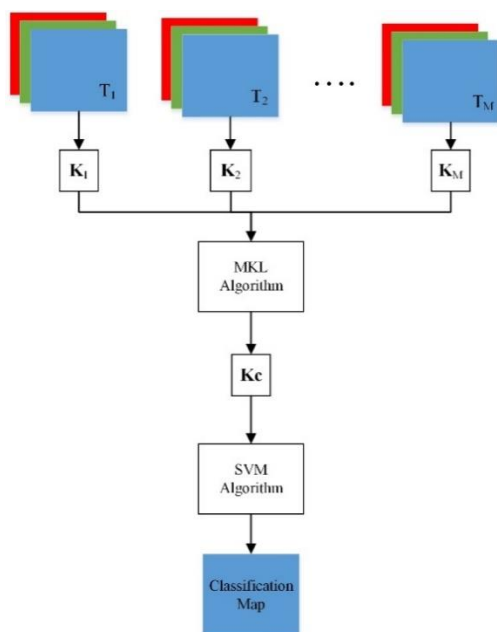
همان طور که از معادله (۵) قابل مشاهده است، این بهینه سازی نسبت به ضربگرهای لاگرانژ از نوع بهینه سازی محدب و نسبت به ضرایب کرنل های مبنا از نوع مقعر است. برای حل همزمان این بهینه سازی نسبت به هر دو دسته (ضربگرهای لاگرانژ و ضرایب کرنل)، می توان از استراتژی بهینه سازی متناوب استفاده نمود. در این استراتژی ابتدا مقادیر ضرایب کرنل معلوم فرض می شود (در اولین مرحله همه ضرایب برابر با هم و یا به صورت اتفاقی اختصاص می یابند) و با استفاده از یک بهینه سازی محدب ضرایب لاگرانژ تخمین زده می شوند. در مرحله دوم با ثابت فرض کردن ضرایب لاگرانژ و با فرض اینکه جمله تنظیم کننده نسبت به ضرایب مشتق پذیر باشد، از روش بیشینه شیب برای تخمین ضرایب کرنل ها استفاده می شود. این دو مرحله ی بهینه سازی تا یک تعداد تکرار معین یا برقراری شرطی برای خاتمه عملیات تکرار می شوند.

الگوریتم GMKL مزایای زیادی در مقایسه با سایر الگوریتم های MKL دارد. اول اینکه برای تخمین ضربگرهای لاگرانژ، می توان در آن از روش های حل SVM استفاده نمود. چرا که مسئله بهینه سازی معادله (۵) با فرض معلوم

بودن ضرایب کرنل های مبنا تبدیل به مسئله بهینه سازی الگوریتم SVM می شود. ثانیاً در این الگوریتم، می توان از هر تابع مشتق پذیر نسبت به $\mathbf{d}$ به عنوان عبارت تنظیم کننده استفاده کرد. در انتها، این الگوریتم می تواند برای ترکیب غیرخطی کرنل ها نیز مورد استفاده قرار بگیرد.

۲-۳- طبقه بندی سری های زمانی با استفاده از GMKL

برای طبقه بندی سری های زمانی چندمتغیره با استفاده از الگوریتم های MKL، محاسبه کرنل های مبنا از اهمیت ویژه ای برخوردار است. برای استخراج بیشینه اطلاعات ممکن از داده، باید ابتدا آن را به کوچکترین اجزای حاوی اطلاعات تقسیم نمود و سپس اطلاعات هر یک از این اجزا را استخراج و با هم ترکیب نمود [۲۷]. برای سری های زمانی چندمتغیره، داده اخذ شده در تاریخ های مختلف سری زمانی را می توان به عنوان اجزای اطلاعاتی در نظر گرفت. پس در روش پیشنهادی، ابتدا از هر تاریخ داده سری زمانی یک کرنل به عنوان کرنل مبنا محاسبه می شود. سپس، این کرنل ها با استفاده از الگوریتم GMKL (یا هر الگوریتم MKL دیگری) با هم ترکیب می شوند و در نهایت کرنل ترکیبی حاصل از این الگوریتم برای آموزش الگوریتم SVM و طبقه بندی نهایی داده استفاده می شود. شکل (۱)، مراحل طبقه بندی یک سری زمانی با M تاریخ داده را با استفاده از روش پیشنهادی نشان می دهد.



شکل ۱- مراحل اجرایی الگوریتم پیشنهادی برای طبقه بندی سری زمانی چندمتغیره

^۱ Regularization Term

۳- داده و پیاده‌سازی

۳-۱- داده‌ی سری زمانی

برای ارزیابی دقت الگوریتم GMKL برای طبقه‌بندی محصولات کشاورزی، دو سری زمانی از تصاویر ماهواره‌ای از منطقه‌ای کشاورزی در نظر گرفته شده است. هر دو سری‌زمانی متشکل از ده تصویر زمین-مرجع شده هستند که توسط سنجنده‌های RapidEye و در سال زراعی ۲۰۱۲، در منطقه Winnipeg واقع در استان Manitoba در کشور کانادا، اخذ شده‌اند. زمان اخذ تصاویر در جدول (۱) نشان داده شده است. داده‌ها با قدرت تفکیک ۵ متر و در پنج باند طیفی اخذ شده‌اند. به دلیل وسعت منطقه و همچنین گوناگونی محصولات موجود، دو منطقه کوچک‌تر با ابعاد ۲۵ کیلومترمربع برای ارزیابی استفاده شده است. این دو منطقه که دربردارنده محصولات کشاورزی عمده منطقه هستند، در این مقاله با نام‌های P1 و P2 نامگذاری شده‌اند. برای هر دو منطقه مورد مطالعه، با استفاده از نقشه محصولات موجود، مزارع شناسایی شده و از آنها نمونه‌های آموزشی و نمونه‌های ارزیابی استخراج شده است. نام محصولات هر منطقه و همچنین تعداد نمونه‌های انتخاب شده به عنوان نمونه‌ی ارزیابی و آموزشی در هر منطقه در جدول (۲) و جدول (۳) نشان داده شده است. علاوه بر این تصاویر هر تاریخ داده از هر منطقه در شکل (۲) و (۳) نشان داده شده است.

جدول ۱- تاریخ اخذ داده‌ها در سری‌های زمانی	
شماره	تاریخ تصویربرداری
۱	May ۱۴
۲	May ۲۰
۳	June ۴
۴	June ۲۸
۵	July ۵
۶	July ۲۱
۷	August ۱۹
۸	August ۲۹
۹	September ۹
۱۰	September ۱۴

جدول ۲- نام محصولات و تعداد نمونه‌های آموزشی و ارزیابی به ازای هر کلاس برای سری‌زمانی P1

شماره	نام محصول	تعداد نمونه آموزشی	تعداد نمونه ارزیابی
۱	ذرت	۱۳۷	۱۴۱۱
۲	کائولا	۱۴۹	۴۱۰۵
۳	گندم	۱۴۷	۲۲۴۷
۴	سویا	۱۵۲	۲۰۹۰
۵	جو دو سر	۱۳۰	۱۲۵۲
۶	علوفه	۱۰۰	۱۶۴
جمع		۸۱۵	۱۱۲۶۹

جدول ۳- نام محصولات و تعداد نمونه‌های آموزشی و ارزیابی به ازای هر کلاس برای سری‌زمانی P2

شماره	نام محصول	تعداد نمونه آموزشی	تعداد نمونه ارزیابی
۱	کائولا	۲۱۲	۴۴۹۳
۲	گندم	۱۷۵	۲۹۴۸
۳	سویا	۲۱۷	۴۰۶۲
۴	جو دوسر	۱۷۸	۲۶۴۲
۵	علوفه	۳۵	۴۳۹
۶	آفتابگردان	۹۹	۲۲۲
۷	جو	۹۵	۲۲۴
جمع		۱۰۱۱	۱۵۰۳۰

۳-۲- پیاده‌سازی

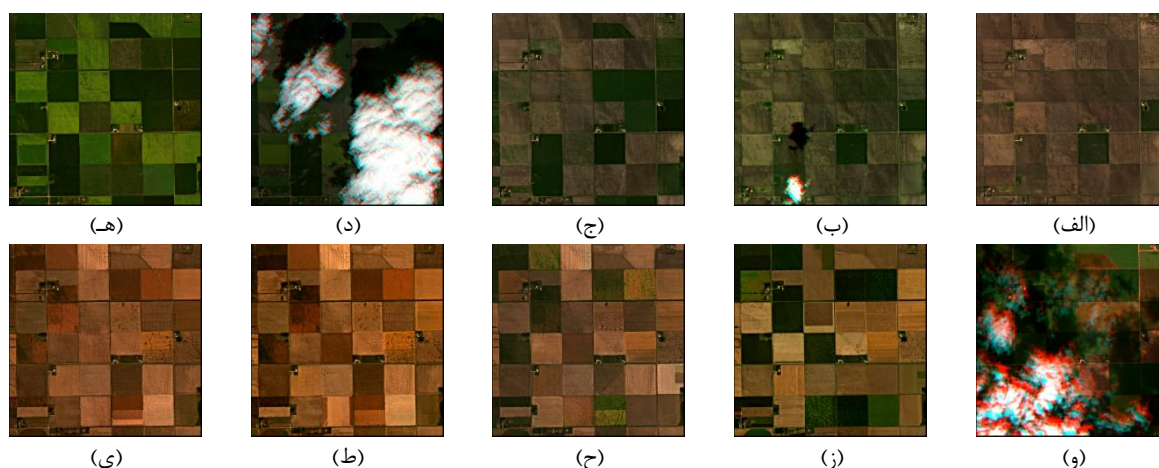
پیاده‌سازی‌های این مقاله، به منظور بررسی و مقایسه دقت طبقه‌بندی سری‌های زمانی چندمتغیره با استفاده از الگوریتم GMKL و روش الحاق داده‌ها و همچنین مقایسه کارایی بین الگوریتم‌های مختلف MKL انجام شده است. به این منظور در این مقاله، دقت طبقه‌بندی الگوریتم SVM آموزش دیده با کرنل‌های ترکیبی الگوریتم‌های MKL مختلف و همچنین کرنل حاصل از روش الحاق داده‌ها به عنوان معیار ارزیابی استفاده شده است. برای محاسبه تمام کرنل‌ها (کرنل‌های مبنا و همچنین کرنل مکعب داده حاصل از الحاق داده‌ها)، از تابع کرنل RBF استفاده شده است. این کرنل به دلیل داشتن نتایج خوب و همچنین پیچیدگی محاسباتی کم انتخاب شده است [۲۸]. پارامتر این کرنل با استفاده از روش cross-validation، ۵- لایه و از مجموعه اعداد بین ۰/۱ تا ۱۰ با افزایش ۰/۵ انتخاب شده است. پارامتر تنظیم‌کننده SVM، نیز با استفاده از روش

۴- نتایج

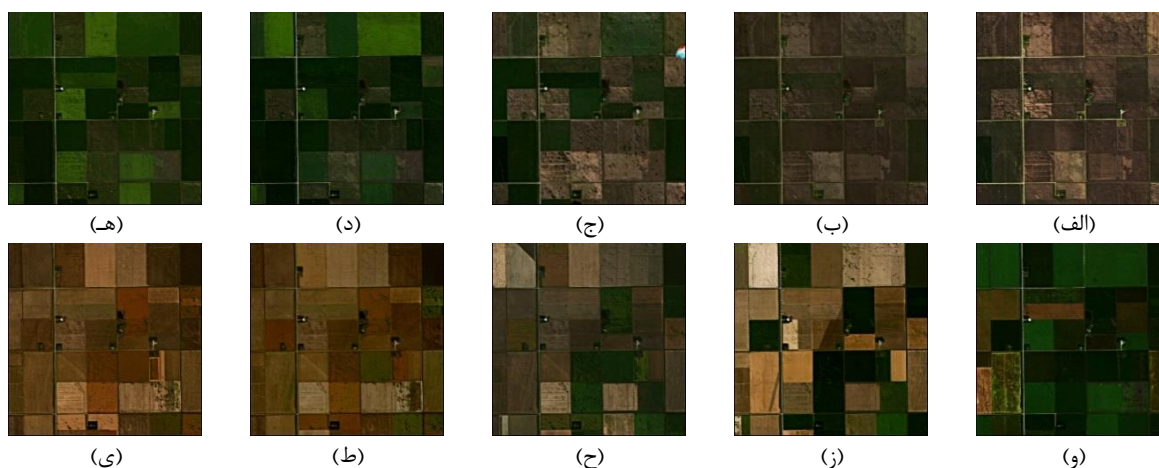
جدول ۴ نتایج حاصل از طبقه‌بندی را برای سرزمانی‌های P1 و P2 نشان می‌دهند. در این جدول دقت کلی طبقه‌بندی (بر حسب درصد)، شاخص Kappa، شاخص میانگین دقت هر کلاس (بر حسب درصد) و همچنین زمان محاسبات (برای الگوریتم‌های MKL و برحسب ثانیه) گزارش شده است.

cross-validation، ۵- لایه و از مجموعه اعداد بین بازه $0/01$ - 200 با افزایش ۵۰ واحدی انتخاب شده است.

در این مقاله نتایج الگوریتم GMKL با چندین الگوریتم MKL دیگر در قالب روش پیشنهادی مقایسه شده‌اند که عبارتند از: SimpleMKL [۲۹]، KL-sum [۱۶]، KA-MKL [۱۶]، [۲۴] Repretative MKL (RMKL). این مقایسه بر مبنای دقت طبقه‌بندی و همچنین زمان محاسبات هر الگوریتم انجام شده است. در اینجا برای طبقه‌بندی چندکلاسه از استراتژی یک-در-برابر همه استفاده شده است.



شکل ۲- تصاویر RGB از تاریخ‌های مختلف سری زمانی P1، (الف) May-۱۴، (ب) May-۲۰، (ج) Jun-۴، (د) Jun-۲۸، (ه) Jul-۵، (و) Jul-۲۱، (ز) Aug-۱۹، (ح) Aug-۲۹، (ط) Sep-۹، (ی) Sep-۱۴



شکل ۳- تصاویر RGB از تاریخ‌های مختلف سری زمانی P3، (الف) May-۱۴، (ب) May-۲۰، (ج) Jun-۴، (د) Jun-۲۸، (ه) Jul-۵، (و) Jul-۲۱، (ز) Aug-۱۹، (ح) Aug-۲۹، (ط) Sep-۹، (ی) Sep-۱۴

جدول ۴- نتایج حاصل از طبقه‌بندی دو سری زمانی بر حسب شاخص‌های مختلف

سری زمانی	شاخص دقت	الگوریتم استفاده شده				
		الحاق داده	MKL-sum	SimpleMKL	KA-MKL	RMKL
P1	دقت کلی (%)	۶۴/۶۶	۷۹/۲۲	۷۹/۴۷	۸۰/۳۹	۷۹/۹۰
	شاخص Kappa	۰/۵۶	۰/۷۳	۰/۷۴	۰/۷۴	۰/۷۸
	میانگین دقت کلاس‌ها (%)	۶۲/۴۲	۷۴/۷۱	۷۳/۱۳	۷۵/۷۷	۷۵/۲۶
	زمان (ثانیه)	—	۰	۵/۳	۸/۳	۰/۳
	۳/۸					
P2	دقت کلی (%)	۸۱/۲۴	۸۶/۹۳	۸۷/۱۵	۸۶/۱۸	۸۶/۷۹
	شاخص Kappa	۰/۷۶	۰/۸۳	۰/۸۳	۰/۸۲	۰/۸۳
	میانگین دقت کلاس‌ها (%)	۸۳/۳۰	۸۷/۴۲	۸۷/۴۲	۸۶/۹۵	۸۷/۳۳
	زمان (ثانیه)	—	۰	۱۳۴/۲	۱۲/۱۱	۰/۳
	۳۱/۳					

کرنل‌هایی که حاوی اطلاعات مفید برای طبقه‌بندی نیستند. حساس است. پس نتایج مطلوب این الگوریتم در اینجا را می‌توان به دلیل عدم وجود چنین کرنلی در بین کرنل‌های مینا دانست.

در مقایسه زمان محاسباتی الگوریتم‌های MKL، بجز الگوریتم MKL-sum که هیچ زمانی برای محاسبات نیاز ندارد، الگوریتم SimpleMKL بیشترین و الگوریتم RMKL کمترین زمان محاسباتی را داشتند. الگوریتم GMKL و KA-MKL نیز زمان محاسباتی متوسطی داشتند.

در مورد الگوریتم RMKL ذکر این نکته ضروری بنظر می‌رسد که گرچه این الگوریتم (و سایر الگوریتم‌های MKL زیرفضا) گاهاً دقت‌های قابل قبولی را ارائه می‌کنند. اما این روش‌ها فاقد زمینه‌ی نظری مناسب هستند. چرا که در هیچ یک از مقالاتی که از این الگوریتم‌ها استفاده شده است، توضیح قابل قبولی برای دلیل تبدیل کرنل‌های مینا به صورت بردار ارائه نشده است. ثانیاً بر خلاف ادعای ارائه‌دهندگان این روش‌ها، ضرایب تبدیل زیرفضاهای استخراج شده، همیشه مثبت نیست، که استفاده از این ضرایب باعث می‌شود که کرنل ترکیبی حاصل، مثبت-معین نباشد.

برای مقایسه بهتر کارایی الگوریتم‌ها در طبقه‌بندی هر محصول، میانگین دقت‌های محصولات (در این مقاله، درصد نمونه‌های درست طبقه‌بندی هر کلاس به عنوان دقت آن کلاس در نظر گرفته شده است.) در جدول (۴) گزارش شده است. همانطور که از این شاخص نیز مشخص است، الگوریتم GMKL به دقت بالاتری در طبقه‌بندی هر کلاس رسیده است. نقشه طبقه‌بندی حاصل از روش پیشنهادی و روش الحاق داده در شکل (۴) و شکل (۵) نشان داده شده است.

با توجه به نتایج گزارش شده در این جدول می‌توان به راحتی مشاهده نمود که تمامی الگوریتم‌های MKL دقت بسیار بالاتری از روش الحاق داده‌ها نشان داده‌اند. دلیل این امر را می‌توان ناتوانی روش الحاق داده‌ها در استخراج اطلاعات از داده‌های ناهمگون سری زمانی دانست. همان‌طور که نتایج نشان می‌دهد، الگوریتم GMKL برای طبقه‌بندی سری زمانی P1 دقت بسیار بالاتری از بقیه روش‌های MKL مورد مقایسه داشته است. این در حالی است که در طبقه‌بندی سری زمانی P2، تنها شاهد افزایش جزئی دقت با استفاده از این الگوریتم هستیم. برای نمونه با در نظر گرفتن دقت کلی، الگوریتم GMKL برای سری زمانی P1 به دقت ۸۳/۳۲ درصد رسید که این دقت تقریباً سه درصد بیشتر از الگوریتم KA-MKL است که رده دوم بالاترین دقت را داشته است. اما برای سری زمانی P2، مقایسه دقت الگوریتم GMKL با دقت دومین الگوریتم (SimpleMKL) تنها ۰/۱ درصد افزایش نشان داده است. در توضیح این پدیده می‌توان عنوان کرد که همان‌طور که از شکل (۲) مشخص است، سری زمانی P1 به دلیل وجود ابر و سایه ابر در سه تاریخ، سطح نویز بسیار بالاتری را در مقایسه با سری زمانی P2 دارد. پس می‌توان افزایش دقت الگوریتم GMKL در سری زمانی P1 را نشان دهنده حساسیت کمتر این الگوریتم به نویز موجود در داده دانست.

از دیگر نکات قابل توجه این نتایج، دقت بالای الگوریتم MKL-sum در طبقه‌بندی داده‌های سری زمانی است. علاوه بر این، این الگوریتم به دلیل استفاده از حاصل جمع کرنل‌های مینا به عنوان کرنل ترکیبی، نیاز به بهینه‌سازی ندارد و بسیار سریع قابل اجراست. با این وجود، این الگوریتم بسیار به وجود کرنل‌های ضعیف

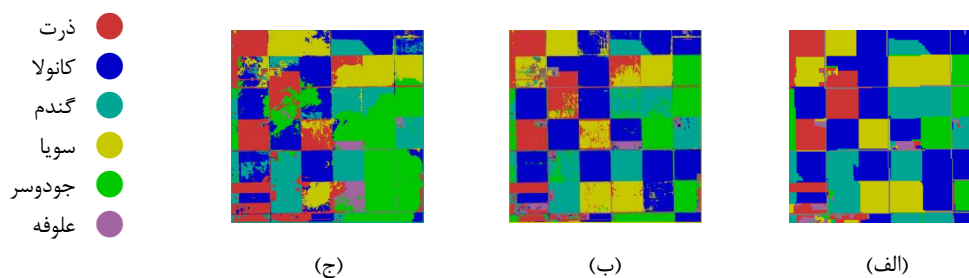
۴- نتیجه گیری

داده‌های سری زمانی چندمتغیره قادر به ارائه اطلاعات بسیار ارزشمندی برای طبقه‌بندی و شناسایی محصولات کشاورزی هستند. با این وجود طبقه‌بندی آنها به دلیل ویژگی‌های خاص این نوع سری زمانی با مشکلات بسیاری روبرو است. در این مقاله، روشی برای طبقه‌بندی سری‌های زمانی چندمتغیره با استفاده از الگوریتم‌های یادگیری چند کرنلی ارائه شده است. اگرچه تمام الگوریتم‌های MKL در چارچوب روش پیشنهادی قابل استفاده هستند، اما در اینجا الگوریتم یادگیری چندکرنلی تعمیم یافته (GMKL) برای این منظور معرفی شده است. این الگوریتم یکی از موفق‌ترین الگوریتم‌های MKL است که تا کنون برای طبقه‌بندی سری‌های زمانی بررسی نشده است. در این مقاله همچنین علاوه بر الگوریتم GMKL، کارائی چندین

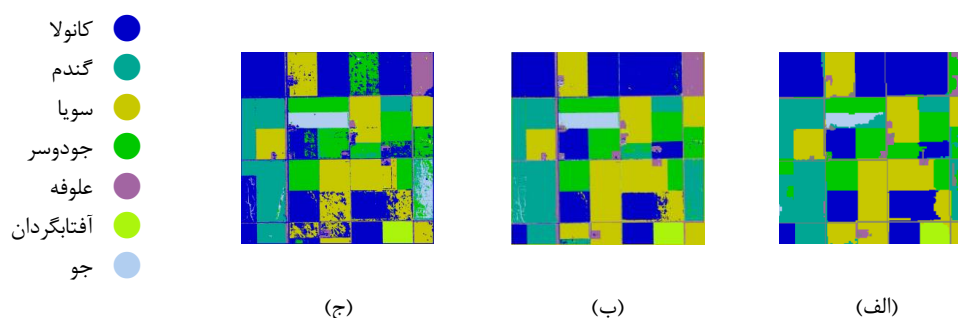
الگوریتم MKL دیگر نیز برای طبقه‌بندی سری‌های زمانی چندمتغیره ارزیابی و با کارائی GMKL مقایسه شده است. نتایج طبقه‌بندی دو سری زمانی برای تشخیص محصولات کشاورزی، نشان داد که روش پیشنهادی برای طبقه‌بندی سری‌های زمانی بسیار موثرتر از روش الحاق داده است. علاوه بر این، بدست آوردن بالاترین دقت طبقه‌بندی با استفاده از الگوریتم GMKL در هر دو سری زمانی نشانگر توانایی بالاتر این الگوریتم در مقایسه با سایر الگوریتم‌های MKL است. این الگوریتم همچنین دقت بالاتری در طبقه‌بندی داده‌های نویزی (به دلیل وجود ابر) داشت.

سپاسگزاری

نویسندگان از خانم دکتر Jiali Shang از موسسه Agri-food کشور کانادا، بخاطر فراهم آوردن امکان استفاده از داده‌های سری زمانی این مقاله کمال تشکر را دارند.



شکل ۴- نقشه‌های طبقه‌بندی حاصل از الگوریتم‌های مختلف برای سری زمانی P1، الف) نقشه واقعیت زمینی، ب) نقشه حاصل از روش پیشنهادی با استفاده از الگوریتم GMKL، ج) نقشه حاصل از روش الحاق داده‌ها



شکل ۵- نقشه‌های طبقه‌بندی حاصل از الگوریتم‌های مختلف برای سری زمانی P2، الف) نقشه واقعیت زمینی، ب) نقشه حاصل از روش پیشنهادی با استفاده از الگوریتم GMKL، ج) نقشه حاصل از روش الحاق داده‌ها

مراجع

- [1] Q. Li, X. Cao, K. Jia, M. Zhang, and Q. Dong, "Crop type identification by integration of high-spatial resolution multispectral data with features extracted from coarse-resolution time-series vegetation index data," International Journal of Remote Sensing, vol. 35, pp. 6076-6088, 2014.
- [2] V. Simonneau, B. Duchemin, D. Helson, S. Er-Raki, A. Oliso, and A. G. Chehbouni, "The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco," International Journal of Remote Sensing, vol. 29, pp. 95-116, 2008/01/01 2008.

- [3] D. K. Bolton and M. A. Friedl, "Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics," *Agricultural and Forest Meteorology*, vol. 173, pp. 74-84, 2013.
- [4] A. Kross, H. McNairn, D. Lapen, M. Sunohara, and C. Champagne, "Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops," *International Journal of Applied Earth Observation and Geoinformation*, vol. 34, pp. 235-248, 2015.
- [5] A. Verhegghen, S. Bontemps, and P. Defourny, "A global NDVI and EVI reference data set for land-surface phenology using 13 years of daily SPOT-VEGETATION observations," *International Journal of Remote Sensing*, vol. 35, pp. 2440-2471, 2014.
- [6] P. Jonsson and L. Eklundh, "TIMESAT - a program for analyzing time-series of satellite sensor data," *Computers & Geosciences*, vol. 30, pp. 833-845, 2004.
- [7] M. G. Baydogan and G. Runger, "Learning a symbolic representation for multivariate time series classification," *Data Mining and Knowledge Discovery*, vol. 29, pp. 400-422, 2015.
- [8] R. Adhikari and R. Agrawal, "An introductory study on time series modeling and forecasting," *arXiv preprint arXiv:1302.6613*, 2013.
- [9] R. Geerken, B. Zaitchik, and J. Evans, "Classifying rangeland vegetation type and coverage from NDVI time series using Fourier Filtered Cycle Similarity," *International Journal of Remote Sensing*, vol. 26, pp. 5535-5554, 2005.
- [10] J. Tigges, T. Lakes, and P. Hostert, "Urban vegetation classification: Benefits of multitemporal RapidEye satellite data," *Remote Sensing of Environment*, vol. 136, pp. 66-75, 2013.
- [11] J. P. Spruce, S. Sader, R. E. Ryan, J. Smoot, P. Kuper, K. Ross, et al., "Assessment of MODIS NDVI time series data products for detecting forest defoliation by gypsy moth outbreaks," *Remote sensing of environment*, vol. 115, pp. 427-437, 2011.
- [12] S. Jamali, J. Seaquist, L. Eklundh, and J. Ardö, "Automated mapping of vegetation trends with polynomials using NDVI imagery over the Sahel," *Remote Sensing of Environment*, vol. 141, pp. 79-89, 2014.
- [13] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE transactions on information theory*, vol. 14, pp. 55-63, 1968.
- [14] M. Gönen and E. Alpaydın, "Multiple kernel learning algorithms," *Journal of Machine Learning Research*, vol. 12, pp. 2211-2268, 2011.
- [15] M. Varma and B. R. Babu, "More generality in efficient multiple kernel learning," in *Proceedings of the 26th Annual International Conference on Machine Learning*, 2009, pp. 1065-1072.
- [16] S. Niazmardi, B. Demir, L. Bruzzone, A. Safari, and S. Homayouni, "A comparative study on Multiple Kernel Learning for remote sensing image classification," in *Geoscience and Remote Sensing Symposium (IGARSS)*, 2016 IEEE International, 2016, pp. 1512-1515.
- [17] J. Milgram, M. Cheriet, and R. Sabourin, "'One Against One" or "One Against All": Which One is Better for Handwriting Recognition with SVMs?," in *Tenth International Workshop on Frontiers in Handwriting Recognition*, 2006.
- [18] S. S. Bucak, R. Jin, and A. K. Jain, "Multiple kernel learning for visual object recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, pp. 1354-1369, 2014.
- [19] M. E. Abbasnejad, D. Ramachandram, and R. Mandava, "A survey of the state of the art in learning the kernels," *Knowledge and information systems*, vol. 31, pp. 193-221, 2012.
- [20] G. Camps-Valls, L. Gomez-Chova, J. Muñoz-Marí, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, pp. 93-97, 2006.
- [21] D. Tuia, F. Ratle, A. Pozdnoukhov, and G. Camps-Valls, "Multisource composite kernels for urban-image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, pp. 88-92, 2010.
- [22] G. Camps-Valls, L. Gómez-Chova, J. Muñoz-Marí, J. L. Rojo-Álvarez, and M. Martínez-Ramón, "Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, pp. 1822-1835, 2008.
- [23] D. Tuia, G. Camps-Valls, G. Matasci, and M. Kanevski, "Learning relevant image features with multiple-kernel classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, pp. 3780-3791, 2010.
- [24] Y. Gu, C. Wang, D. You, Y. Zhang, S. Wang, and Y. Zhang, "Representative multiple kernel learning for classification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 2852-2865, 2012.

- [25] Q. Wang, Y. Gu, and D. Tuia, "Discriminative Multiple Kernel Learning for Hyperspectral Image Classification," IEEE Transactions on Geoscience and Remote Sensing, vol. PP, pp. 1-16, 2016.
- [26] G. Yanfeng, W. Qingwang, W. Hong, Y. Di, and Z. Ye, "Multiple Kernel Learning via Low-Rank Nonnegative Matrix Factorization for Classification of Hyperspectral Imagery," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 8, pp. 2739-2751, 2015.
- [27] Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, pp. 1798-1828, 2013.
- [28] S. Niazmardi, S. Homayouni, and A. Safari, "An Improved FCM Algorithm Based on SVDD for Unsupervised Hyperspectral Data Classification," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 6, pp. 831-839, 2012.
- [29] A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet, "SimpleMKL," Journal of Machine Learning Research, vol. 9, pp. 2491-2521, 2008.