

# کاهش شکاف مفهومی در بازیابی تصویر با رویکرد بهبود اثر تلفیق در یادگیری منیفلد

محمدعلی زارع چاهوکی<sup>۱</sup>

## چکیده

بازیابی تصویر مبتنی بر محتوی، به معنای بازیابی تصاویر با استفاده از ویژگی‌های سطح پایین همچون رنگ، بافت و شکل می‌باشد. در این نوع بازیابی، شکاف معنایی به معنای اختلاف در تفسیر تصاویر، بین انسان و الگوریتم کامپیوتری می‌باشد. در این حوزه، نگاشت غیرصحیح ویژگی‌های سطح پایین تصویر به معنای سطح بالا، سبب می‌شود تا شکاف معنایی افزایش یابد. در بازیابی تصویر، چنانچه با تغییر در بافت، رنگ و یا شکل تصویر، از نظر انسان معنای تصویر تغییری نکند، ولی به دلیل تغییر غیر پیوسته در بردار ویژگی‌های سطح پایین تصویر، بازیابی تصاویر مشابه در تمامی موارد به درستی انجام نمی‌پذیرد. در این مقاله جهت همسوسازی نسبت تغییرات در فضای ویژگی متناسب با فضای معنا، از رویکرد کاهش ابعاد بردارهای ویژگی به صورت غیر خطی استفاده می‌شود. کاهش ابعاد بردارهای ویژگی به صورت غیرخطی که یادگیری منیفلد نیز نامیده می‌شود به معنای جستجوی ساختارهایی با ابعاد کم است که به صورت ذاتی و غیرخطی در مشاهدات با ابعاد بالا وجود دارد. نوآوری اصلی ارائه شده در این مقاله، استخراج یک فضای ویژگی از چند فضای ویژگی می‌باشد که با دو روش ارائه شده، اثر منفی نویز در دقت یادگیری منیفلد کاهش می‌یابد. در ارزیابی دو روش پیشنهادی، از دادگان‌های بخش B از MPEG-7 و Fish استفاده شده است که نتایج تجربی بیانگر موثر بودن روش‌های پیشنهادی می‌باشد.

## کلید واژه‌ها

بازیابی تصویر مبتنی بر محتوی، بازیابی شکل، استخراج غیرخطی ویژگی‌ها، کاهش ویژگی‌ها، یادگیری منیفلد

## ۱ مقدمه

سیگنال‌های نوری را دریافت می‌کنند، مغز انسان، چنانچه تصویر را قبلاً دیده و یا راجع به آن شنیده باشد، با سرعت ادراک صحیحی از آن پیدا می‌کند و می‌تواند با موضوعات قبلی که در خاطر دارد آن را مرتبط سازد. امروزه با گسترش داده‌ها در انواع مختلف و به خصوص در انواع چندرسانه‌ای، این نیاز که بتوانیم با دقت و سرعت، این نوع داده‌ها را بازیابی نماییم امری مهم می‌باشد. در فضای اینترنت، موتورهای جستجو با روش‌های مختلف جستجو در متن و تشخیص درجه ارتباط صفحات با اطلاعات آورده شده در پرس‌وجو، می‌توانند داده‌های متنی را با دقت و سرعت مطلوب بازیابی نمایند. چنانچه تصویری به عنوان پرس‌وجو به موتور جستجوگر داده شود و تصاویر مرتبط با تصویر

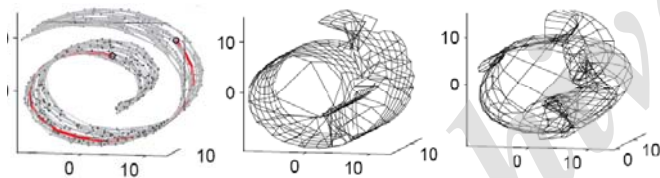
کامپیوترهای امروزی در مقایسه با انسان از سرعت بسیار بیشتری در پردازش‌های ریاضی برخوردار هستند، ولی فعالیت‌های بازشناسی و تشخیص که انسان در برخورد با تصویر و صوت دارد در بسیاری از موارد در دانش محاسبات کامپیوتری، دور از دسترس می‌باشد. هر بار که چشم انسان گشوده می‌شود و حسگرهای بینایی

این مقاله در تیرماه سال ۱۳۹۲ دریافت، در اردیبهشت‌ماه سال ۱۳۹۳ بازنگری و در تیرماه سال ۱۳۹۳ پذیرفته شد.

<sup>۱</sup> استادیار دانشکده مهندسی برق و کامپیوتر، دانشگاه یزد  
رایانامه: [chahooki@yazd.ac.ir](mailto:chahooki@yazd.ac.ir)

می‌باشد، بهبود یابد. در روش‌هایی که تاکنون برای یادگیری منیفلد ارائه گردیده است، توجهی به توصیف نمونه‌ها در فضاهای مختلف ویژگی نشده است. دو روشی که در این مقاله ارائه می‌گردد با بهره‌گیری از چند فضای ویژگی، اثر نویز را در دقت یادگیری منیفلد کاهش می‌دهد. رویکرد ارائه شده برای کاهش حساسیت به نویز در یادگیری منیفلد برای نمونه‌هایی که در چند فضای ویژگی توصیف می‌شوند، نوآوری اصلی این مقاله می‌باشد. نگاشت چندگانه ارائه شده در این پژوهش، راهکاری برای استخراج یک فضا از چند فضای ویژگی می‌باشد.

دو روش  $FoGD^1$  و  $FoD^2$  که برای استخراج فضای ویژگی به صورت غیر خطی پیشنهاد می‌شوند مبتنی بر رویکرد استفاده شده در الگوریتم ایزومپ می‌باشند [۶]. این الگوریتم روشی بدون نظارت در یادگیری منیفلد می‌باشد که رویکرد آن مبتنی بر فاصله بین نمونه‌ها در گراف می‌باشد. استفاده از این الگوریتم در داده‌های واقعی همانند تصاویری که در شرح‌گذاری استفاده می‌شوند، محدودیت‌هایی را ایجاد می‌کند. در روش ایزومپ ساخت گراف همسایگی از اهمیت بالایی برخوردار است. همانگونه که در شکل ۱ آمده است، این گراف پایه محاسبه فاصله کوتاه‌ترین مسیر بین نمونه‌ها می‌باشد که در ایزومپ "کوتاه‌ترین مسیر در سطح"<sup>۷</sup> به آن گفته می‌شود، می‌باشد. باتوجه به شکل ۱، چنانچه گراف همسایگی به درستی ایجاد نگردد، کوتاه‌ترین مسیر بین نمونه‌ها نمی‌تواند به خوبی بیان کنند کوتاه‌ترین مسیر در سطح بین آنها باشد.



(الف) (ب) (ج)

شکل ۱ اهمیت ساخت گراف همسایگی در محاسبه کوتاه‌ترین مسیر در سطح بین نمونه‌ها. در (الف) و (ب) گراف همسایگی به درستی ساخته نشده است بنابراین در این دو گراف مشابه (ج) کوتاه‌ترین مسیر در سطح بین نمونه‌ها با محاسبه کوتاه‌ترین مسیر بین نمونه‌ها بدست نمی‌آید [۶].

علاوه بر ساخت گراف همسایگی، که اهمیت بالایی در محاسبه کوتاه‌ترین مسیر در سطح بین نمونه‌ها دارد، هر چه فاصله دویه دوی نمونه‌ها به فاصله معنایی آنها نزدیکتر باشد، کوتاه‌ترین مسیر بین نمونه‌ها که برآیند جمع فاصله بین دویه‌دوی نمونه‌ها است، به کوتاه‌ترین مسیر در سطح بین نمونه‌ها نزدیکتر می‌شود. بنابراین واژه‌ای تحت عنوان "داده‌های نویزی" در یادگیری منیفلد مبتنی بر فاصله را در این مقاله به صورت زیر تعریف می‌کنیم:

پرس‌وجو تنها با توجه اطلاعات تصویری، بازیابی گردند در پژوهش‌ها تحت عنوان "بازیابی مبتنی بر محتوی"<sup>۱</sup> (CBIR) نام برده می‌شود. در پژوهش‌هایی که در حوزه CBIR انجام می‌گیرند سعی بر آن است تا با استخراج ویژگی‌های رنگ، بافت و شکل تصویر، در بازیابی تصاویری را به کاربر نشان دهند که بیشترین ارتباط را با تصویر پرس و جو داشته باشند. ادراکی که الگوریتم کامپیوتری از تصویر پیدا می‌کند تنها بر اساس ویژگی‌های استخراج شده می‌باشد و بازیابی را نیز بر آن مبنا انجام می‌دهد. بنابراین نتیجه بازیابی با نتیجه مورد انتظار کاربر انسانی لزوماً همخوانی بالایی ندارد. این مطلب که ادراک انسان از تصویر با ادراک کامپیوتر از تصویر تفاوت دارد در پژوهش‌ها تحت عنوان "شکاف معنایی"<sup>۲</sup> نام برده می‌شود که الگوریتم‌های مختلف با رویکردهایی نظیر استفاده از بازخورد<sup>۳</sup> کاربر سعی دارند تا این فاصله را کاهش دهند.

رویکردی که در این مقاله جهت حل مسأله نگاشت غیر صحیح ویژگی‌های سطح پایین به معانی سطح بالا دنبال گردیده است در راستای اهمیت فضای ویژگی در بهبود دقت خوشه‌بندی می‌باشد. در این پژوهش جهت همسوسازی فضای ویژگی با فضای معنا، از رویکرد کاهش ابعاد بردارهای ویژگی به صورت غیر خطی استفاده می‌شود. در حوزه کاهش ابعاد بردار ویژگی، کاهش ابعاد به صورت غیرخطی، یادگیری منیفلد<sup>۴</sup> نیز نامیده می‌شود [۴۲]. یادگیری منیفلد به معنای جستجوی ساختارهایی با ابعاد کم است که به صورت ذاتی و غیرخطی در مشاهدات با ابعاد بالا وجود دارند [۱]. مسأله اساسی در استفاده از روش‌های پایه یادگیری منیفلد مربوط به ماهیت بردار ویژگی بدست آمده برای داده‌هایی مانند تصویر نسبت به بردارهای ویژگی داده‌های آزمایشگاهی است که روش‌های پایه یادگیری منیفلد از آنها استفاده کرده‌اند. در این مقاله از این مسأله با عنوان "حساسیت به نویز" نام برده شده است [۲، ۳، ۴، ۵].

فرض روش‌های یادگیری منیفلد آن است که نمونه‌های  $x_i \in R^D$  که در فضای مشاهده در ابعاد  $D$  وجود دارند، با تابعی غیرخطی  $f$  از نمونه‌های اولیه  $\tau_i \in R^d$  با اضافه شدن نویز  $\epsilon_i$  با رابطه  $x_i = f(\tau_i) + \epsilon_i$  بدست آمده‌اند. نویزی که در دادگان روش‌های پایه یادگیری منیفلد [۶، ۷، ۸، ۹ و ۱۰] در نظر گرفته شده است، محدود بوده و از توزیع از پیش تعریف شده‌ای تبعیت می‌کنند ولی این فرض در دادگان تصویر برقرار نمی‌باشد.

هدف ما در این مقاله یافتن نگاشتی است که با اعمال آن بر بردار ویژگی تصاویر، فاصله بین بردارهای ویژگی در فضای جدید به فاصله معنایی بین آنها نزدیکتر گردد. این امر سبب می‌گردد تا نگاشت غیر صحیح ویژگی‌های سطح پایین به معانی سطح بالا که اساسی‌ترین مسأله در روش‌های شرح‌گذاری خودکار تصویر

<sup>1</sup>Content Based Image Retrieval (CBIR)

<sup>2</sup>Semantic gap

<sup>3</sup>Feed back

<sup>4</sup>Manifold learning

<sup>5</sup>Fusion of Dissimilarities (FoD)

<sup>6</sup>Fusion of Geodesic Distances (FoGD)

<sup>7</sup>Geodesic Distance

<sup>۱</sup>RBF استفاده می‌شود. برای این منظور ۴۳۵ ویژگی تصویر پرس جو به ۶۰ ویژگی کاهش داده می‌شود. در [۱۱] نتیجه گیری که بر اساس نتایج تجربی ارائه شده است، معادل بودن نتیجه بازیابی در هر دو رویکرد می‌باشد. بنابراین می‌توان به جای محاسبات زمانگیر کوتاه‌ترین مسیر در هر پرس جو، کوتاه‌ترین فاصله معنایی را در فضای کاهش بعد داده شده بر اساس فاصله اقلیدسی انجام داد.

در [۱۲]، علاوه بر بازخورد ربط، از بازخورد بی‌ربط که توسط کاربر داده می‌شود، استفاده می‌شود. برای بیان ارتباط بین داده‌ها از سه گراف استفاده می‌شود. در اولین گراف، داده‌ها با توجه به فاصله اقلیدسی شان به هم متصل می‌شوند. وزنی که روی یال‌ها قرار می‌گیرد بیان کننده این فاصله می‌باشد. در گراف دوم، چنانچه دو تصویر در بازخوردها به صورت مرتبط شناخته شده باشند در این صورت عدد ۱ روی یالی که آنها را به هم متصل می‌کند قرار می‌گیرد. در گراف سوم که برای تصاویر غیر مرتبط می‌باشد، عدد ۱- روی یالی که آنها را به هم وصل می‌کند قرار می‌گیرد. کاهش ویژگی با نگاشت ویژه لاپلاسی انجام می‌شود ولی ماتریس لاپلاسی به جای اینکه تنها از یک گراف بدست آید از ترکیب هر سه گراف حاصل می‌شود.

در [۱۳] برای هر داده  $x_i$ ، دو نوع همسایگی تعریف شده است. همسایگی‌هایی که با  $x_i$  دارای برچسب غیر یکسانی هستند (قبلاً به صورت غیر مرتبط مشخص شده اند) به مجموعه  $N_b(x_i)$  و همسایگی‌هایی که شرط قبلی را نداشته باشند به مجموعه  $N_w(x_i)$  افزوده می‌شوند. با توجه به این دو نوع همسایگی، دو گراف ایجاد می‌شود. در گراف  $G_b$  داده‌هایی به هم وصل می‌شوند که در همسایگی  $N_b$  قرار دارند و در گراف  $G_w$  داده‌هایی به هم وصل می‌شوند که در همسایگی  $N_w$  قرار گرفته‌اند. در ادامه کار، مسأله ای که می‌بایست حل شود نگاشت دو گراف  $G_b$  و  $G_w$  به صورت یک خط می‌باشد و این خط می‌بایست به گونه ای حاصل شود که نقاطی از  $G_b$  که به هم متصل می‌باشند، تا حد ممکن از هم دور باشند و نقاط  $G_w$  تا حد ممکن به هم نزدیک باشند. بنابراین تابع هدفی تعریف می‌شود و بردارهایی با ابعاد کاهش یافته  $d$  که این تابع هدف را بهینه می‌کنند، جواب‌های فرایند کاهش ویژگی می‌باشند.

### ۳ تلفیق فضای ویژگی‌ها در یادگیری منیفلد

نتایج ارائه شده توسط روش‌های پایه یادگیری منیفلد بر روی داده‌های بدون نویز (مانند رول سوئیسی) ارائه گردیده است. بردارهای ویژگی که از داده‌هایی مانند تصویر با روش‌های مختلف به دست می‌آیند، داده‌های نویزی می‌باشند. این بدان معناست که نویز موجود در بردارهای ویژگی ( $\epsilon_i$ ) در کنترل ما نمی‌باشد و

چنانچه بردارهای ویژگی توصیف کننده نمونه‌ها (داده‌ها) در فضای ویژگی به گونه‌ای نباشند که پس از ساخت گراف همسایگی، محاسبه کوتاه‌ترین مسیر بین نمونه‌ها بیان کننده "کوتاه‌ترین مسیر در سطح" بین نمونه‌ها باشد، گفته می‌شود که داده‌ها نویزی می‌باشند. در ادامه این مقاله، در بخش ۲ مروری بر روش‌های بازیابی تصویر آورده می‌شود که در آنها از رویکرد یادگیری منیفلد در راستای کاهش شکاف مفهومی استفاده شده است. در بخش ۳ به بیان دو روش پیشنهادی می‌پردازیم. پیاده‌سازی و ارزیابی نتایج تجربی در بخش ۴ آورده شده است. نتیجه‌گیری و پیشنهادهایی برای کارهای آینده نیز در بخش ۵ آورده شده است.

## ۲ استخراج غیرخطی فضای ویژگی در بازیابی تصویر

در سال‌های اخیر استفاده از رویکرد یادگیری منیفلد در کاربردهای مختلف مورد توجه قرار گرفته است. در بازیابی تصویر مبتنی بر محتوی، کاهش ابعاد بردار ویژگی به صورت غیرخطی در [۱۱]، [۱۲] و [۱۳] آورده شده است. استفاده از این رویکرد اولین بار در [۱۱] ارائه گردید. در این تحقیق، داده‌ها کل تصویر هستند و ۴۳۵ ویژگی به دست آمده از رنگ و بافت تصویر در هر تصویر به ۶۰ ویژگی کاهش داده شده است. برای وزن دهی یال‌های گرافی که ارتباط داده‌ها را بیان می‌کند، هم از فاصله اقلیدسی بردار ویژگی تصاویر و هم از بازخورد ربطی که توسط کاربر داده می‌شود، استفاده می‌شود. چنانچه فاصله اقلیدسی دو تصویر از مقدار کوچکی چون  $\epsilon$  کمتر باشد، در این صورت دو تصویر با یالی که وزن آن برابر با فاصله اقلیدسی دو تصویر است مرتبط می‌شوند. همچنین هنگامی که دو تصویر در بازخورد کاربر توسط سیستم مرتبط شناخته می‌شود، وزن یالی که این دو تصویر را به هم وصل می‌کند،  $\epsilon$  در نظر گرفته می‌شود.

در [۱۱]، با دو رویکرد به مسأله بازیابی پرداخته است و در انتها نتیجه‌گیری که بر اساس نتایج تجربی ارائه نموده‌است، معادل بودن این دو رویکرد می‌باشد. در رویکرد اول، با آمدن هر تصویر پرس‌وجو، تصویر مربوطه به گراف اضافه می‌شود و فاصله اقلیدسی تصویر پرس‌وجو با تمامی تصاویر گراف محاسبه می‌شود. در این رویکرد برای بازیابی تصاویری که دارای کمترین فاصله معنایی با تصویر پرس‌جو هستند، کوتاه‌ترین مسیر بین تصویر پرس‌جو با تمامی تصاویر محاسبه می‌شود و نتیجه بازیابی، تصاویری هستند که کمترین کوتاه‌ترین مسیر را با تصویر پرس‌وجو دارند. در رویکرد دوم با روش نگاشت ویژه لاپلاسی [۸]، که از روش‌های پایه در یادگیری منیفلد است، ابعاد بردار ویژگی هر تصویر کاهش داده می‌شود. در این رویکرد تمامی محاسبات برای یافتن شباهت با فاصله اقلیدسی در فضای کاهش بعد داده شده انجام می‌شود. در رویکرد دوم برای یادگیری چگونگی کاهش بعد از شبکه عصبی

<sup>۱</sup>Radial Basis Function (RBF)

"نگاشت چندگانه بر مجموعه‌ای از فضاها و ویژگی، نگاشتی است که یک فضای ویژگی را از چند فضای ویژگی با ابعاد مختلف استخراج می‌کند."

رابطه ۱ بیان کننده این نگاشت می‌باشد:

$$f: (R^{D_1}, R^{D_2}, \dots, R^{D_M}) \rightarrow R^d \quad (1)$$

در این رابطه،  $M$  بیانگر تعداد فضاها و ویژگی،  $D_i$  بیانگر ابعاد بردار ویژگی (تعداد ویژگی‌های هر نمونه) در فضای  $i$ ام و  $d$  بیانگر ابعاد بردار ویژگی در فضای جدید می‌باشد.

در هر کدام از فضاها، فاصله اقلیدسی بین نمونه‌ها تا حدی بیان کننده فاصله معنایی بین نمونه‌ها می‌باشد. هر چند فاصله اقلیدسی در هیچکدام از فضاها، بیان کننده فاصله معنایی بین نمونه‌ها نمی‌باشد. با دو روشی که در این مقاله ارائه می‌شود هر نمونه با روشی بدون نظارت در یک فضا توصیف می‌شود که این فضا را فضای منیفلد می‌نامیم. نتایج بدست آمده از پیاده‌سازی و ارزیابی در بخش بعد بیانگر افزایش دقت در بازیابی تصویر و در نتیجه کاهش شکاف معنایی بین نتایج بازیابی و نظر انسان می‌باشد.

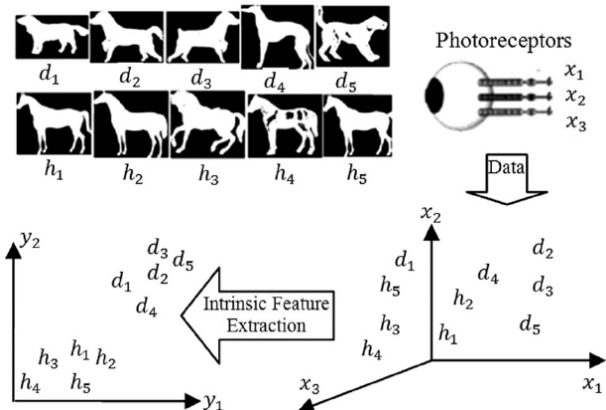
در این بخش دو روش برای یادگیری منیفلد آورده می‌شود. روش اول تنها بر مبنای تلفیق<sup>۱</sup> عدم شباهت<sup>۲</sup>ها در فضاها مختلف ویژگی می‌باشد، درحالی‌که روش دوم علاوه بر این نوع تلفیق از تلفیق<sup>۳</sup> "کوتاه‌ترین مسیر در سطح" بین نمونه‌ها نیز استفاده می‌کند. برای سادگی بیان، روش اول را یادگیری منیفلد مبتنی بر تلفیق عدم شباهت<sup>۳</sup> (FoD) و روش دوم را یادگیری منیفلد مبتنی بر تلفیق کوتاه‌ترین مسیر در سطح<sup>۴</sup> (FoGD) می‌نامیم.

در ادامه این بخش، در ابتدا نحوه تلفیق بردارهای ویژگی و ساخت ماتریس عدم شباهت تلفیقی<sup>۵</sup> که در هر دو روش استفاده می‌شود را در بخش ۳-۱ بیان می‌کنیم. سپس دو روش FoD و FoGD به ترتیب در دو بخش ۳-۲ و ۳-۳ آورده می‌شوند. در بخش ۳-۴ به جمع‌بندی مدل پیشنهادی در یادگیری منیفلد و تاثیر آن در بازیابی تصویر مبتنی بر محتوی می‌پردازیم.

### ۳-۱ تلفیق عدم شباهت

در این بخش به بیان چگونگی ساخت ماتریس عدم شباهت تلفیقی می‌پردازیم. این ماتریس، در روش‌های FoD و FoGD که در دو بخش بعدی به توصیف آنها می‌پردازیم استفاده می‌شود. تلفیق اطلاعات<sup>۶</sup> به معنای ترکیب اطلاعاتی است که از منابع مختلف به دست آمده است تا بتوانیم بر اساس آنها تصمیم‌گیری را بهبود بخشیم [۱۵]. کارهای اولیه که در این زمینه در حوزه پردازش تصویر انجام پذیرفته است مربوط به کاربردهای نظامی می‌باشد که

مشخص نیست که در چه بازه‌ای تغییر می‌کنند. بنابراین استفاده از روش‌های پایه یادگیری منیفلد برای این داده‌ها مناسب نمی‌باشد.



شکل ۱ مثالی از استخراج ویژگی به گونه‌ای که فاصله اقلیدسی بین نمونه‌ها در فضای جدید به فاصله معنایی آنها نزدیکتر می‌باشد [۱۴]

شکل ۲، هدف از روش‌های یادگیری منیفلد را در بازیابی تصویر نشان می‌دهد. ذهن انسان نیز به صورت مشابه و به صورت دائمی در حال حل چنین مسأله‌ای می‌باشد. در هر لحظه بیش از سی هزار عصب دریافت کننده شنوایی و بیش از ده میلیون عصب دریافت کننده بینایی اطلاعات را از محیط دریافت می‌کنند و آنها را به ویژگی‌هایی تبدیل می‌کنند که در فضای معنا قابل درک می‌باشند [۶]. در این شکل، مجموعه نمونه‌ها شامل پنج تصویر سگ و پنج تصویر اسب می‌باشد. هر چند ذهن انسان در فضای معنا بین تصاویر هرکدام از دو خانواده تصویری ارتباط پیوسته‌ای را در نظر می‌گیرد، ولی فاصله اقلیدسی بین نمونه‌ها در فضای ویژگی اولیه به طور کامل بیانگر ارتباط معنایی تصاویر نمی‌باشد. به عبارتی دیگر، تمامی نمونه‌های مربوط به هرکدام از دو نوع حیوان در یک خوشه قرار نمی‌گیرند. در فضای ویژگی جدید که با ابعادی کمتر از ابعاد فضای اولیه ترسیم گردیده است، فاصله اقلیدسی به فاصله معنایی نزدیکتر شده است.

در کاربردهایی مانند بازیابی تصویر، هر تصویر با روش‌های مختلفی می‌تواند توصیف شود. توصیف نمونه‌ها با روش‌های مختلف، بردارهای ویژگی را در فضاها متفاوت ویژگی در اختیار قرار می‌دهد. در روش‌هایی که تاکنون برای یادگیری منیفلد ارائه گردیده است، توجهی به توصیف نمونه‌ها در فضاها مختلف ویژگی نشده است. با این توضیح، دو روشی که در این بخش ارائه می‌گردد با بهره‌گیری از فضاها و ویژگی، اثر نویز را در دقت استخراج ویژگی به صورت غیرخطی کاهش می‌دهد. هدف از ارائه این دو روش آن است تا فاصله اقلیدسی را در فضای منیفلد به فاصله معنایی نزدیک سازیم. نگاشت نمونه‌هایی که توصیف آنها در چند فضای ویژگی صورت پذیرفته است را به یک فضا، نگاشت چندگانه تعریف می‌کنیم که به صورت زیر بیان می‌شود.

<sup>1</sup>Fusion

<sup>2</sup>Dissimilarity

<sup>3</sup>manifold learning based on Fusion of Dissimilarities (FoD)

<sup>4</sup> manifold learning based on Fusion of Geodesic Distances (FoGD)

<sup>5</sup>Fused Dissimilarity Matrix

<sup>6</sup>Information Fusion



شبهات هر نمونه با نمونه‌هایی که در همسایگی نزدیکی با آن قرار دارند، کاهش یابد. به تعبیری دیگر، هر نمونه، نمونه‌هایی را که انتظار می‌رود با هم معنای مشترکی دارند را جذب می‌کند. همچنین  $\theta \times N$  مقدار انتهایی لیست مرتب شده را در عدد بزرگی مانند  $1 > Inf$  ضرب می‌کنیم. بنابراین هر نمونه، نمونه‌هایی را که انتظار می‌رود معنای مشترکی با هم ندارند را دفع می‌کند. سپس مقادیر متناظر یال‌ها را جایگزین می‌کنیم. گراف جدید که در آن مقادیر یال‌های آن نسبت به گراف  $G$  تغییر کرده است،  $\bar{G}$  نامیده می‌شود.

**گام ۳-** به منظور حذف یال‌هایی که در محاسبه کوتاه‌ترین مسیر در سطح بین نمونه‌های گراف  $\bar{G}$  اثر منفی دارند، گراف همسایگی  $\bar{N}\bar{G}$  از گراف  $\bar{G}$  ساخته می‌شود. در این گام نیز از روش ارائه شده در روش ایزومپ استفاده می‌شود. برای این منظور دو گره از  $\bar{N}\bar{G}$  به هم متصل می‌مانند چنانچه یکی از آنها در  $k$  همسایگی نزدیکتر دیگری قرار داشته باشد.

**گام ۴-** محاسبه کوتاه‌ترین مسیر در سطح نیز مشابه روش ایزومپ با محاسبه کوتاه‌ترین مسیر بین دو گره‌ها الگوریتم دایکسترا<sup>۵</sup> انجام می‌شود. این فاصله به شکاف معنایی بین دو گره نزدیکتر می‌باشد. کوتاه‌ترین مسیر در سطح محاسبه شده بین نمونه‌ها را در ماتریس  $FoD$  قرار می‌دهیم.

**گام ۵-** برای رسیدن به نقاطی که فاصله آنها در فضای ویژگی به کوتاه‌ترین مسیر در سطح که در گام قبل محاسبه شد نزدیک باشد از روشی مبتنی بر MDS<sup>۶</sup> استفاده می‌کنیم. روشی که در این گام برای استخراج بردارهای ویژگی استفاده می‌کنیم MDS کلاسیک می‌باشد. ماتریس  $FoD$  که از گام قبل محاسبه گردید دربرگیرنده کوتاه‌ترین مسیر در سطح بین نمونه‌ها می‌باشد. از آنجاییکه این ماتریس متقارن نمی‌باشد بنابراین با مقارنه‌ی مجدد این ماتریس به صورت  $FoD(i, j) = \min(FoD(i, j), FoD(j, i))$ ، ماتریس حاصل متقارن می‌گردد. چنانچه فرض کنیم  $Y = [y_1, y_2, \dots, y_N]^T$  ماتریس نمونه‌ها در فضای جدید (فضای منیفلد) باشد، استخراج ویژگی‌ها در فضای منیفلد با تجزیه ماتریس  $T = YY^T$  با روشی که در ادامه به توضیح آن می‌پردازیم انجام می‌شود. در فضای جدید، فاصله اقلیدسی بین دو نمونه از رابطه زیر به دست می‌آید. در این رابطه به دلیل اینکه انتظار داریم فاصله اقلیدسی در فضای جدید همان کوتاه‌ترین مسیر در سطح بین نمونه‌ها باشد به جای آنمقادیر ماتریس  $FoD$  را جایگذاری می‌کنیم.

$$FoD_{ij}^2 = T_{ii} + T_{jj} - 2T_{ij} \\ T_{ij} = \sum_{k=1}^d y_{ik}y_{jk} \quad (4)$$

در این رابطه،  $d$  ابعاد بردارهای ویژگی در فضای جدید می‌باشد. چنانچه فرض کنیم مرکز ثقل نقاط در فضای جدید در مبدا مختصات قرار دارد، بنابراین می‌توانیم مقادیر ماتریس  $T$  را بر اساس مقادیر ماتریس  $FoD$  به صورت زیر مقدار دهی کنیم [۱۷].

$$T_{ij} = \frac{-1}{2} [FoD_{ij}^2 - FoD_{i,i}^2 - FoD_{j,j}^2 - FoD_{i,i}^2]$$

برای ردیابی خودکار<sup>۱</sup>، بازشناسی<sup>۲</sup> و شناسایی<sup>۳</sup> اهداف و همچنین دیده‌بانی از میدان نبرد<sup>۴</sup> استفاده گردیده است [۱۶].

فرض می‌کنیم  $X = \{x_1, x_2, \dots, x_N\}^T$  مجموعه‌ای از  $N$  تصویر می‌باشد. هر تصویر  $x_i$  در فضای مشاهده با  $M$  بردار ویژگی به صورت  $\{\vec{x}_i^1 \in S_1, \vec{x}_i^2 \in S_2, \dots, \vec{x}_i^M \in S_M\}$  تعریف می‌شود. بنابراین، فرض ما بر این است که هر تصویر با  $M$  روش مختلف توصیف تصویر آورده شده است که هر روش توصیف، بردارهای ویژگی تمامی نمونه‌ها را در فضای متناظر خود تولید می‌کند. در هر فضای  $S_i$ ،  $d_i(\vec{x}_j^i, \vec{x}_k^i)$  را فاصله بین دو نمونه  $x_j$  و  $x_k$  تعریف می‌کنیم.

به دلیل اینکه روش‌های استخراج بردار ویژگی متمایز از هم می‌باشند بنابراین  $d_i$ ها از هم مستقل می‌باشند. پیش از تلفیق، می‌بایست فاصله بین نمونه‌ها در هر فضا نرمال گردد. نرمال سازی در هر فضا به گونه‌ای انجام می‌شود که مجموع فاصله‌های هر نمونه با بقیه نمونه‌ها برابر یک شود.  $\hat{d}_i(\vec{x}_j^i, \vec{x}_k^i)$  را فاصله نرمال بین دو نمونه  $x_j$  و  $x_k$  تعریف می‌کنیم که با رابطه ۲ بدست می‌آید.

$$\hat{d}_i(\vec{x}_j^i, \vec{x}_k^i) = \hat{d}_i(\vec{x}_j^i, \vec{x}_k^i) = \frac{d_i(\vec{x}_j^i, \vec{x}_k^i)}{\sum_{k=1}^N d_i(\vec{x}_j^i, \vec{x}_k^i)} \quad (2)$$

تابع تلفیق عدم شبهات بین دو نمونه  $x_j$  و  $x_k$  را با رابطه ۳ تعریف می‌کنیم.

$$fusion(x_j, x_k) = \frac{1}{M} \sum_{i=1}^M w_i \hat{d}_i(\vec{x}_j^i, \vec{x}_k^i) \quad (3)$$

در این رابطه،  $w_i$  بیانگر میزان اهمیت فضای ویژگی  $S_i$  (روش  $i$ ام توصیف ویژگی) می‌باشد. به صورت پیش فرض این مقدار برای تمامی فضاها یک می‌باشد و در نتایج تجربی چنانچه مقداری دیگر در نظر گرفته شود، نحوه محاسبه آن بیان می‌گردد.

این روش در پنج گام به صورت زیر انجام می‌شود.

**گام ۱-** در این روش تمامی نمونه‌ها در گراف  $G(V, E)$  مدل می‌شوند. در این گراف،  $V$  مجموعه‌ای از  $N$  نمونه می‌باشد و  $E = V \times V$  مجموعه یال‌های این گراف می‌باشد. چنانچه  $x_j$  و  $x_k$  دو نمونه باشند، مقدار نسبت داده شده به یال  $(x_j, x_k) \in E$  را با  $edge_{jk}$  بیان می‌کنیم. مقادیر یال‌ها با مقدار متناظر در ماتریس عدم شبهات تلفیقیبه صورت  $edge_{jk} = D_{jk}$  مقدار دهی اولیه می‌شوند.

**گام ۲-** هدف از این مرحله، اصلاح مقادیر یال‌ها به منظور کاهش حساسیت به نویز در محاسبه کوتاه‌ترین مسیر در سطح داده‌ها می‌باشد. برای این منظور یال‌های متصل به هر گره گراف را بر اساس مقادیرشان به صورت صعودی مرتب می‌کنیم. سپس  $\mu \times N$  مقدار ابتدایی لیست مرتب شده را در عدد کوچکی مانند  $0 < \epsilon < 1$  ضرب می‌کنیم. این ضرب سبب می‌شود تا عدم

<sup>1</sup>automatically tracking

<sup>2</sup>recognizing

<sup>3</sup>identifying

<sup>4</sup>battlefield surveillance

<sup>5</sup>dijkstra

<sup>6</sup>Multidimensional Scaling

نوآوری‌ایکه در این روش نسبت به روش قبل وجود دارد، بهبود اثر نویز در هرفضا به صورت جداگانه می‌باشد. در روش قبل اصلاح اثر نویز، پس از تلفیق عدم شباهت‌ها در گراف  $G$  انجام می‌پذیرفت. از طرف دیگر، اصلاح اثر نویز قبل از تلفیق سبب می‌شود تا از انتشار اثر نویز در فرایند تلفیق جلوگیری شود. بنابراین انتظار می‌رود یادگیری منیفلد با دقت بالاتری انجام شود. در این گام نیز در هر گراف  $G_i$ ، یال‌های متصل به هر گره به صورت صعودی مرتب می‌شوند. بر خلاف روش قبل، مرتب سازی یال‌ها بر اساس وزن یال‌ها در همان گراف انجام نمی‌پذیرد. بلکه یال‌های گراف  $G_i$  بر اساس وزن یال‌های متناظرشان در گراف  $G$  مرتب می‌شوند. این رویکرد سبب می‌شود تا یافتن همسایه‌های نزدیکتر و دورتر در هر فضا، با اطلاعات جامع‌تری که از تلفیق به دست آمده است انجام پذیرد. سپس  $\mu \times N$  مقدار ابتدایی لیست مرتب شده در هر فضا که مرتبط با هر گره می‌باشد را در عدد کوچکی مانند  $\varepsilon$  ضرب می‌کنیم و به همین ترتیب  $\theta \times N$  مقدار انتهایی لیست را در عدد بزرگی مانند  $Inf$  ضرب می‌کنیم. سپس مقادیر بهنگام شده یال‌ها را در  $E_i$  جایگزین می‌کنیم.

**گام ۴-** به منظور حذف یال‌هایی که در محاسبه کوتاه‌ترین مسیر در سطح بین نمونه‌های گراف  $G_i$  در فضای  $s_i$  اثر منفی دارند، گراف همسایگی  $\bar{N}G_i$  از گراف  $G_i$  ساخته می‌شود. در این گام نیز از روش ارائه شده در روش ایزومپ استفاده می‌شود. برای این منظور دو گره از  $\bar{N}G_i$  به هم متصل می‌مانند چنانچه یکی از آنها در  $k$  همسایگی نزدیکتر دیگری قرار داشته باشد.

**گام ۵-** محاسبه کوتاه‌ترین مسیر در سطح بین نمونه‌های گراف  $G_i$  در فضای  $s_i$  نیز مشابه روش ایزومپ با محاسبه کوتاه‌ترین مسیر بین دو گره با الگوریتم دایکسترا انجام می‌شود. این فاصله به شکاف معنایی بین دو گره نزدیکتر می‌باشد. کوتاه‌ترین مسیر در سطح محاسبه شده بین نمونه‌ها را در ماتریس  $GD_i$  قرار می‌دهیم.

**گام ۶-** تلفیق کوتاه‌ترین مسیرهای در سطح که در هر فضا بدست آمده است با رابطه زیر انجام می‌شود.

$$FoGD_{N \times N} = \frac{1}{M} \sum_{i=1}^M GD_i \quad (8)$$

کوتاه‌ترین مسیر در سطح محاسبه شده در هر گراف، نزدیکی بیشتر با شکاف معنایی دارد چراکه از این فاصله در آخرین گام روش FoD برای استخراج ویژگی‌ها استفاده شده است. بنابراین تلفیق در سطح کوتاه‌ترین مسیر در سطح در روش FoGD نسبت به تلفیق در سطح فاصله بین گره‌های گراف در روش FoD، دارای اعتبار بیشتری می‌باشد. این برتری در نتایج تجربی نیز نشان داده شده است.

**گام ۷-** در این گام نیز مشابه آخرین گام در روش FoD، از MDS کلاسیک جهت استخراج  $N$  بردار ویژگی در فضایی با ابعاد  $d$  استفاده می‌کنیم. فاصله اقلیدسی بین بردارهای ویژگی در این فضا، از کوتاه‌ترین مسیر در سطح محاسبه شده در گام قبل پیروی می‌کند. بنابراین فاصله اقلیدسی در منیفلد استخراج شده با روش FoGD نسبت به روش FoD به شکاف معنایی نزدیکتر می‌باشد.

$$FoD_i^2 = \frac{1}{N} \sum_{j=1}^N FoD_{ij}^2$$

$$FoD_j^2 = \frac{1}{N} \sum_{i=1}^N FoD_{ij}^2$$

$$FoD^2 = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N FoD_{ij}^2 \quad (5)$$

از آنجاییکه ماتریس  $T$ ، ماتریسی متقارن با مقادیر حقیقی می‌باشد بنابراین می‌تواند به صورت  $T = U\Lambda U^T$  نوشته شود. در این رابطه، ستون‌های  $U$ ، بردارهای ویژگی ماتریس  $T$  می‌باشد و  $\Lambda$  ماتریسی قطری می‌باشد که مقادیر ویژه ماتریس  $T$ ، مقادیر قطر آن را تشکیل می‌دهند. بنابراین با توجه به رابطه  $T = YY^T$ ، نمونه‌ها در فضای جدید با رابطه  $Y = U\Lambda^{\frac{1}{2}}$  به دست می‌آیند.

ما به دنبال فضای جدید با ابعاد کم می‌باشیم، بنابراین از بردارهای ویژه‌ای استفاده می‌کنیم که متناظر با مقادیر ویژه بزرگتر می‌باشند. ابعاد بردارهای ویژگی در فضای منیفلد،  $d$ ، را از رابطه زیر بدست می‌آوریم.

$$\sum_{i=1}^{d-1} \lambda_i < \alpha \sum_{i=1}^N \lambda_i < \sum_{i=1}^d \lambda_i \quad (6)$$

در رابطه بالا،  $\lambda_i$  مقادیر ویژه ماتریس  $FoD$  بوده و  $0 < \alpha < 1$  مقدار آستانه‌ای برای انتخاب ابعاد بردار ویژگی در فضای جدید می‌باشد. نمونه‌ها در فضای منیفلد با ابعاد  $d$  با رابطه زیر بدست می‌آیند. در این رابطه  $u_i$ ، بردار ویژه متناظر با مقدار ویژه  $\lambda_i$  می‌باشد.

$$Y = [u_1, \dots, u_d] \text{diag} \left( \lambda_1^{\frac{1}{2}}, \dots, \lambda_d^{\frac{1}{2}} \right) = U_d \Lambda_d^{\frac{1}{2}} \quad (7)$$

زمانیکه تمامی مقادیر ویژه، مثبت باشند و یا تعداد کمی از آنها منفی باشند، در این صورت این روش کارایی مناسبی در استخراج ویژگی دارد [۱۷]. بنابراین، چنانچه تعداد زیادی از آنها منفی باشند و یا قدرمطلق مقادیر منفی نسبت به مقادیر مثبت بزرگتر باشد، این روش برای استخراج ویژگی مناسب نمی‌باشد.

### ۲-۳ یادگیری منیفلد مبتنی بر تلفیق کوتاه‌ترین مسیر در سطح (FoGD)

این روش نیز در پنج گام به صورت زیر انجام می‌شود:

**گام ۱-** در گام اول تمامی نمونه‌ها در هرکدام از فضاهای ویژگی  $s_i$ ، در گرافی جداگانه نمایش داده می‌شوند. گراف‌های  $G_i$  می‌باشند. در این گراف‌ها،  $V_i$  مجموعه‌ای از  $N$  نمونه و  $E_i$  مجموعه یال‌ها در هر فضا می‌باشد. چنانچه  $x_j$  و  $x_k$  دو نمونه باشند، مقدار نسبت داده شده به یال  $(x_j, x_k) \in E_i$  را با  $edge_{jk}^i$  بیان می‌کنیم. مقادیر یال‌ها با فاصله نرمال شده بین نمونه‌ها در هر فضا که از رابطه (۲-۳) محاسبه شده‌است، به صورت  $edge_{jk}^i = \hat{d}_i(\vec{x}_j^i, \vec{x}_k^i)$  مقدار دهی اولیه می‌شوند.

**گام ۲-** گراف  $G(V, E)$  همانند گام اول در روش FoD ساخته می‌شود.

**گام ۳-** در این گام نیز همانند روش FoD، نویزی که در مقادیر یال‌های گراف وجود دارند اصلاح می‌شوند.

### ۳-۳ تحلیل عملکرد دو روش پیشنهادی

در این بخش دو روش مبتنی بر حفظ فاصله با رویکرد بدون نظارت برای استخراج ویژگی‌ها به صورت غیرخطی ارائه گردید. این دو روش جهت حل اثر نویز در روش‌های پایه‌ای یادگیری مینفولد ارائه گردیده‌اند. در کاربردهایی مانند بازیابی تصویر، هر نمونه با روش‌های مختلفی توصیف می‌شود. بنابراین دو روش ارائه شده در راستای تلفیق فضاهای مختلف ویژگی و استخراج یک فضای مینفولد از چند فضای ویژگی می‌باشند. **نگاشت چندگانه** که در مدل پیشنهادی آمده است، راهکاری برای استخراج یک فضا از چند فضای ویژگی می‌باشد.

هرچند که در هر کدام از فضاهای ویژگی، فاصله اقلیدسی بین نمونه‌ها تا حدی بیان‌کننده شکاف معنایی بین آنها می‌باشد، ولی فاصله کمتری در فاصله بین نمونه‌ها در فضای واحد مینفولد و فضای معنا وجود دارد. نوآوری اصلی ارائه شده در این مدل، استفاده مؤثر از توصیف نمونه‌ها در فضاهای مختلف ویژگی در راستای استخراج فضای مینفولد می‌باشد. دو تفاوت در روش FoGD نسبت به FoD وجود دارد. تفاوت اول ناشی از رویکرد پیشگیرانه FoGD در هموار سازی اثر نویز می‌باشد. این هموار سازی سبب می‌شود اثر نویز در هر فضا به صورت جداگانه بررسی گردد و از تجمیع آنها در گراف تلفیقی پیشگیری شود. این رویکرد سبب می‌شود تا فاصله بین نمونه‌ها با رویکردی بدون نظارت به شکاف معنایی نزدیکتر گردد. تفاوت دوم به نحوه تلفیق فضاهای ویژگی مرتبط می‌شود. به جای تلفیق گراف‌های بدست آمده پس از هموار سازی اثر نویز، ابتدا کوتاه‌ترین مسیر در سطح بین نمونه‌ها محاسبه می‌شود و سپس تلفیق در سطح کوتاه‌ترین مسیر در سطح بین نمونه‌ها انجام می‌پذیرد. ولی در روش FoD تلفیق در سطح فاصله اقلیدسی بین نمونه‌ها انجام می‌شود و سپس کوتاه‌ترین مسیر در سطح بین نمونه‌ها محاسبه می‌شود. به دلیل نزدیکی کوتاه‌ترین مسیر در سطح به شکاف معنایی، بنابراین تلفیق در سطح کوتاه‌ترین مسیر در سطح مؤثر تر از تلفیق در سطح فاصله اقلیدسی می‌باشد. نتایج بدست آمده از پیاده‌سازی و ارزیابی این دو روش در فصل بعد بیانگر افزایش دقت در بازیابی تصویر و در نتیجه کاهش شکاف معنایی بین نتایج بازیابی و نظر انسان می‌باشد. این نتایج همچنین مؤثر تر بودن روش FoGD نسبت به FoD را نیز نشان می‌دهد.

### ۴ پیاده‌سازی و ارزیابی نتایج تجربی

در این بخش جهت ارزیابی مؤثر بودن دو روش FoD و FoGD، نتایج تجربی آنها بر روی دو دادگان بخش B از MPEG-7 و Fish آورده می‌شود. در ادامه این بخش، ابتدا در بخش ۴-۱ به معرفی دو دادگان ذکر شده می‌پردازیم. سپس در بخش ۴-۲، نتایج تجربی و ارزیابی دو روش پیشنهادی آورده می‌شود.

### ۴-۱ دادگان‌های استفاده شده در ارزیابی

در این بخش ابتدا دادگان بخش B از MPEG-7 و سپس دادگان Fish معرفی می‌شود.

#### الف) دادگان بخش B از MPEG-7

این دادگان شامل ۱۴۰۰ تصویر دودویی می‌باشد که در ۷۰ کلاس، طبقه بندی شده‌اند [۱۸]. بنابراین هر کلاس شامل ۲۰ تصویر می‌باشد. این دادگان از پرستفاده‌ترین دادگان‌ها در بازیابی شکل می‌باشد و روش‌های متعددی نتایج خود را در این دادگان ارائه داده‌اند.

ارزیابی در این دادگان، با رویکرد بدون نظارت، با آزمون<sup>۱</sup> BEP انجام می‌شود. در این آزمون هرکدام از ۱۴۰۰ تصویر به عنوان پرس‌وجو داده می‌شود ( $x_i$ ) و از ۴۰ تصویری که بیشترین شباهت را با آن دارند، تعداد تصاویری که در همان کلاس تصویر پرس‌وجو قرار دارند، ثبت می‌شود ( $bep_i$ ). نتیجه این آزمون طبق رابطه ۹ بدست می‌آید.

$$BEP = \frac{\sum_{i=1}^{1400} bep_i}{1400 \times 20} \times 100 \quad (9)$$

همچنین جهت مقایسه رویکردهای مختلف در پیاده‌سازی، از نمودار "دقت<sup>۲</sup> و فراخوان<sup>۳</sup>" نیز استفاده می‌شود. برای این منظور، مشابه آزمون BEP، هرکدام از ۱۴۰۰ تصویر به عنوان پرس‌وجو داده می‌شود ( $x_i$ ) و از  $t$  تصویری که بیشترین شباهت را با آن دارند، تعداد تصاویری که در همان کلاس تصویر پرس‌وجو قرار دارند، ثبت می‌شود ( $c_i$ ). دقت و فراخوان مطابق رابطه ۱۰ بدست می‌آیند.

$$Precision = \frac{\sum_{i=1}^{1400} c_i}{1400 \times t}$$

$$Recall = \frac{\sum_{i=1}^{1400} c_i}{1400 \times 20} \quad (10)$$

#### ب) دادگان Fish

این دادگان شامل ۱۱۰۰ فایل متنی است که شامل مختصات نقاط لبه حیوانات دریایی می‌باشند [۱۹]. تعداد نقاط لبه‌ای هر شکل از ۲۵۶ تا ۱۶۵۳ نقطه، متغیر می‌باشند. در دادگان اصلی، حیوانات به کلاس‌های مشخصی تخصیص داده نشده‌اند. در ارزیابی، از دسته‌بندی که در [۲۰] ارائه شده است، استفاده شده است. در این دسته‌بندی، حیوانات به ۱۰ دسته معنایی، طبقه‌بندی شده‌اند که شامل اسب دریایی (۵ تصویر)، پروانه دریایی (۶ تصویر)، کوسه (۵۸ تصویر)، ماهی پهن (۵۲ تصویر)، ماهی زبان (۱۹ تصویر)، سخت‌پوستان (۴ تصویر)، مارماهی (۲۵ تصویر)، مارماهی خمیده (۲۵ تصویر)، ماهی لوله‌ای (۱۶ تصویر) و ماهی چهارگوش (۴۱ تصویر) می‌باشند. ۸۴۸ تصویر به کلاس مشخصی

<sup>۱</sup>Bulls-Eye Performance (BEP)

<sup>۲</sup>Precision

<sup>۳</sup>Recall

مبتنی بر ناحیه، دقت بالاتری در بازیابی شکل ارائه می‌دهند [۲۲]. بنابراین مقادیر  $w_i$  را با توجه به شرایط زیر تعیین گردیده است.

$$\begin{aligned} \text{I)} w_1 > w_2 \\ \text{II)} w_3 > w_4 \\ \text{III)} w_3 + w_4 > w_1 + w_2 \\ \text{IV)} 0 \leq w_i \leq 4 \quad i = 1, 2, 3, 4 \end{aligned} \quad (13)$$

در آزمایشاتی که در ادامه این بخش در مدل بدون نظارت انجام می‌شود، مقادیر  $w_1=3.5$ ،  $w_2=1$ ،  $w_3=3.5$  و  $w_4=3.25$  به عنوان ضرایب فرایند تلفیق در نظر گرفته می‌شوند. همچنین، در هر دو روش  $FoD$  و  $FoGD$ ، مقادیر  $\mu = 0.009$ ،  $\theta = 0.9$ ،  $\varepsilon = 10^{-0.5}$  و  $Inf = 3.5$  جهت کاهش حساسیت به نویز و نیز  $\alpha = 0.83$  برای انتخاب ابعاد بردار ویژگی، استفاده می‌شوند. لازم به ذکر است که در آزمایشات، برای سنجش عدم شباهت بین دو بردار ویژگی از فاصله سیتی بلوک<sup>۱</sup> استفاده می‌شود. انتخاب این روش، بواسطه مقایسه‌های انجام شده با دیگر روش‌ها، انجام پذیرفته است [۲۳].

#### الف) ارزیابی دو روش پیشنهادی در دادگان بخش B از MPEG-7

نتایج دقت بازیابی با آزمون BEP در این دادگان در جدول (۱) نشان داده شده است.

جدول ۱ مقایسه دقت بازیابی در دادگان بخش B از MPEG-7 با استفاده از آزمون BEP.

ردیف	روش	دقت بازیابی با آزمون BEP
۱	توصیف نمونه‌ها در فضای ویژگی $s_1$	۶۹٪
۲	توصیف نمونه‌ها در فضای ویژگی $s_2$	۶۴٪
۳	توصیف نمونه‌ها در فضای ویژگی $s_3$	۷۰٪
۴	توصیف نمونه‌ها در فضای ویژگی $s_4$	۶۸٪
۵	بازیابی با ماتریس عدم شباهت تلفیقی	۸۱/۰۱٪
۶	استخراج بردار ویژگی با روش MDS	۸۱/۶۳٪
۷	استخراج بردار ویژگی با روش ایزومپ ( $d = 24$ )	۸۴/۹۴٪
۸	استخراج بردار ویژگی با روش FoD ( $d = 31$ )	۸۸/۸۶٪
۹	استخراج بردار ویژگی با روش FoGD ( $d = 24$ )	۸۹/۷۳٪

نتایج این جدول نشان می‌دهد که تلفیق چهار فضای ویژگی سبب می‌شود که حداقل ۱۱٪ بهبود نسبت به هرکدام از فضاهای ویژگی حاصل شود. در این رویکرد، تابع عدم شباهت دو نمونه  $x_j$  و  $x_k$  در فضای  $i$ ام  $(d_i(\vec{x}_j^i, \vec{x}_k^i))$ ، بخشی از ویژگی‌های شکل را بیان می‌کند. بنابراین، ترکیب آنها سبب می‌شود که شکاف معنایی در بازیابی کاهش یابد.

در استخراج ویژگی با روش ایزومپ به دلیل ساخت گراف همسایگی و محاسبه کوتاه‌ترین مسیر بین نمونه‌ها در آن گراف، برآورد بهتری از کوتاه‌ترین مسیر در سطح بین نمونه‌ها محاسبه می‌شود. بنابراین دقت بازیابی با روش ایزومپ نسبت به MDS به

تخصیص داده نشده است. بنابراین این تصاویر در کلاس پیش فرضی قرار داده شده‌اند.

ارزیابی این دادگان در رویکرد بدون نظارت انجام شده است. نحوه پرس‌وجو از این دادگان، همانند روش ارائه شده در [۲۰] می‌باشد. برای این منظور، ۳۰ تصویر پرس‌وجو از ۱۰ کلاس انتخاب می‌شوند. در این روش از هر کلاس نبایستی بیشتر از ۳ تصویر برای پرس و جو انتخاب شود. بنابراین هرکدام از ۳۰ تصویر به عنوان پرس‌وجو داده می‌شود  $(x_i)$  و از  $t$  تصویری که بیشترین شباهت را با آن دارند، تعداد تصاویری که در همان کلاس تصویر پرس‌وجو قرار دارند، ثبت می‌شود  $(c_i)$ . دقت و فراخوانمطابق رابطه ۱۱ بدست می‌آیند:

$$\begin{aligned} Precision &= \frac{\sum_{i=1}^{30} c_i}{30 \times t} \\ Recall &= \frac{\sum_{i=1}^{30} c_i}{\sum_{i=1}^{30} g_i} \end{aligned} \quad (11)$$

در این رابطه،  $g_i$  بیانگر تعداد تصاویری است که با  $x_i$  در یک کلاس قرار دارند.

در این دادگان، دقت و فراخوان برای کلاس مشخصی از ماهی‌ها نیز محاسبه می‌شود. برای این منظور تمامی تصاویر کلاس مورد نظر به عنوان پرس‌وجو داده می‌شوند. رابطه ۱۲، دقت و فراخوان را برای تصاویر کلاس A محاسبه می‌کند.

$$\begin{aligned} Precision_A &= \frac{\sum_{i=1}^{size(A)} c_i}{size(A) \times t} \\ Recall_A &= \frac{\sum_{i=1}^{size(A)} c_i}{size(A)^2} \end{aligned} \quad (12)$$

در این رابطه،  $size(A)$  بیانگر تعداد تصاویر کلاس A می‌باشد.

#### ۴-۲ نتایج تجربی و ارزیابی دو روش پیشنهادی

در این بخش به بیان نتایج تجربی و ارزیابی دو مدل بدون نظارت، در دادگان شکل می‌پردازیم. دو دادگان بخش B از MPEG-7 و Fish در ارزیابی دو مدل بدون نظارت آورده شده‌اند. برای استخراج بردارهای توصیف کننده شکل، از روش‌های حوزه فرکانس و حوزه تصویر استفاده شده است. ترکیب اطلاعات حوزه‌های فرکانس و تصویر سبب می‌شود تا اطلاعات از دو حوزه مختلف با هم ترکیب شوند و شناسایی شکل با دقت بالاتری انجام پذیرد. در این مقاله از دو روش مبتنی بر محیط و دو روش مبتنی بر ناحیه که در [۲۱] به تفصیل آورده شده‌اند استفاده می‌شود بنابراین،  $M = 4$  می‌باشد. با انتخاب مقادیر بهینه برای  $w_1$ ،  $w_2$ ،  $w_3$  و  $w_4$ ، دقت بازیابی افزایش می‌یابد. در ارزیابی مدل بدون نظارت، انتخاب این مقادیر می‌تواند با الگوریتم‌های تکاملی و یا با روش جستجوی سراسری انجام پذیرد که در این پیاده‌سازی از جستجوی سراسری استفاده شده است. براساس نتایج بدست آمده از جدول (۱) جدول، فضاهای ویژگی  $s_1$  و  $s_3$ ، دقت بیشتری در بازیابی نسبت به فضاهای  $s_2$  و  $s_4$  دارند. همچنین، روش‌های

<sup>۱</sup>City block



در ارزیابی، علاوه بر آزمون BEP، از نمودار دقت و فراخوان نیز استفاده شده است. شکل ۳، مقایسه دقت و فراخوان دو روش پیشنهادی را با روش‌های دیگر را نشان می‌دهد. کاهش شکاف معنایی دو روش یادگیری منیفلد نسبت به ارزیابی بر اساس ماتریس تلفیقی عدم شباهت، در شکل ۳ الف نشان داده شده است. همچنین، همانگونه که در شکل ۳ ب نشان داده شده است، دو روش پیشنهادی نسبت به ایزومپ، در ارزیابی شکل، دقت و فراخوان بالاتری را دارا می‌باشند. شکل ۳ ج برتری روش پیاده‌سازی شده را نسبت به دیگر روش‌ها در ارزیابی بدون نظارت شکل، نشان می‌دهد.

مقایسه روش FoGD با دیگر روش‌ها در ارزیابی بدون نظارت در جدول (۲) آورده شده است. همانند [۲۴] و [۲۵]، در آزمایشات از مقادیر عدم شباهت بین نمونه‌ها با روش [۲۶ IDSC] به جای مقادیر بدست آمده از فضای  $S_2$  استفاده کرده‌ایم. نتیجه بدست آمده از آزمون BEP با روش FoGD به  $93/96\%$  می‌رسد که بهترین نتیجه گزارش شده در این دادگان می‌باشد.

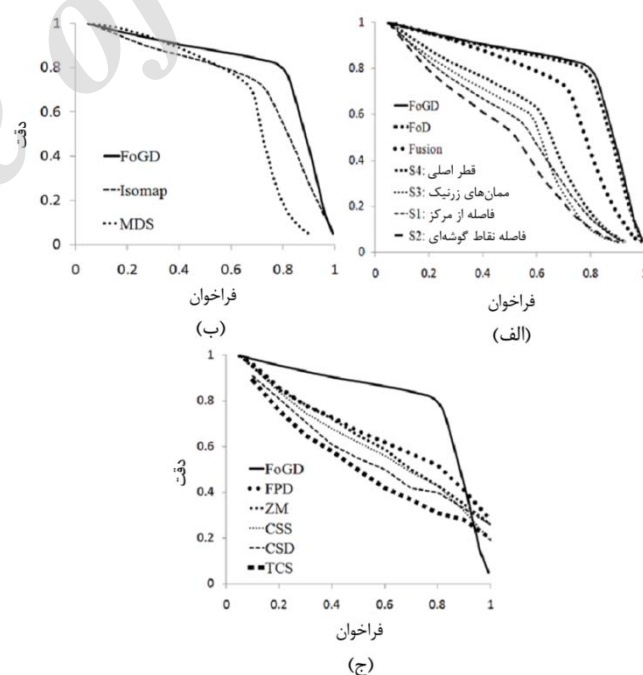
جدول ۲ مقایسه روش FoGD با دیگر روش‌های بدون نظارت در دادگان بخش B از MPEG-7 با آزمون BEP

ردیف	روش	دقت
۱	[۲۷] Shape matching using shape contexts (SC+TPS)	۷۶/۵۱%
۲	[۲۸] On aligning curves	۷۸/۱۷%
۳	[۲۹] Distance sets for recognition	۷۸/۳۸%
۴	[۳۰] Morphological Curvature Scale Spaces (MCSS)	۷۸/۸%
۵	[۳۱] Shape-Based Object Detection via Boundary Structure Segmentation	۸۰/۸۵%
۶	[۳۲] Curvature Scale Space Representation (CSS)	۸۱/۱۲%
۷	[۳۳] Beam Angle Statistics (BAS)	۸۲/۳۷%
۸	[۳۴] Multiscale Closed Contour (MCC)	۸۴/۹۳%
۹	[۳۵] An Adaptive Approach	۸۴/۹۷%
۱۰	[۳۶] Dynamic Space Warping (DSW)	۸۵/۳%
۱۱	[۳۷] Robust symbolic representation for shape recognition and retrieval	۸۵/۹۲%
۱۲	[۳۸] Triangle Area Representation and Dynamic Space Warping (TAR+DSW)	۸۷/۱۳%
۱۳	[۳۹] Shape Matching and Recognition Using Group-Wised Points	۸۷/۱۵%
۱۴	روش پیشنهادی (FoGD)	۸۹/۷۳%
۱۵	[۲۶] Inner Distance Shape Classification and Dynamic Programming (IDSC)	۸۵/۴%
۱۶	[۲۴] IDSC + Learning Graph Transduction	۹۱%
۱۷	[۲۵] IDSC + Learning context sensitive shape similarity by graph transduction	۹۱/۶۱%
۱۸	[۴۰] IDSC+Learning context-sensitive similarity by shortest path propagation	۹۳/۳۵%
۱۹	روش پیشنهادی (FoGD) + IDSC	۹۳/۹۶%

### ب) ارزیابی دو روش پیشنهادی در دادگان Fish

نتایج ارزیابی دو روش بدون نظارت FoGD و FoD در دادگان Fish، در شکل ۴ آورده شده است. کاهش شکاف معنایی این دو روش یادگیری منیفلد، نسبت به ارزیابی بر اساس ماتریس تلفیقی عدم شباهت، در شکل ۴ الف شکل نشان داده شده است. همچنین، همانگونه که در شکل ۴ ب شکل نشان داده شده است، روش FoGD نسبت به ایزومپ، در ارزیابی شکل، دقت و فراخوان بالاتری را دارا می‌باشد. همچنین، شکل ۴ ج شکل، ده برتری روش‌های پیاده‌سازی شده را نسبت به نتایج روش دیگر در این دادگان [۴۱]، نشان می‌دهد.

میزان  $3/2\%$  افزایش می‌یابد. استخراج ویژگی با روش FoD نسبت به ایزومپ، سبب بهبود دقت ارزیابی به میزان  $3/9\%$  می‌شود. این افزایش دقت، به سبب کاهش حساسیت به نویز در عدم شباهت بین نمونه‌ها با روش ارائه شده در FoD می‌باشد. کاهش حساسیت به نویز در عدم شباهت بین نمونه‌ها در روش FoGD با رویکردی متفاوت نسبت به FoD انجام پذیرفته است. این امر سبب گردیده است که تا افزایش دقتی به میزان  $5/2\%$  نسبت به روش ایزومپ حاصل گردد. نتایج تجربی در این دادگان و دادگان Fish، بیانگر برتری نسبی روش FoGD نسبت به FoD می‌باشد. کاهش حساسیت به نویز در FoGD، در هر فضای ویژگی به صورت جداگانه انجام می‌پذیرد و این در حالی است که نویز در روش FoD، پس از تلفیق عدم شباهت‌ها انجام می‌پذیرد. بنابراین در روش FoGD، به صورت پیشگیرانه از تجمیع اثر نویز در عدم شباهت نمونه‌ها در هر فضا، به گراف عدم شباهت تلفیقی، جلوگیری می‌شود. همچنین، در FoGD، کوتاه‌ترین مسیر در سطح در هر فضا جداگانه محاسبه می‌شود. از آنجا که کوتاه‌ترین مسیر در سطح نسبت به عدم شباهت بین نمونه‌ها، به شکاف معنایی نزدیکتر می‌باشد، بنابراین، تلفیق کوتاه‌ترین مسیر در سطح نسبت به تلفیق عدم شباهت، اثر بخشی بهتری دارد.

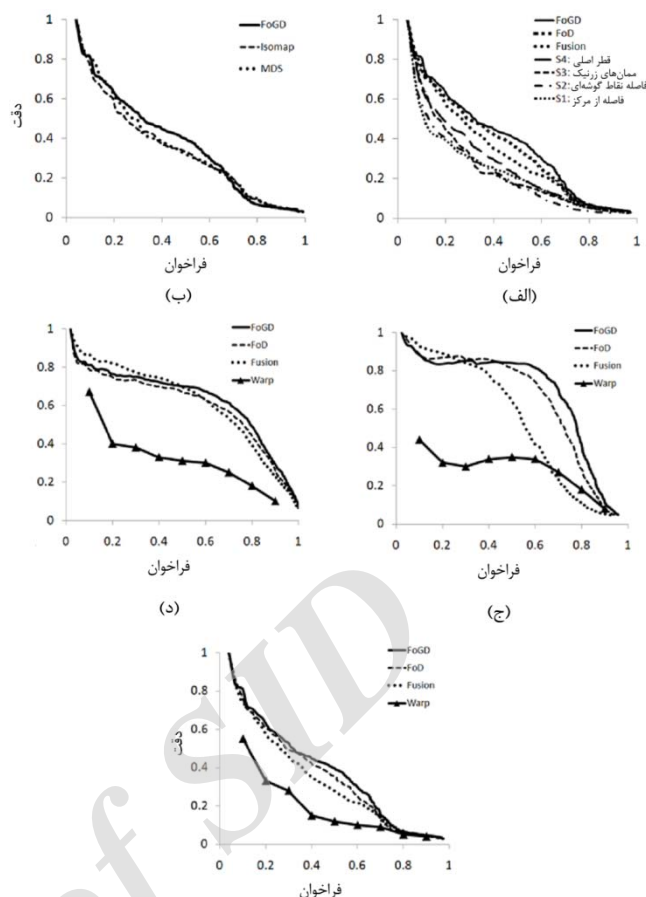


شکل ۳ مقایسه دقت و فراخوان دو روش یادگیری منیفلد پیشنهادی در رویکرد بدون نظارت در دادگان بخش B از MPEG-7. (الف) مقایسه دو روش پیشنهادی در یادگیری منیفلد با ارزیابی بر اساس ماتریس تلفیقی عدم شباهت و با چهار روش توصیف شکل، (ب) مقایسه روش یادگیری منیفلد پیشنهادی با ایزومپ و MDS و (ج) مقایسه روش پیشنهادی با دیگر روش‌ها در ارزیابی بدون نظارت شکل

بنابراین در روش FoGD، به صورت پیشگیرانه از تجمیع نویز در عدم شباهت نمونه‌ها در هر فضا، به گراف عدم شباهت تلفیقی، جلوگیری می‌شود. همچنین، در FoGD، کوتاه‌ترین مسیر در سطح در هر فضا جداگانه محاسبه می‌شود. از آنجا که کوتاه‌ترین مسیر در سطح نسبت به عدم شباهت بین نمونه‌ها، به شکاف معنایی نزدیکتر می‌باشد، بنابراین، تلفیق کوتاه‌ترین مسیر در سطح نسبت به تلفیق عدم شباهت، اثر بخشی بهتری دارد.

## ۵ نتیجه‌گیری و پیشنهادهایی برای کارهای آینده

در بازیابی تصویر مبتنی بر محتوی (CBIR)، تصویر پرس‌وجو به پایگاه تصاویر داده می‌شود و تصاویر مشابه با آن بازیابی می‌شود. در این حوزه، نگاشت غیرصحیح ویژگی‌های سطح پایین تصویر به معانی سطح بالا، سبب می‌شود تا دقت بازیابی کاهش یابد. از آنجا که ساختارهای نهفته شده که بایستی از نمونه‌های فضای ویژگی استخراج شوند، لزوماً ساختارهای خطی نیستند بنابراین با روش‌های استخراج ویژگی به صورت خطی مانند PCA، قادر به کشف آنها نمی‌باشیم. پس از مرور روش‌های غیرخطی استخراج ویژگی که به روش‌های یادگیری منیفلد نیز معروف هستند، مشخص گردید که الگوریتم‌های پایه یادگیری منیفلد برای همین هدف ارائه گردیده‌اند. مشکلی که در رابطه با استفاده از این الگوریتم‌ها در بازیابی تصویر وجود داشت، حساس بودن آنها به نویز می‌باشد. نتایج تجربی که از این الگوریتم‌ها ارائه شده است، بر روی دادگان‌های آزمایشگاهی مانند رول سوئیسی می‌باشد. بنابراین برای بردارهای ویژگی که از دادگان‌های متنوع تصویری به دست می‌آید، دارای عملکرد مطلوبی نمی‌باشند. در این پژوهش جهت حل مشکل حساسیت به نویز، دو روش FoD و FoGD ارائه گردید. در این مدل، مفهوم نگاشت چندگانه معرفی گردید که به معنای نگاشتی است که یک فضای ویژگی را از چند فضای ویژگی با ابعاد مختلف استخراج می‌کند. با آزمون این دو روش بر دو دادگان شکل، مشخص گردید که دو روش پیشنهادی به نحو مؤثری استخراج ویژگی را در فضای منیفلد انجام می‌دهند. هر چند این دو روش توسعه‌ای بر روش ایزومپ می‌باشند، ولی در نتایج تجربی آزمایش‌ها، مشخص گردید که رویکرد حذف اثر نویز در این دو روش توانسته است به نحو مؤثری شکاف معنایی را در بازیابی کاهش دهد. مؤثر بودن این روش در بازیابی شکل مبتنی بر محتوی و مقایسه آن با دیگر روش‌ها نیز در نتایج تجربی آزمایش‌ها آورده شده است. در ادامه پژوهش در این حوزه به دنبال آن هستیم تا با فرضی دیگر متفاوت با فرضی که در الگوریتم‌های یادگیری منیفلد وجود دارد، استخراج ویژگی را انجام دهیم. در روش‌های یادگیری منیفلد، وجود ابعاد یکسان برای تمامی نمونه‌ها در فضای منیفلد در نظر گرفته می‌شود. با نقض این فرض، بردار ویژگی هر نمونه می‌تواند تعدادی از ویژگی‌ها را شامل نباشد. مثلاً چنانچه بردار ویژگی تصویر متعلق به انسان باشد، ویژگی جنسیت در آن معنا پیدا می‌کند و برای دیگر معانی هرچند



شکل ۴ مقایسه دقت و فراخوان دو روش یادگیری منیفلد پیشنهادی در رویکرد بدون نظارت در دادگان ماهی. (الف) مقایسه دو روش پیشنهادی در یادگیری منیفلد با بازیابی براساس ماتریس تلفیقی عدم شباهت و با چهار روش توصیف شکل، (ب) مقایسه روش یادگیری منیفلد پیشنهادی با ایزومپ و MDS، (ج)، (د) و (ه) مقایسه دو روش پیشنهادی و همچنین بازیابی براساس ماتریس عدم شباهت تلفیقی با روش Warp. در (ج)، تصویر پرس‌وجو می‌تواند متعلق به هر کلاسی باشد. در (د) تصویر پرس‌وجو متعلق به کلاس پهن‌ماهی بوده و در (ه) تصویر پرس‌وجو به کلاس ماهی‌های چهارگوش متعلق می‌باشد.

## ج) تحلیل نتایج

استخراج ویژگی با روش FoD نسبت به ایزومپ، سبب بهبود دقت بازیابی می‌شود. این افزایش دقت، به سبب کاهش حساسیت به نویز در عدم شباهت بین نمونه‌ها با روش ارائه شده در FoD می‌گردد. کاهش حساسیت به نویز در عدم شباهت بین نمونه‌ها در روش FoGD با رویکردی متفاوت نسبت به FoD انجام پذیرفته است. نتایج تجربی بیانگر برتری نسبی روش FoGD نسبت به FoD می‌باشد. کاهش حساسیت به نویز در FoGD، در هر فضای ویژگی به صورت جداگانه انجام می‌پذیرد و این در حالی است که نویز در روش FoD، پس از تلفیق عدم شباهت‌ها انجام می‌پذیرد.

- [14] M. A. Z. Chahooki, N. M. Charkari, "Learning the Shape Manifold to Improve Object Recognition", *Machine Vision and Applications*, vol. 24, pp. 33-46, 2013.
- [15] I. Bloch, "Information Fusion in Signal and Image Processing", Wiley, 2008, ISBN: 9781848210196.
- [16] D. L. Hall, J. Llinas, "An introduction to multisensor data fusion", *Proceedings of the IEEE*, vol. 85(1), pp. 6-23, 1997.
- [17] A. Webb, "Statistical pattern recognition", Wiley, 1999, ISBN: 0470845139.
- [18] L. J. Latecki, R. Lakamper, "Shape similarity measure based on correspondence of visual parts", *IEEE Transactions Pattern Analysis Machine Intelligence*, vol. 22(10), pp. 1185-1190, 2000.
- [19] F. Mokhtarian, S. Abbasi, J. Kittler, "Robust and efficient shape indexing through curvature scale space", *Proceedings of British Machine Vision Conference*, pp. 53-62, 1996.
- [20] I. Bartolini, P. Ciaccia, M. Patella, "Warp: accurate retrieval of shapes using phase of Fourier descriptors and time warping distance", *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 27(1), pp. 142-147, 2005.
- [21] م. ع. زارع چاهوکی، ن. مقدم چرکری، تلفیق شاخص‌های توصیف‌کننده شکل در بازیابی تصویر مبتنی بر محتوی، ششمین کنفرانس بینایی ماشین و پردازش تصویر ایران، دانشگاه اصفهان، ۱۳۸۹
- [22] K. Mikolajczyk, C. Schmid, "Indexing based on scale invariant interest points", *Proceedings of 8th IEEE Computer Vision*, pp. 525-531, 2001.
- [23] M. A. Z. Chahooki, N. M. Charkari, "Shape Retrieval Based on Manifold Learning by Fusion of Dissimilarity Measures", *IET Image Processing*, vol. 6(4), pp. 327-336, 2012.
- [24] X. Yang, X. Bai, L. Latecki, Z. Tu, "Improving shape retrieval by learning graph transduction", *Computer Vision ECCV*, pp. 788-801, 2008.
- [25] X. Bai, X. Yang, L. J. Latecki, W. Liu, Z. Tu, "Learning context-sensitive shape similarity by graph transduction", *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 32(5), pp. 861-874, 2009.
- [26] H. Ling, D. W. Jacobs, "Shape classification using the inner-distance", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29(2), pp. 286-299, 2007.
- [27] S. Belongie, J. Malik, J. Puzicha, "Shape matching and object recognition using shape contexts", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24(4), pp. 509-522, 2010.
- [28] T. B. Sebastian, P. N. Klein, B. B. Kimia, "On aligning curves", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25(1), pp. 116-125, 2003.

این ویژگی مفهوم ندارد ولی ویژگی‌های معنا دار دیگری مطرح می‌باشد.

## مراجع

- [1] J. B. Tenenbaum, "Mapping a manifold of perceptual observations", *Advances in neural information processing systems*, pp. 682-688, 1998.
- [2] X. Geng, D. C. Zhan, Z. H. Zhou, "Supervised nonlinear dimensionality reduction for visualization and classification", *IEEE Transactions on Systems, Man, And Cybernetics-Part B: Cybernetics*, vol. 35(6), pp. 1098 - 1107, 2005.
- [3] C. G. Li and J. GUO, "Supervised Isomap with explicit mapping", *Proceedings of the First International Conference on Innovative Computing, Information and Control*, pp. 345-348, 2006.
- [4] H. Zhaoa, S. Sunb, Z. Jinga, J Yang, "Local structure based supervised feature extraction", *Pattern Recognition*, vol. 39(8), pp. 1546-1550, 2006.
- [5] C. Orsenigo, C. Vercellis, "An effective double-bounded tree-connected Isomap algorithm for microarray data classification", *Pattern Recognition Letters*, vol. 33(1), pp. 1-16, 2012.
- [6] J. B. Tenenbaum, V. Silva, J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction", *Science*, vol. 290(5500), pp. 2319-2323, 2000.
- [7] S. T. Roweis, L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding", *Science*, vol. 290(5500), pp. 2323-2326, 2000.
- [8] M. Belkin, P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation", *Neural computation*, vol. 15(6), pp. 1373-1396, 2003.
- [9] D. L. Donoho, C. Grimes, "Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 55-91, 2003.
- [10] K. Q. Weinberger, L. K. Saul, "Unsupervised learning of image manifolds by semidefinite programming", *International Journal of Computer Vision*, vol. 70(1), pp. 77-90, 2006.
- [11] X. He, W. Y. Ma, H. J. Zhang, "Learning an image manifold for retrieval", *Proceedings of 12th annual ACM international conference on Multimedia*, pp. 17-23, 2004.
- [12] Y. Y. Lin, T. L. Liu, and H. T. Chen, "Semantic manifold learning for image retrieval", *Proceedings of 13th annual ACM international conference on Multimedia*, pp. 249-258, 2005.
- [13] X. He, D. Cai, J. Han, "Learning a maximum margin subspace for image retrieval," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20(2), pp. 189-201, 2008.



محمد علی زارع چاهوکی کارشناسی خود را در رشته مهندسی کامپیوتر گرایش نرم‌افزار در سال ۱۳۷۸ از دانشگاه شهید بهشتی دریافت کردند. سپس کارشناسی ارشد و دکتری خود را در همان رشته و گرایش در سال‌های ۱۳۸۳ و ۱۳۹۲ از دانشگاه

تربیت مدرس اخذ نمودند. در حال حاضر ایشان استادیار گروه مهندسی کامپیوتر دانشگاه یزد هستند و علاقه‌مند به پژوهش در حوزه‌های یادگیری ماشین، مهندسی نرم‌افزار (متدولوژی و کاربردهای مرتبط با یادگیری)، بینایی ماشین (بازیابی تصویر و نظارت دیداری) و نهان‌نگاری هستند.

- [29] C. Grigorescu, N. Petkov, "Distance sets for shape filters and shape recognition". IEEE Transactions on Image Processing, vol. 12(10), pp. 1274-1286, 2003.
- [30] A. C. Jalba, M. H. F. Wilkinson, J. Roerdink, "Shape representation and recognition through morphological curvature scale spaces", IEEE Transactions on Image Processing, vol. 15(2), pp. 331-341, 2006.
- [31] A. Toshev, B. Taskar, K. Daniilidis, "Shape-Based Object Detection via Boundary Structure Segmentation", International journal of computer vision, pp. 1-24, 2012.
- [32] F. Mokhtarian, M. Bober, "Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization", Kluwer Academic Publishers, vol. 25, 2003.
- [33] N. Arica, F. T. Y. Vural, "BAS: a perceptual shape descriptor based on the beam angle statistics", Pattern Recognition Letters, vol. 24(9-10), pp. 1627-1639, 2003.
- [34] T. Adamek, N. E. O'Connor, "A multiscale representation method for nonrigid shapes with a single closed contour", IEEE Transactions on Circuits and Systems for Video Technology, vol. 14(5), pp. 742-753, 2004.
- [35] A. Bandera, E. Antnez, R. Marfil, "An Adaptive Approach for Affine-Invariant 2D Shape Description", Pattern Recognition and Image Analysis, vol. 5524, pp. 417-424, 2009.
- [36] N. Alajlan, I. El Rube, M. S. Kamel, G. Freeman, "Shape retrieval using triangle-area representation and dynamic space warping", Pattern Recognition, vol. 40(7), pp. 1911-1920, 2007.
- [37] M. R. Daliri, V. Torre, "Robust symbolic representation for shape recognition and retrieval", Pattern Recognition, vol. 41(5), pp. 1782-1798, 2008.
- [38] N. Alajlan, M. S. Kamel, G. H. Freeman, "Geometry-based image retrieval in binary image databases", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30(6), pp. 1003-1013, 2008.
- [39] J. Wang, Y. Zhou, X. Bai, W. Liu, W. "Shape matching and recognition using group-wised points", Advances in Image and Video Technology, pp. 393-404, 2012.
- [40] J. Wang, Y. Li, X. Bai, Y. Zhang, C. Wang, N. Tang, "Learning context-sensitive similarity by shortest path propagation", Pattern Recognition, vol. 44(10), pp. 2367-2374, 2011.
- [41] I. Bartolini, P. Ciaccia, M. Patella, "Warp: accurate retrieval of shapes using phase of Fourier descriptors and time warping distance", IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 27(1), pp. 142-147, 2005.
- [42] J. J. A. Lee, M. Verleysen, "Nonlinear dimensionality reduction", Springer, 2007.