

تاریخ دریافت مقاله: ۹۷/۰۴/۱۲
تاریخ پذیرش مقاله: ۹۷/۰۶/۰۷

شناسایی هویت از روی نحوه راه رفتن با استفاده از مدل سلسله مراتبی فازی - اسپایکی

رنا صنیعی

آزمایشگاه پردازش تصاویر، دانشکده مهندسی و علوم کامپیوتر، دانشگاه شهید بهشتی، تهران، ایران
پست الکترونیکی: saniei.r@gmail.com

عباس قبله

آزمایشگاه پردازش تصاویر، دانشکده مهندسی و علوم کامپیوتر، دانشگاه شهید بهشتی، تهران، ایران
پست الکترونیکی: abbas.ghebleh@gmail.com

محسن ابراهیمی مقدم*

آزمایشگاه پردازش تصاویر، دانشکده مهندسی و علوم کامپیوتر، دانشگاه شهید بهشتی، تهران، ایران
پست الکترونیکی: m_moghadam@sbu.ac.ir

چکیده

داده به ترتیب چالش‌های تغییر سرعت راه رفتن و تغییر نوع پوشش را زیر ذره بین قرار می‌دهند. معیارهای ارزیابی به کار گرفته شده در این مقاله نرخ شناسایی (به‌طور خاص در رتبه ۱ و رتبه ۵) و منحنی CMC می‌باشد. در نهایت، نتایج به دست آمده بر روی مجموعه داده A و B با استفاده از مدل پیشنهاد شده، با نتایج روش‌های دیگر مقایسه شده است. این مقایسه نشان می‌دهد مدل HFS تقریباً نسبت به تمامی روش‌ها به نرخ شناسایی بهتری در رتبه ۱ و رتبه ۵ دست می‌یابد.

واژگان کلیدی: شناسایی الگو، ویژگی‌های کرتکس، ساختار سلسله مراتبی، شناسایی هویت از طریق نحوه راه رفتن، روش‌های فازی

۱. مقدمه

مغز پیچیده‌ترین عضو بدن انسان و مرکز سیستم عصبی می‌باشد که با وجود پیشرفت‌های چشمگیر علوم

شناسایی هویت از روی نحوه راه رفتن، از این نظر که در روش‌های از راه دور نیز قابل محاسبه است، یکی از ویژگی‌های زیست‌سنجی (بیومتریکی) مهم به حساب می‌آید. تحقیقات گسترده‌ای در این زمینه صورت گرفته و روش‌های متعددی ارائه شده است. اما تعداد کمی از این تحقیقات بر استفاده از تکنیک‌های شناسایی اشیاء توسط مغز تمرکز کرده‌اند. در این مقاله، مدلی برای شناسایی اشیاء بر مبنای ساختار سلسله مراتبی و پیچیده ویژگی‌های کرتکس مغز، ارائه شده است. مدل ارائه شده که آن را مدل سلسله مراتبی فازی - اسپایکی می‌نامیم، بر سازوکار تولید پاسخ نورون‌ها در ویژگی‌های کرتکس و ماهیت فازی‌شان تمرکز دارد. یکی از نوآوری‌ها در این مدل، پیشنهاد یک لایه دسته‌بند فازی مشابه سیستم استنتاج عصبی-فازی تطبیق‌پذیر است. برای ارزیابی مدل ارائه شده، از مجموعه داده‌های تردمیل A و B استفاده شده است. این دو مجموعه

* نویسنده مسئول

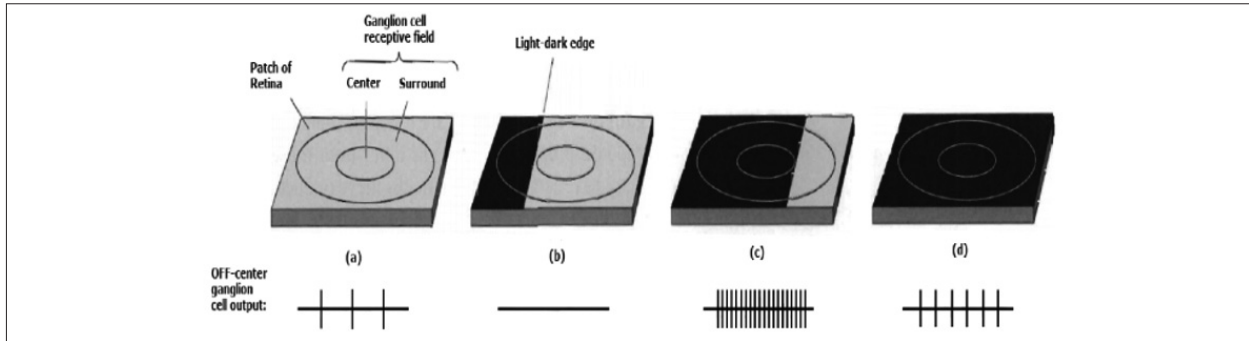
پزشکی و اعصاب، تا امروز بخش زیادی از عملکردش برای دانشمندان ناشناخته است. مغز یک انسان بالغ از حدود یک تریلیون سلول تشکیل شده که ۵۰ تا ۱۰۰ میلیون (۵ تا ۱۰ درصد) آن‌ها نورون‌ها هستند. مغز مانند یک پردازشگر قوی عمل کرده و به انسان این قدرت را می‌دهد که تقریباً بدون هیچ گونه زحمتی، اشیاء، افراد و سایر چیزها را شناسایی کند. ویژوال کرتکس^۱ بخشی از مغز است که مسئولیت پردازش اطلاعات مربوط به بینایی را بر عهده دارد. متخصصان علم بینایی ماشین همواره به ساختار مغز و عملکرد آن به‌عنوان منبع الهامی برای حل مسایل مطرح در بینایی ماشین می‌نگرند. بدون شک الهام گرفتن از جنبه زیست‌شناختی سیستم‌های بینایی در جهت حل مسایل اساسی بینایی ماشین وابسته به پیشرفت علمی همچون فیزیولوژی و علوم اعصاب^۲ می‌باشد. از این رو کاری که هابل^۳ و ویزل^۴ (هابل و ویزل، ۱۹۶۲) در جهت رمزگشایی آنچه در قسمت ویژوال کرتکس مغز رخ می‌دهد انجام دادند، اولین و مهم‌ترین محرک در جامعه دانشمندان علم بینایی ماشین برای پیوند این علم با مفاهیم زیست‌شناسانه به حساب می‌آید. در آن زمان بر اساس نتایج به دست آمده توسط این دو نفر، تلاش‌هایی برای ارائه ساختاری سلسله‌مراتبی، مشابه آنچه در مغز وجود دارد صورت گرفت، اما جزییات ناقص و کمبود اطلاعات دقیق در این مورد، باعث شد تحقیقات دانشمندان به کندی پیش رفته و تا سال‌ها پیشرفت چشمگیری در این زمینه وجود نداشته باشد (کروگر و همکاران، ۲۰۱۳). اما امروزه با توجه به اطلاعات کامل‌تری که از عملکرد بخش ویژوال کرتکس مغز و ساز و کار دقیق نورون‌های زیست‌شناختی در اختیار است، سیستم‌های متنوعی ارائه شده‌اند. با قطعیت می‌توان گفت در همه این سیستم‌ها، مدل سلول‌های ساده^۵ و پیچیده^۶، ارائه شده توسط هابل و ویزل، نقش اصلی را

ایفا می‌کند. نئوکاگنیترون^۷ (فوکوشیما و همکاران، ۱۹۸۳)، شبکه‌های عصبی کانولوشنال^۸ (لکن و همکاران، ۱۹۹۸) و HMAX (سره و همکاران، ۲۰۰۴) نمونه‌هایی از معروف‌ترین مدل‌های شناسایی الگو بر اساس ساختار سلسله‌مراتبی ویژوال کرتکس هستند. قبل از معرفی مدل ارائه شده در این مقاله، ذکر این نکته حائز اهمیت است که برای ارزیابی این مدل، مسئله شناسایی هویت برپایه نحوه راه رفتن^۹ یا گیت، انتخاب شده است. در کنار ویژگی‌های زیست‌سنجی رایجی مانند اثر انگشت، قرنیه و صورت، گیت یک برتری قابل توجه دارد: مهم‌ترین مزیت این خصیصه این است که از فاصله دور قابل تشخیص بوده، نیاز به همکاری و مشارکت فرد مورد نظر ندارد (عبدالحمید و همکاران، ۲۰۱۲). با این حال، گیت نقاط ضعف مخصوص به خود را نیز دارد. برای مثال تضمینی وجود ندارد که مشخصات سوژه در مجموعه آزمون^{۱۰} مشابه با مشخصات آن در مجموعه آموزش^{۱۱} باقی بماند. این مشخصات شامل تغییرات در پوششش، تغییرات در سرعت راه رفتن، حمل اشیاء (مانند کیف) و تغییرات در زاویه سوژه نسبت به دوربین عکس‌برداری می‌باشد. چالش اصلی در شناسایی گیت، یافتن ویژگی‌هایی است که نسبت به تغییرات ذکر شده مقاوم باشند. در این بررسی، بر روی تغییرات در پوششش و سرعت تمرکز شده و برای ارزیابی، مجموعه داده بزرگ OU-ISIR Treadmill مورد استفاده قرار گرفته است (حسین و همکاران، ۲۰۱۰).

ادامه این مقاله به شکل زیر سازمان‌دهی شده است. در بخش ۲ برخی از مهم‌ترین مدل‌های شناسایی الگو که ساختار خود را از ویژوال کرتکس مغز الهام گرفته‌اند، معرفی می‌شوند. سپس چند روش مهم در شناسایی گیت مورد بررسی قرار خواهند گرفت. بخش ۳ به معرفی مدل ارائه شده و شرح جزییات هر لایه آن اختصاص دارد. در بخش ۴ نتایج آزمایش‌های انجام گرفته به منظور ارزیابی

7- NeoCognitron
8- Convolutional Neural Network (CNN)
9- Gait Recognition
10- Test Set
11- Train Set

1- Visual Cortex
2- Neuroscience
3- Hubel
4- Wiesel
5- Simple Cells
6- Complex Cells



شکل ۱: نمایش منطقه گیرنده یک سلول گانگلیون شبکه‌ای و اثر تابش نور بر قسمت‌های مختلف آن. از چپ به راست می‌توان دید هنگامی که لبه روشن-تاریک به سمت مرکز مثبت منطقه گیرنده حرکت می‌کند، فرکانس تولید پاسخ افزایش می‌یابد. با پوشانیده شدن کل منطقه گیرنده توسط لبه تاریک-روشن، به تدریج فرکانس تولید پاسخ کاهش می‌یابد (رودیگر).

مکان منطقه تحریک‌پذیر^{۱۴} (مثبت) و بازدارنده^{۱۵} (منفی) است. سلول‌های گانگلیون منطقه گیرنده‌ای متقارن و دایره شکل دارند. در وسط آن‌ها یک مرکز و در پیرامون آن‌ها یک منطقه پیرامونی وجود دارد که علامت‌های متضاد دارند. در سلول‌های ساده و پیچیده وضعیت پیچیده‌تری وجود دارد. شکل ۱ منطقه گیرنده دو نوع سلول گانگلیون را نمایش می‌دهد: مرکز روشن^{۱۶} و مرکز خاموش^{۱۷}. شکل ۲ منطقه گیرنده یک سلول ساده و پاسخش به محرک‌های مختلف را نمایش می‌دهد. سلولی که به‌طور خاص حساس به خطوط جهت‌دار است. همان‌طور که در شکل مشخص است، محرکی که بخش بیشتری از منطقه تحریک‌پذیر را پوشاند، نسبت به محرکی که بخش کمتری از این منطقه را می‌پوشاند، منجر به تولید پاسخ قوی‌تری از سلول می‌شود. اگر سلول تحت تاثیر محرکی باشد که هر دو منطقه تحریک‌پذیر و بازدارنده را به‌طور کامل پوشانده، پاسخی تولید نخواهد کرد.

سلول‌های پیچیده، ساز و کاری به مراتب پیچیده‌تر از سلول‌های ساده دارند. این نوع سلول‌ها ورودیشان را از گروهی از سلول‌های ساده دریافت می‌کنند و مانند آن‌ها، هرکدام به بخش کوچکی از محدوده بینایی انسان حساس هستند. اما نمی‌توان عملکردشان را با تقسیم منطقه گیرنده‌شان به دو بخش کاملاً مجزای تحریک‌پذیر و بازدارنده توصیف کرد (هابل و همکاران، ۱۹۹۵). عملکرد

مدل پیشنهادی را ارائه خواهیم کرد. در پایان، در بخش ۶ به نتیجه‌گیری و بررسی کارهای آینده می‌پردازیم.

۲. مرور کارهای پیشین

در ادامه مروری بر کارهای پیشین در دو حیطه مدل‌های شناسایی الگوی الهام گرفته از زیست‌شناسی بر مبنای ساختار و ویژوال کورتکس و مدل‌های شناسایی گیت خواهیم داشت.

۱-۲ مدل‌های شناسایی الگوی الهام گرفته از زیست‌شناسی

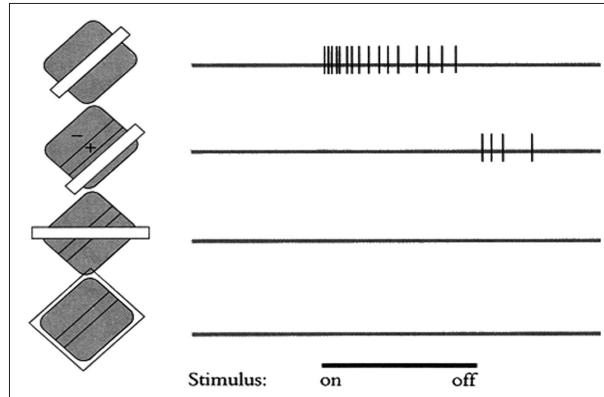
از این‌گونه مدل‌ها می‌توان به شبکه‌های کانولوشنال، نئوکانیترون و HMAX به‌عنوان سه نمونه از شناخته‌شده‌ترین مدل‌های ارائه‌شده با الهام از زیست‌شناسی اشاره کرد. همان‌طور که قبلاً اشاره شد، اصلی‌ترین محرک برای ساخت این مدل‌ها، مدل سلول‌های ساده و پیچیده‌ای است که هابل و ویزل ارائه کردند. هر سلول ساده منطقه گیرنده^{۱۲} ای کوچک و کاملاً مشخص دارد که در اثر تحریک شدن با یک نقطه نورانی و بسته به این‌که نور در کجای منطقه گیرنده می‌افتد، پاسخ مثبت یا منفی تولید می‌کند (هابل و همکاران، ۱۹۹۵). ورودی این سلول‌ها از گروهی از سلول‌ها به نام گانگلیون^{۱۳} که در لایه‌های ابتدایی‌تر قرار دارند، تامین می‌شود. تفاوت این سلول‌ها با سلول‌های ساده و پیچیده در هندسه آن‌ها و

14- Inhibitory
15- Excitatory
16- On-Center
17- Off-Center

12- Receptive Field
13- Ganglion

همسایگی‌هایی مربعی از تصویر ورودی (متناظر با منطقه گیرنده) متصل هستند.

این خروجی معادل با ضرب داخلی وزن‌های شبکه در این لایه و مقادیر همسایگی مربعی است که به‌عنوان ورودی به هر نورون متصل هستند. در لایه همجوشی با یک سازوکار نمونه‌برداری، مقداری جدید به ازای هر بلوک از نتیجه مرحله قبل محاسبه می‌شود. این دو لایه به صورت پشت سر هم تکرار شده، در نهایت خروجی آن‌ها به یک شبکه پرسپترون کاملاً متصل داده می‌شود. این شبکه وظیفه رده‌بندی^{۲۴} را بر عهده دارد. HMAX مدلی پیش‌بازخورد است که ساختار سلسله مراتبی و یژوال کرتکس را با لایه‌های متوالی S و C متناظر با سلول‌های ساده و پیچیده در منطقه V1 مغز، شبیه‌سازی می‌کند. V1 اولین منطقه‌ای در ناحیه کرتیکال مغز است که اطلاعات تصویری را پردازش می‌کند (کروگر و همکاران، ۲۰۱۳). این منطقه از سلول‌هایی حساس به لبه^{۲۵}، خط^{۲۶} یا شبکه‌ای از خطوط^{۲۷} (برای مثال بافتی از خطوط منحنی) تشکیل شده است. با پیش‌رفتن در طول سلسله مراتب این مدل، تغییرناپذیری^{۲۸} و به‌گزینی^{۲۹} افزایش می‌یابد. در لایه S1 پالایه گابور در مقیاس و زوایای مختلف بر تصویر ورودی اعمال می‌شود تا به‌گزینی مدل افزایش پیدا کند. سپس واحدهای C1 عمل بیشینه‌گیری را روی ورودی‌های خود از زوایای یکسان، همسایگی‌های محلی و مقیاس‌های مجاور، اعمال می‌کنند تا تغییرناپذیری مدل در برابر جابجایی افزایش یابد. واحدهای S2 مشابه نورون‌های تنظیم زاویه در منطقه اینفروتمپرال در کرتکس عمل می‌کنند. ورودی هر واحد S2 یک تکه از تصویر از خروجی C1ها در مرحله قبل است. تکه‌هایی که مرتبط با زوایای متفاوت و مقیاس یکسانی هستند (سره و همکاران، ۲۰۰۷). خروجی لایه C2 با یک بیشینه‌گیری عمومی روی تمامی مقیاس‌ها و مکان‌ها



شکل ۲: نمایش منطقه گیرنده سلول‌های ساده جهت‌دار و چگونگی تولید پاسخ آن توسط محرک‌های مختلف (هابل و همکاران، ۱۹۹۵)

سلول‌های پیچیده به صورت یک لایه نمونه‌برداری، برای مثال میانگین یا بیشینه گروهی از خروجی‌های لایه قبل، مدل‌سازی می‌شود.

نئوکائیترون، شبکه‌های عصبی کانولوشنال و HMAX همگی بر اساس مفاهیم یکسانی بنا شده‌اند و تفاوت آن‌ها در برخی جزئیات و نحوه پیاده‌سازی است. نئوکائیترون یک مدل شبکه عصبی در جهت شناسایی الگو طبق ساختار و یژوال کرتکس است. این مدل، شبکه‌ای چندلایه‌ای، تشکیل شده از ارتباطات آبخاری^{۳۸} بین چندین مرحله پردازشی است که قادر است الگوهای تصویری را به خوبی از طریق آموزش، شناسایی کند. هر مرحله از یک لایه S-cell (سلول‌های ساده) و یک لایه C-cell (سلول‌های پیچیده) به دنبال آن، تشکیل شده است. S-cellها استخراج‌کننده ویژگی هستند و C-cellها عملیات تارکردگی^{۳۹} را انجام می‌دهند. بنابراین خروجی C-cellها در مقایسه با خروجی S-cellها، به تغییر مکان^{۲۰} الگوی ورودی کمتر حساسیت دارد. CNNها از تعدادی لایه تشکیل شده‌اند که هر یک دارای یک یا چند صفحه هستند. یک CNN به‌طور معمول دارای سه نوع لایه است: لایه کانولوشنال^{۲۱}، لایه همجوشی^{۲۲} و شبکه پرسپترون کاملاً متصل^{۲۳}. لایه کانولوشنال خروجی نورون‌هایی را محاسبه می‌کند که به

24- Classification
25- Edge
26- Bar
27- Gratings
28- Invariancy
29- Selectivity

18- Cascade Connection
19- Blurring
20- Shift
21- Convolutional
22- Pooling
23- Fully Connected Multi-Layer Perceptron

محاسبه می‌شود. بنابراین ویژگی‌های نهایی به دست آمده، نمایشی مقاوم نسبت به مکان و مقیاس از تصویر ورودی هستند.

۲-۲ مدل‌های شناسایی گیت

مدل‌های شناسایی گیت بر اساس مرحله استخراج ویژگی به دو دسته کلی مبتنی بر مدل^{۳۰} و جامع^{۳۱} (بدون مدل) تقسیم می‌شوند. روش‌های مبتنی بر مدل بر مدل کردن قسمت‌های مختلف بدن مانند پاها، سر، آرنج و ردیابی آن‌ها در تمامی قاب‌ها تمرکز دارد. در این روش‌ها سپس از انواع مختلفی از پارامترهای پویا یا ایستای ساختار بدن، مانند اندازه ارتفاع، بیشترین فاصله بین لگن و پاها، فاصله بین سر و لگن و فاصله بین پاها استفاده می‌شود (جانسون و همکاران، ۲۰۰۱). برخی دیگر از روش‌ها، از مدل‌های ساده‌ای مانند دو پاندول متصل به هم برای مدل‌سازی پاها و نحوه حرکت آن‌ها استفاده می‌کنند (کندو و همکاران، ۱۹۹۷، کندو و همکاران، ۲۰۰۳، نش و همکاران، ۱۹۹۷، یام و همکاران، ۲۰۰۲). روش‌های مبتنی بر مدل معمولاً محاسبات سنگین‌تری از روش‌های دسته دوم دارند. به علاوه به علت خطای ذاتی که در تطبیق مدل وجود دارد، عموماً این روش‌ها دقت پایین‌تری نسبت به روش‌های جامع خواهند داشت. راهکارهای بدون مدل را مبتنی بر پیکره نیز می‌نامند. تمرکز این مدل‌ها بر شکل پیکره یا کل حرکت بدن فرد بدون در نظر گرفتن ساختار آن است. در بیشتر این روش‌ها، نمایشی از تمامی گیت‌های متوالی از یک سوژه ساخته می‌شود. تصویر انرژی گیت^{۳۲} یکی از کارآمدترین و در عین حال ساده‌ترین این نمایش‌هاست. برای محاسبه GEI ابتدا پیکره فرد از هر قاب که به صورت دودویی در آمده، استخراج می‌شود. پس از تراز کردن همه تصاویر، GEI از رابطه زیر محاسبه می‌شود:

$$GEI = \frac{1}{T} \sum_{t=1}^T I(x, y, t). \quad (1)$$

در این رابطه، T تعداد کل قاب‌ها در یک سیکل و $I(x, y, t)$ مقدار روشنایی پیکسل با مختصات x و y در قاب t ام است. اگر چه تصویر انرژی گیت به دلیل عمل میانگین‌گیری نسبت به نوفه مقاوم است، اما در (یو و همکاران، ۲۰۰۶، مان و همکاران، ۲۰۰۶) نشان داده شده است که این نمایش، نسبت به ویژگی‌های متغیر سوژه حساس است. برای حل این مشکل، نمایش دیگری به نام تصویر آنتروپی گیت^{۳۳} (بشیر و همکاران، ۲۰۰۹) ارائه شد که با محاسبه آنتروپی شانون^{۳۴} در هر پیکسل از تصویر به دست می‌آید. بنابراین در تصویر آنتروپی گیت یا GENI بیشتر اطلاعات مربوط به قسمت‌های محرک تصویر به دست آمده و نسبت به ویژگی‌های متغیری که ظاهر سوژه را تغییر می‌دهند، مقاوم است (بشیر و ژانگ، ۲۰۰۹). در روش دیگری با استخراج ویژگی‌های حوزه فرکانس و استفاده از یک مدل تبدیل دید^{۳۵}، نمایشی مقاوم در برابر تغییر زاویه دید ارائه شده است (اکیهارا و همکاران، ۲۰۰۶). در تحقیقی دیگر با توجه به این نکته که بخش‌های حرکتی بدن نسبت به بخش‌های ساکن کمتر دچار تغییر می‌شوند، نمایش EndFT را معرفی کردند (روکانوجامان و همکاران، ۲۰۱۳). در این نمایش ویژگی‌های حوزه فرکانس و GENI استخراج می‌شوند تا بیشتر روی نواحی ساکن بدن تمرکز شود. در پژوهشی دیگر که بر پایه نظریه مجموعه اطلاعات^{۳۶} بنا شده است، آئورا و همکاران تصویر اطلاعات گیت^{۳۷} یا به اختصار GI را پیشنهاد دادند (آرورا و همکاران، ۲۰۱۵). برای محاسبه GI از یک تابع آنتروپی برای نشان دادن عدم قطعیت در منابع اطلاعاتی که در اینجا قاب‌های گیت هستند، استفاده می‌شود. در این روش از دو ویژگی دیگر به نام‌های GI^{۳۸} و EF و GI-SF^{۳۹} نیز استفاده می‌شود که به ترتیب بر پایه ویژگی انرژی و سیگموییدی بنا شده‌اند. آئورا و همکاران همچنین نمایشی معرفی نمودند که از ترکیب هیستوگرام

33- Gait Entropy Image (GEI)

34- Shannon Entropy

35- View Transformation Model

36- Information Set Theory

37- Gait Information Image

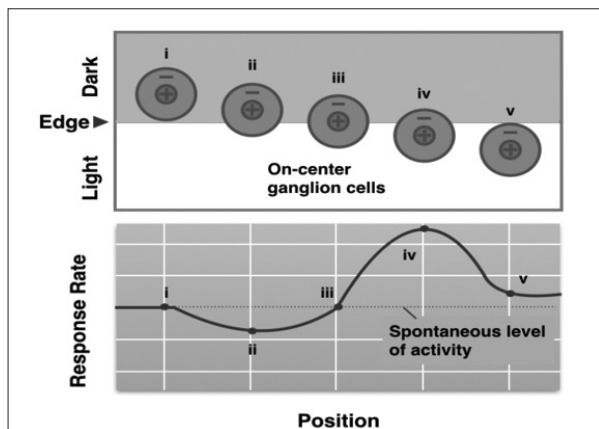
38- Gait Information Image with Energy Feature

39- Gait Information Image with Sigmoid Feature

30- Model-based

31- Holistic

32- Gait Energy Image (GEI)



شکل ۳: ماهیت فازی عملکرد سلول‌ها گانگلیون. پاسخ قوی به معنای پاسخی با فرکانس بالا، پاسخ ضعیف به معنای پاسخی با فرکانس پایین است.

در این مقاله، نتایج با روش‌هایی مقایسه شده‌اند که بر هر دو چالش نوع پوشش و سرعت تمرکز دارند. چرا که فکر می‌کنیم در دنیای واقعی در نظر گرفتن هر دو این چالش‌ها مورد اهمیت است.

۳. روش پیشنهادی

۳-۱ نوآوری

طبق آنچه در بخش ۱ شرح داده شد، فرکانس پاسخ‌دهی سلول‌های عصبی در ویژوال کورتکس، ماهیت فازی دارد (شکل ۳). عملکرد سلول‌های عصبی را می‌توان مانند توابع تعلق^{۴۲} موجود در سیستم‌های فازی دانست که به فرکانس اسپایک‌های تولید شده توسط ورودیشان حساس هستند. بنابراین می‌توان این سازوکار را با تبدیل مقادیر روشنایی تصاویر تولید شده توسط لایه S به فرکانس اسپایک و تزریق آن‌ها به یک دسته‌بند فازی، شبیه‌سازی کرد.

مهم‌ترین نوآوری در این تحقیق افزودن لایه تولید اسپایک به مدل HMAX و استفاده از یک دسته‌بند فازی در لایه آخر است. به دلیل ماهیت مدل پیشنهادی، آن را مدل سلسله مراتبی فازی-اسپایکی^{۴۳} یا به اختصار مدل HFS می‌نامیم. شکل ۴ مراحل پردازش در مدل HFS و شکل ۵ تصویری کلی از آن و اجزای تشکیل دهنده‌اش را نمایش می‌دهد. توابع تعلق به کار رفته در این سیستم

گرادیان جهت‌دار و تصویر گوسی گیت به دست می‌آید (آرورا و همکاران، ۲۰۱۵). این نمایش به اختصار GHGI^{۴۰} نامیده می‌شود. در این روش توزیع گوسی گیت برای هر سیکل به دست آمده، سپس هیستوگرام گرادیان برای هر نقطه از تصویر محاسبه می‌شود.

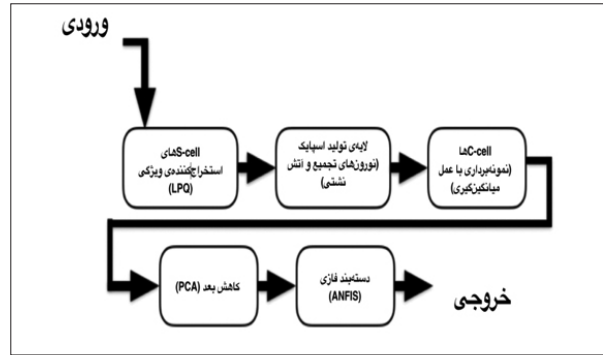
روش‌های ذکر شده در بالا، برای هر دو چالش مربوط به شناسایی گیت یعنی تغییر سرعت و تغییر پوشش، مورد ارزیابی قرار گرفته‌اند. اما برخی راه‌حل‌های ارائه شده تنها بر یکی از چالش‌ها تمرکز دارند. برای مثال در یکی از پژوهش‌ها بخش‌هایی از بدن که در حین حرکت کمترین تغییر را دارند شناسایی شده و با استفاده از آن‌ها یک سیستم مقاوم در برابر تغییر پوشش ارائه می‌شود. در این روش، کل بدن به ۸ ناحیه تقسیم شده و نواحی موثر برای هر نوع لباس، استخراج می‌شود. این نواحی در فاز شناسایی بیشتری وزن و اهمیت را خواهند داشت (حسین و همکاران، ۲۰۱۰). در روشی دیگر روی تمام نمونه‌های مجموعه گالری با پوشش عادی، عمل میانگین‌گیری انجام می‌شود و تصویر انرژی گیت عادی^{۴۱} به دست می‌آید (وایتاک و همکاران، ۲۰۱۳). این تصویر به اختصار TGEI نامیده شده و فرم عادی بدن را نمایش می‌دهد. سپس برای تشخیص نواحی تحت تاثیر پوشش یا وسایل همراه مانند کیف، نمایش GEI تصاویر آزمون از TGEI کم شده و پس از یک آستانه گذاری، نواحی تحت تاثیر پوشش مشخص می‌شوند. سپس تمامی سطرهای تشخیص داده شده از نمایش GEI حذف می‌گردند. روش‌های دیگری نیز هستند که بر چالش سرعت تمرکز دارند. در این چالش فرض می‌شود نمونه‌های موجود در مجموعه آزمون و گالری سرعت‌های متفاوتی دارند. ماکیهارا و همکاران در پژوهشی سعی کرده‌اند این مشکل را با استفاده از ویژگی‌های حوزه فرکانس و مدل تبدیل سرعت مرتفع نمایند (ماکیهارا و همکاران، ۲۰۱۲). آن‌ها در مقاله‌شان اظهار داشته‌اند که استفاده از این مدل منجر به نتایج بهتری شده است.

42- Membership Function

43- Hierarchical Fuzzy Spiking Model

40- Gradient Histogram Gaussian Image

41- Typical Gait Energy Image



شکل ۴: مراحل موجود در مدل سلسله مراتبی فازی-اسپایکی (HFS)

فازی به ازای برخی از مقادیر ورودی بالاترین پاسخ و به ازای بقیه مقادیر، پاسخی در بازه صفر و یک تولید می‌کنند.

نوآوری دیگر در این پژوهش، استفاده از توصیفگر بافت LPQ (آژانسو و همکاران، ۲۰۰۸) برای استخراج ویژگی‌ها در لایه S1 است. طبق دانش ما، تا کنون از یک توصیفگر بافت در لایه S1 استفاده نشده است.

۳-۲ ورودی

در این تحقیق، سیکل‌های گیت استخراج شده تا تعداد نمونه‌ها برای هر سوژه افزایش یابد و از بیش‌برازش^{۴۴} در فاز رده‌بندی جلوگیری شود. برای استخراج یک سیکل در گیت روش‌های متعددی مانند روش پیشنهادی در (لیتل و همکاران، ۱۹۹۸) ارائه شده است. تعریف یک سیکل که واحد اصلی برای توصیف گیت در طول راه رفتن است، بدین صورت است: فاصله بین زمانی که پاشنه یک پا زمین را لمس می‌کند تا زمانی که همان پا دوباره به زمین می‌رسد (ایواشیتا و همکاران، ۲۰۱۳).

در این تحقیق برای استخراج سیکل‌ها، از روشی بر پایه ثابت گشتاور پیوسته^{۴۵} درجه ۱ استفاده شده است (ایواشیتا و همکاران، ۲۰۱۳). ثابت‌های گشتاور پیوسته توصیفگرهایی بر پایه گشتاورها هستند که توصیفگر شکل یک شیء بوده و در برابر تغییرات مقاومند. گشتاور درجه (p+q) برای شیء O از رابطه زیر به دست می‌آید (فلاسر و همکاران، ۲۰۰۹):

$$\mu_{pq} = \sum \sum_{(x,y) \in O} (x - x_g)^p (y - y_g)^q \bar{I}(x,y) \quad (2)$$

در این رابطه، x_g و y_g مرکز شیء می‌باشند که مقدارشان از گشتاور هندسی، از رابطه زیر به دست می‌آید:

$$m_{pq} = \sum \sum_{(x,y) \in O} x^p y^q \bar{I}(x,y) \quad (3)$$

$$x_g = \frac{m_{10}}{m_{00}} \text{ و } y_g = \frac{m_{01}}{m_{00}} \quad (4)$$

مقدار گشتاور پیوسته درجه ۱ از رابطه زیر به دست می‌آید:

$$A_1 = \frac{1}{\mu_{00}^4} (\mu_{20}\mu_{02} - \mu_{11}^2) \quad (5)$$

در این روش، برای استخراج سیکل‌های راه رفتن، مقدار A_1 برای هر قاب از گیت محاسبه می‌شود. همان‌طور که در شکل ۶ مشخص است، مقدار A_1 به‌طور سینوسی تکرار می‌شود و قاب‌هایی که این مقدار در آن‌ها، بیشینه محلی نمودار است، قاب‌های مربوط به قرار گرفتن دو پا روی زمین هستند. بنابراین تصاویر مربوط به سه بیشینه محلی متوالی در نمودار، تصاویر یک سیکل از راه رفتن را تشکیل می‌دهند.

۳-۳ استخراج ویژگی و رده‌بندی

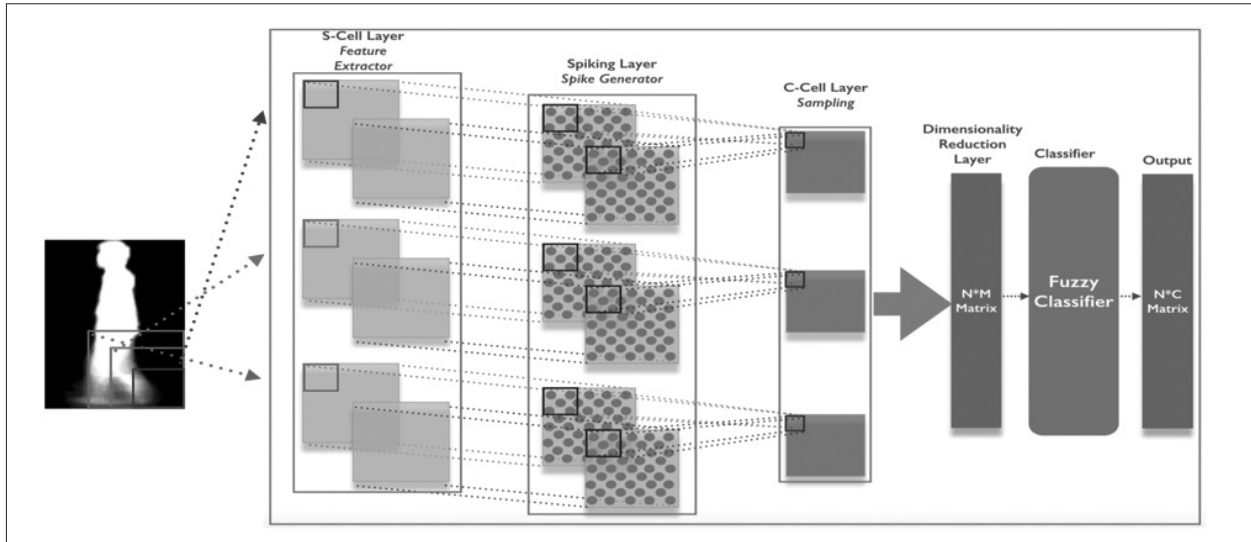
پس از استخراج سیکل‌های گیت برای هر سوژه، GEI تصاویر هر سیکل محاسبه می‌شود. هر GEI یک ورودی برای روش استخراج ویژگی ارائه شده با ۵ لایه است. در ادامه مشخصات هر لایه با جزییات بیشتر شرح داده شده است.

۳-۳-۱ لایه S-cells

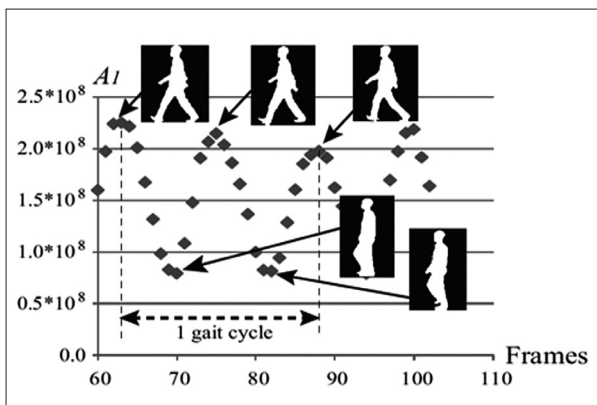
این لایه مشابه سلول‌های S-cell در ویژوال کورتکس عمل می‌کند. همان‌طور که در (سره و همکاران، ۲۰۰۷) گفته شده است، این لایه لبه‌ها و بافت تصاویر را استخراج می‌کند و می‌توان برای شبیه سازی آن از پالایه‌های گابور استفاده کرد. در این تحقیق با تاکید بر استخراج ویژگی‌های بافت، از توصیفگر LPQ استفاده شده است. LPQ به‌طور خاص

44- Overfitting

45- Affine Moment Invariant



شکل ۵: نمایشی از مدل سلسله‌مراتبی فازی-اسپایکی (HFS) و اجزای آن



شکل ۶: نمایش نمودار گشتاور پیوسته درجه ۱ برای قاب‌های مختلف گیت

جدول ۱: اندازه همسایگی‌های به کار رفته در اعمال توصیفگر بافت LPQ

اندازه همسایگی	شماره مقیاس	باند مقیاس
۳×۳	۱	۱
	۵×۵	
۷×۷	۳	۲
	۹×۹	
۱۱×۱۱	۵	۳
	۱۳×۱۳	
۱۵×۱۵	۷	۴
	۱۷×۱۷	
۱۹×۱۹	۹	۵
	۲۱×۲۱	
۲۳×۲۳	۱۱	۶
	۲۵×۲۵	

برای رده‌بندی تصاویر تار ارائه شده است. استفاده از این توصیفگر در مقایسه با پالایه گابور و توصیفگر بافت^{۴۶} LBP، برای هر دو نوع تصاویر تار و غیر تار، نرخ‌های شناسایی بالاتری به دست می‌آید (آژانسیوو و هیکیلا، ۲۰۰۸). LPQ از اطلاعات محلی فاز در حوزه فرکانس در یک همسایگی $M \times M$ استفاده می‌کند. جدول ۱ حاوی اطلاعات دقیق‌تری در مورد همسایگی‌هایی است که این توصیفگر روی آن‌ها اعمال شده است. در این جدول هر دو مقیاس مجاور، یک باند مقیاس را تشکیل می‌دهند که از آن‌ها در لایه C-cells برای نمونه‌برداری استفاده می‌شود.

۳-۳-۲ لایه تولید اسپایک

در این لایه از مدل نورونی تجمیع و انفجار نشستی یا به اختصار LIF برای تولید اسپایک استفاده شده است. صفحات نورونی به تعداد و اندازه‌های مشابه صفحات لایه S-cells، در این لایه از مدل وجود دارد. LIF به دلیل سادگی و کارآمدی یکی از مدل‌هایی است که تا به حال بیشترین استفاده از آن شده است. در LIF، حامل‌های ورودی به غشای نورونی، با بارهای ذخیره شده در یک مدار RC مدل شده‌اند (گرستنر و کیستلر، ۲۰۰۲). در نورون تجمیع و آتش، سلول هم می‌تواند توسط یک جریان خارجی و هم توسط ورودی سیناپسی سلول‌های پیش-سیناپسی

46- Local Binary Pattern

جدول ۲: پارامترهای به کار رفته شده در نورون تجمیع و آتش در این تحقیق

پارامتر	مقدار
τ_m	0.6
γ	0.3 V
u_r	0 V
R	0.5 Ω
گام زمانی	0.1ms
T	10 ms

است که این رابطه بعد از زمانی که $t = t^{(f)}$ می باشد، به اندازه Δ_{abs} ، که مدت زمان مقاومت مطلق است، دچار وقفه شده و پس از آن دوباره از سر گرفته می شود. جدول ۲ مقادیر به کار گرفته شده برای متغیرهای مذکور در نورون LIF را که به صورت تجربی به دست آمده اند نمایش می دهد.

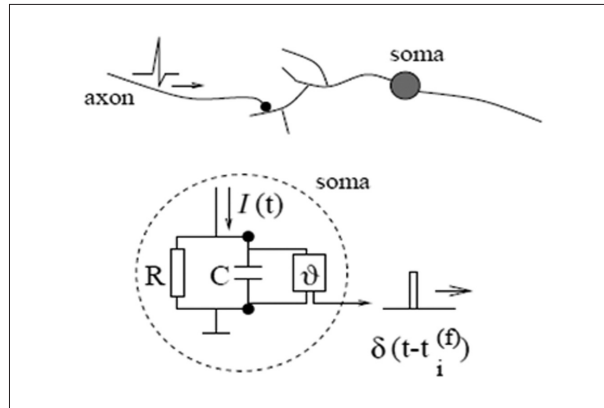
در لایه تولید اسپایک، هر نورون مقدار روشنایی پیکسل متناظرش در لایه قبل را به عنوان ورودی دریافت کرده و در مدت زمان T یک رشته اسپایک تولید می کند. بعد از اتمام زمان T، تعداد پالس های تولید شده در این مدت زمان بر T تقسیم شده و به این ترتیب فرکانس پاسخ هر نورون i در صفحه سلولی به دست می آید. به طور دقیق تر داریم.

$$\text{Frequency}_i = \frac{N_Spike_i}{T} \quad (10)$$

در این رابطه، Frequency_i ، فرکانس پاسخ نورون i ، N_Spike_i تعداد اسپایک های تولید شده توسط نورون i در زمان T و T زمان در نظر گرفته شده برای بررسی رفتار نورون است.

۳-۳-۳ لایه C-cells

این لایه وظیفه ای مشابه سلول های پیچیده در ویژوال کورتکس را داراست. سلول های پیچیده در مغز، ورودی خود را از مجموعه ای از سلول های ساده دریافت می کنند و یک نمونه برداری روی پاسخ تولید شده آن ها اعمال خواهند کرد. در این مقاله عمل نمونه برداری با استفاده از میانگین گیری روی یک همسایگی $N \times N$ در یک باند مقیاس



شکل ۷: معادل مداری مدل تجمیع و آتش نورون زیست شناسانه

تحریک شود. در این مدل از شکل دقیق پالس های خروجی صرف نظر شده و زمان رخداد هر پالس، به عنوان مشخصه اصلی در نظر گرفته می شود. شکل ۷، معادل مداری این مدل را نمایش می دهد.

در ادامه با استفاده از قانون اهم و رابطه $C = q/u$

داریم:

$$I(t) = \frac{u(t)}{R} + C \frac{du}{dt} \quad (6)$$

با ضرب رابطه بالا در R و با توجه به رابطه ثابت زمانی خازن، $\tau_m = RC$ ، به شکل استاندارد زیر می رسمیم:

$$\tau_m \frac{du}{dt} = -u(t) + RI(t) \quad (7)$$

در معادله ۷، u پتانسیل غشایی نورون و τ_m ثابت زمانی غشا می باشد. همچنین اسپایک ها هر کدام به صورت رخدادهایی تعریف می شوند که در زمان آتش کردن $t^{(f)}$ اتفاق افتاده اند. زمان های آتش کردن نورون با توجه به پتانسیل آستانه ϑ تعریف می شود:

$$t^{(f)}: u(t^{(f)}) = \vartheta \quad (8)$$

بلافاصله بعد از زمان $t^{(f)}$ ، پتانسیل غشایی به مقدار u_r که کمتر از ϑ است، تنزل پیدا می کند.

$$\lim_{t \rightarrow t^{(f)}; t > t^{(f)}} u(t) = u(r) \quad (9)$$

هنگامی که $t > t^{(f)}$ رابطه ۹ تکرار می شود تا زمانی که پتانسیل غشایی مجدداً حد آستانه را رد کند. لازم به ذکر

در فاز آموزش این شبکه از روش گرادینان مزدوج افزایشی^{۵۰} (ستیسلی و بارکانا، ۲۰۱۰) استفاده شده است. این روش برای داده‌های با بعد بالا مناسب‌تر از روش‌های بهینه‌سازی بر پایه مشتق درجه دوم است. برای مقدار دهی اولیه قوانین فازی نیز از الگوریتم kmeans استفاده شده است (هارتیگان و ونگ، ۱۹۷۹). در اینجا لازم است قبل از فاز آموزش مقدار K که در اینجا تعداد توابع تعلق برای هر رده است، به‌عنوان ورودی به سیستم داده شود.

۳-۳-۶ همجوشی^{۵۱}

در این روش به دلیل استخراج سیکل‌ها برای هر نمونه، در انتها برای هر نمونه آزمون بیش از یک برچسب خروجی خواهیم داشت. بنابراین به روش همجوشی برای برچسب‌های مربوط به یک نمونه، نیاز داریم. در این تحقیق، از روش همجوشی اکثریت استفاده شده است. در مواردی که بیش از یک برچسب با تعداد اکثریت^{۵۲} وجود دارد، مجموع احتمالات تعلق^{۵۳} برای هر برچسب محاسبه می‌شود. در این حالت برچسب نهایی برچسبی با بیشترین مجموع احتمالات خواهد بود.

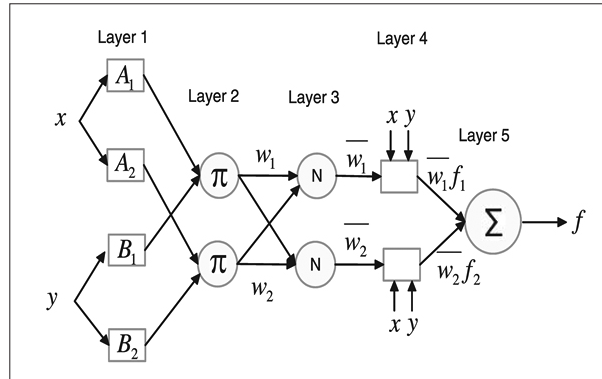
۴. نتایج

همان‌طور که در بخش ۲ عنوان شد، برای ارزیابی سیستم از مسئله شناسایی گیت استفاده شده و چالش‌های تغییر سرعت و تغییر نوع پوشش انتخاب شده‌اند. در ادامه، نتایج روی مجموعه داده OU-ISIR و OU-ISIR Treadmill A و B نیز Treadmill B که به مجموعه داده‌های ترمیل A و B نیز معروفند، ذکر شده است.

۴-۱ مجموعه داده‌ها

مجموعه داده ترمیل A و B، مجموعه داده‌ای بزرگ حاوی تصاویر راه رفتن افراد مختلف با شرایط متفاوت از نظر نوع پوشش، سرعت راه رفتن و زاویه دید است. در تهیه این مجموعه داده، از ۲۵ دوربین برای عکس‌برداری از

50-Scaled onjugate gradient
51-Fusion
52-Majority
53-Membership Probabilities



شکل ۸: نمایش ساختار یک شبکه ANFIS با دو ورودی و یک خروجی (تاگای و سوگنو، ۱۹۸۵)

صورت گرفته است. مشخص است به دلیل عدم تساوی اندازه ماتریس‌های دو مقیاس مجاور، لازم است ابتدا عملیات لایه‌گذاری روی ماتریس کوچک‌تر صورت بگیرد.

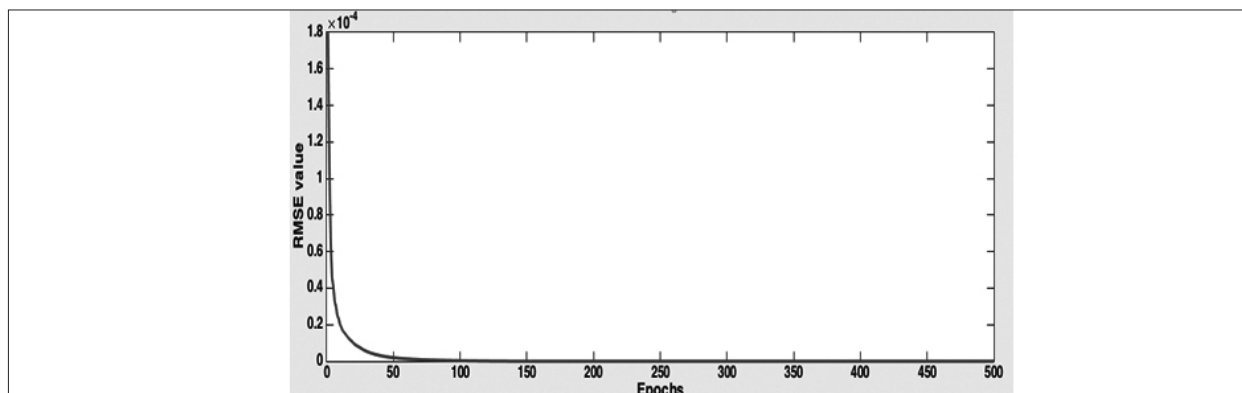
۳-۳-۴ لایه کاهش بعد

یکی از مشکلاتی که در کار با مجموعه داده‌هایی با بعد بسیار زیاد وجود دارد این است که در بسیاری موارد همه متغیرهای اندازه‌گیری شده، در فاز شناسایی مفید نیستند (سرزانو و همکاران، ۲۰۱۴). با این‌که روش‌هایی وجود دارند که از نظر محاسباتی بسیار سنگین بوده و با استفاده از داده‌هایی با بعد بسیار زیاد به دقت بالایی در نتایج می‌رسند، اما در سایر مدل‌ها معمولاً بعد داده‌های اصلی برای هرگونه مدل‌سازی، کاهش پیدا می‌کند. بدین منظور از روش‌های کاهش بعد نظیر PCA^{۴۷} و LDA^{۴۸} (بله‌مر و همکاران، ۱۹۹۷) استفاده می‌شود. در این تحقیق از روش کاهش بعد PCA استفاده شده است.

۳-۳-۵ دسته‌بندی فازی

در این لایه از یک سیستم استنتاج عصبی فازی تطبیق‌پذیر^{۴۹} یا به اختصار ANFIS استفاده شده است (جانگ، ۱۹۹۳). ANFIS یک سیستم استنتاج فازی در چارچوب شبکه‌های انطباقی است که در این مقاله، یک نگاشت ورودی-خروجی با استفاده از داده‌های ورودی و خروجی فاز آموزش فراهم می‌سازد. شکل ۸ ساختار یک ANFIS با دو ورودی و یک خروجی را نشان می‌دهد.

47-Principal Component Analysis
48-Linear Discriminant Analysis
49-Adaptive neuro Fuzzy Inference System



شکل ۹: الف) نمودار تغییرات خطا در فاز آموزش روی مجموعه داده ترمیل A در ۵۰۰ اجرا

به ترتیب برابر با $۱۰^{-۱۵} \times ۳/۶۴۱۲$ و $۱۰^{-۱۴} \times ۱/۶۵۹۹$ می باشد.

برای آزمایش سیستم از مجموعه آزمون با همان تعداد رده در فاز آموزش و ۸۵۶ نمونه مختلف استفاده شده است.

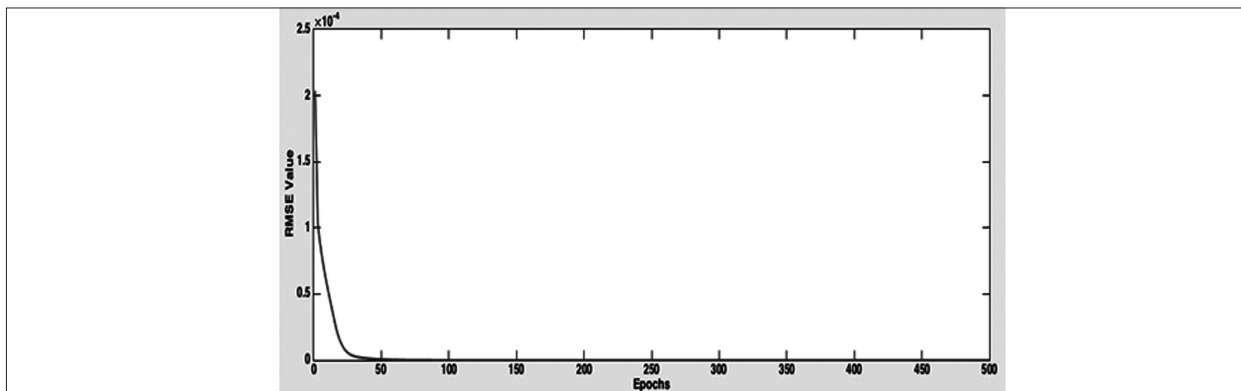
۴-۶ تغییرات سرعت

برای نشان دادن کارایی مدل پیشنهادی، مقایسه‌ای با چندین روش موجود با استفاده از معیار مشخصه تطابق جمع‌شونده یا به اختصار CMC استفاده شده است. CMC نموداری است که نرخ شناسایی را در k رتبه اول مشخص می‌کند. در این تحقیق سرعت‌های ۴ کیلومتر بر ساعت، ۵ کیلومتر بر ساعت و ۶ کیلومتر بر ساعت در نظر گرفته شده است. داده‌های مربوط به رده با سرعت راه رفتن ۵ کیلومتر بر ساعت به عنوان مجموعه آموزش و داده‌های مربوط به رده با سرعت راه رفتن ۴ کیلومتر بر ساعت، ۵ کیلومتر بر ساعت و ۶ کیلومتر بر ساعت، به عنوان مجموعه آزمون در نظر گرفته شده است. شکل ۱۰-الف، منحنی CMC را برای مدل پیشنهادی HFS و سایر روش‌ها، با مجموعه آزمون با سرعت ۴ کیلومتر بر ساعت نمایش می‌دهد. شکل‌های ۱۰-ب و ۱۰-پ نیز مربوط به مجموعه داده‌های آزمون با سرعت ۵ کیلومتر و ۶ کیلومتر بر ساعت هستند. در این نمودارها، نتایج مربوط به روش‌های GENI و GII-SF, GII-EF, DFT, EnDFT, GEI (آرورا و همکاران، ۲۰۱۵) گرفته شده است.

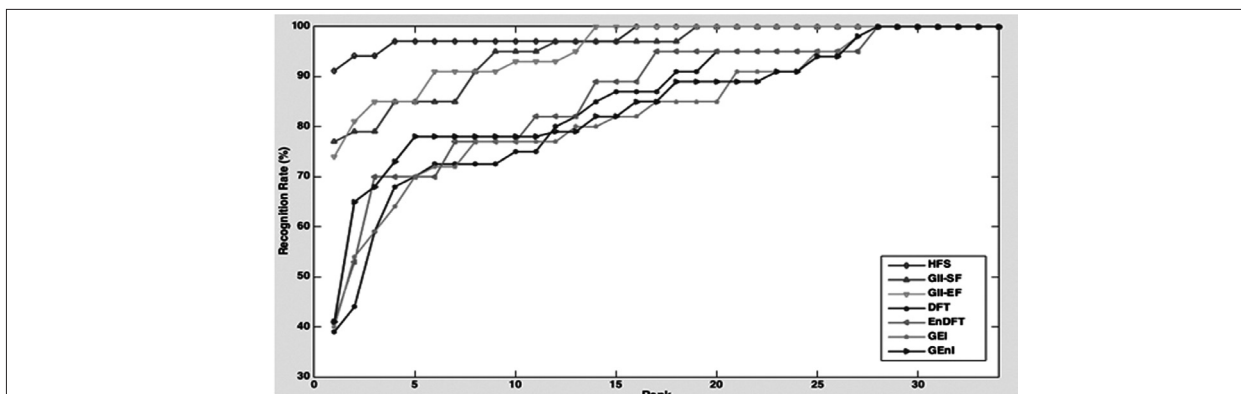
شخص در حال راه رفتن روی ترمیل استفاده شده است. مجموعه داده ترمیل A از ۳۴ رده با سرعت‌های حرکت از ۲ کیلومتر بر ساعت تا ۱۰ کیلومتر بر ساعت با فاصله ۱ کیلومتر بر ساعت، تشکیل شده است. نمونه‌ها برای رسیدن به سرعت‌های ۲ کیلومتر بر ساعت تا ۷ کیلومتر بر ساعت، راه رفته‌اند و برای رسیدن به سرعت‌های ۸ تا ۱۰ کیلومتر بر ساعت دویده‌اند (و یا تند راه رفته‌اند^(۴)). این مجموعه داده بزرگترین مجموعه داده موجود برای کار بر روی تغییر سرعت راه رفتن است.

در این پژوهش، مدل پیشنهادی با تصاویر مجموعه گالری با سرعت m کیلومتر بر ساعت آموزش دیده، سپس با تصاویر مجموعه آزمون با سرعت n کیلومتر بر ساعت ($n \neq m$)، مورد ارزیابی قرار گرفته است.

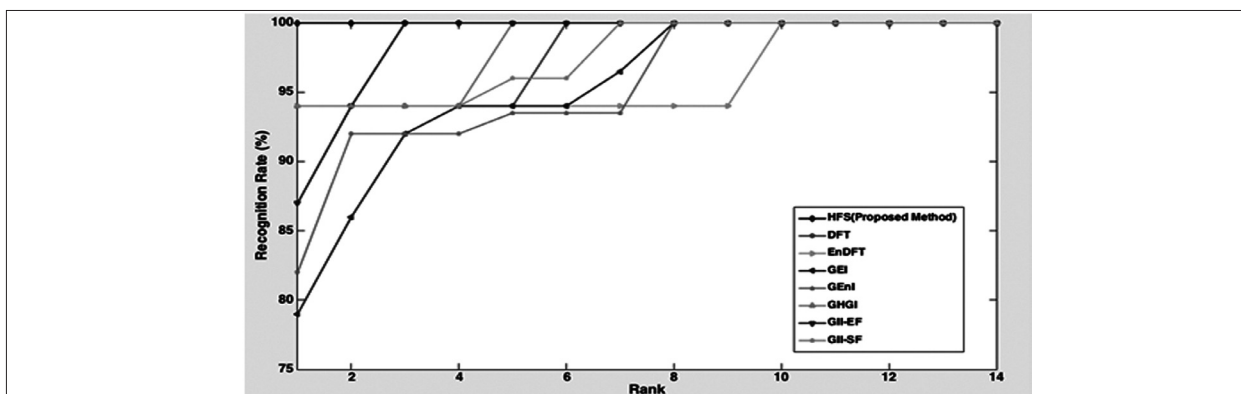
مجموعه داده B در مقایسه با مجموعه داده‌هایی که تا به حال جمع‌آوری شده‌اند، از تنوع بیشتری در نوع پوشش برخوردار است. این مجموعه داده ۶۸ رده با ۳۲ ترکیب مختلف از انواع پوشش‌ها را داراست. در این تحقیق مجموعه آموزشی با ۲۰ رده که دارای پوشش‌های متفاوتی هستند، برای آموزش PCA در فاز کاهش بعد، به کار گرفته شده است. همچنین مجموعه گالری متشکل از ۴۸ رده در فاز آموزش مورد استفاده قرار گرفته است. شکل ۹-الف و شکل ۹-ب نمودار خطای سیستم (RMSE) را به ازای هر بار اجرا در فاز آموزش برای مجموعه داده A و B نمایش می‌دهد. آخرین مقدار RMSE در این نمودارها



شکل ۹: ب) نمودار تغییرات خطا در فاز آموزش روی مجموعه داده ترمبیل B در ۵۰۰ اجرا



شکل ۱۰: الف) نمودار منحنی CMS با استفاده از روش‌های مختلف روی مجموعه داده ترمبیل A، مجموعه آزمون با سرعت ۴ کیلومتر بر ساعت می‌باشد.

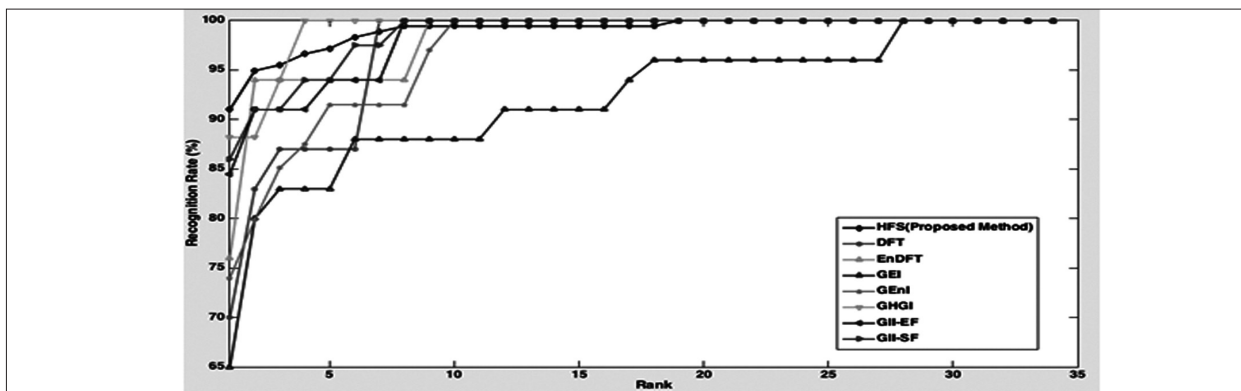


شکل ۱۱: ب) نمودار منحنی CMS با استفاده از روش‌های مختلف روی مجموعه داده ترمبیل A، مجموعه آزمون با سرعت ۵ کیلومتر بر ساعت می‌باشد.

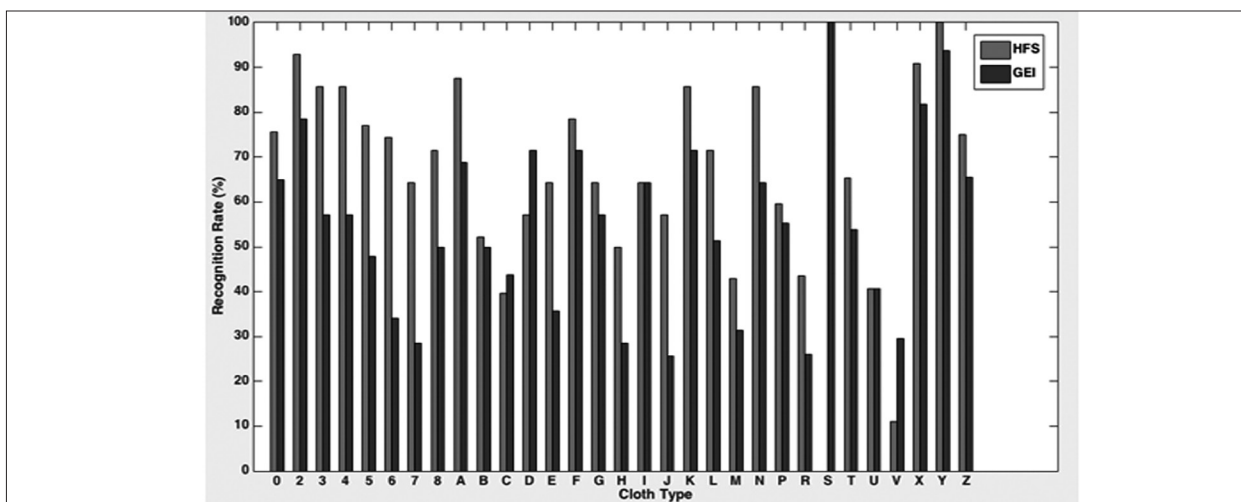
۳-۴ تغییرات در نوع پوشش

مجموعه داده ترمبیل B، مخصوص تغییرات پوشش است. نتایج به دست آمده روی مدل پیشنهادی در مرحله اول با نتایج روش پایه یعنی روش استفاده از GEI به عنوان ویژگی، مورد مقایسه قرار گرفته است. همان‌طور که در شکل ۱۱-الف مشخص است، نرخ شناسایی برای بیشتر انواع پوشش‌ها برای روش ارائه شده بهتر از نتایج

روش GEI می‌باشد. همان‌طور که قبلاً گفته شد، ورودی مدل ویژگی GEI برای هر سیکل است. در اینجا برای مقایسه عادلانه در روش GEI هم، ویژگی GEI برای هر سیکل راه رفتن استخراج شده است. شکل ۱۱-ب منحنی CMC را برای هر دوی این روش‌ها نمایش می‌دهد. طبق این منحنی، نتایج در همه رتبه‌ها، برای روش ارائه شده به طرز چشمگیری بهتر از نتایج روش GEI می‌باشد.



شکل ۱۰: پ) نمودار منحنی CMS با استفاده از روش‌های مختلف روی مجموعه داده ترمیم A، مجموعه‌ی آزمون با سرعت ۶ کیلومتر بر ساعت می‌باشد.



شکل ۱۱: الف) نمودار نرخ شناسایی در رتبه ۱ برای انواع پوشش‌ها در مجموعه داده ترمیم B با استفاده از دو روش HFS و GEI

جدول ۳: مقایسه نتایج روش‌های موجود و روش‌های ارائه شده روی مجموعه داده ترمیم B

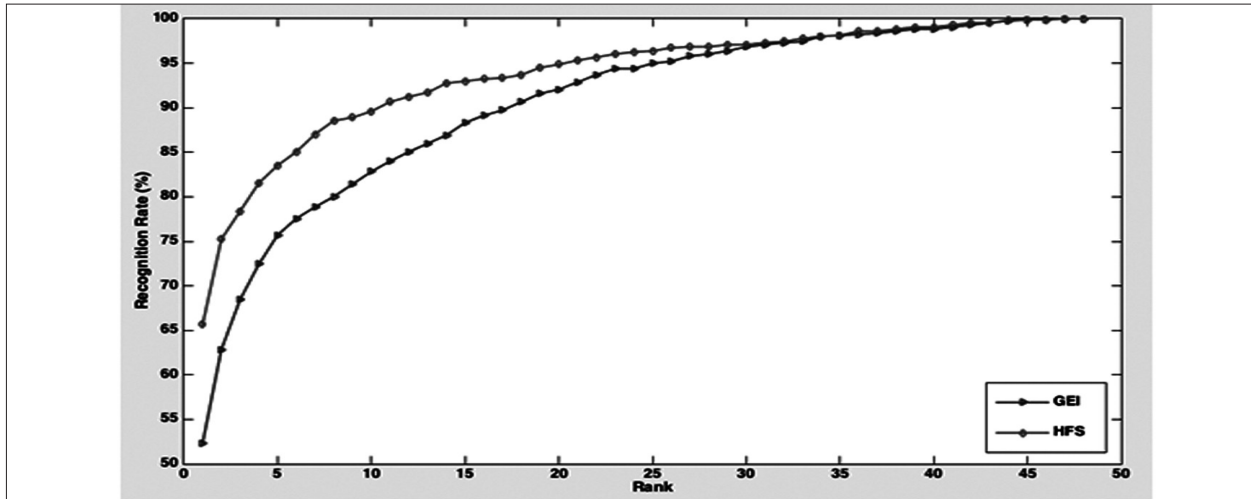
	نرخ شناسایی در رتبه ۵ (%)	نرخ شناسایی در رتبه ۱ (%)
GEnI	75.6	47.3
DFT	73.2	46.6
EnDFT	77.4	52.0
GII-EF	82.6	57.4
GII-SF	85.1	61.2
GHGI	77.0	62.5
HFS	83.53	65.65

لایه از ویژگی‌ها استفاده می‌کنیم، سودمندتر است یا خیر. برای پاسخ به این سوال، نرخ‌های شناسایی با استفاده از ویژگی‌های لایه ۱ و ۲، با نرخ شناسایی حاصل از ویژگی‌های لایه ۳، با استفاده از رده‌بند فازی، روی مجموعه داده B محاسبه شده است. لازم به ذکر است روی

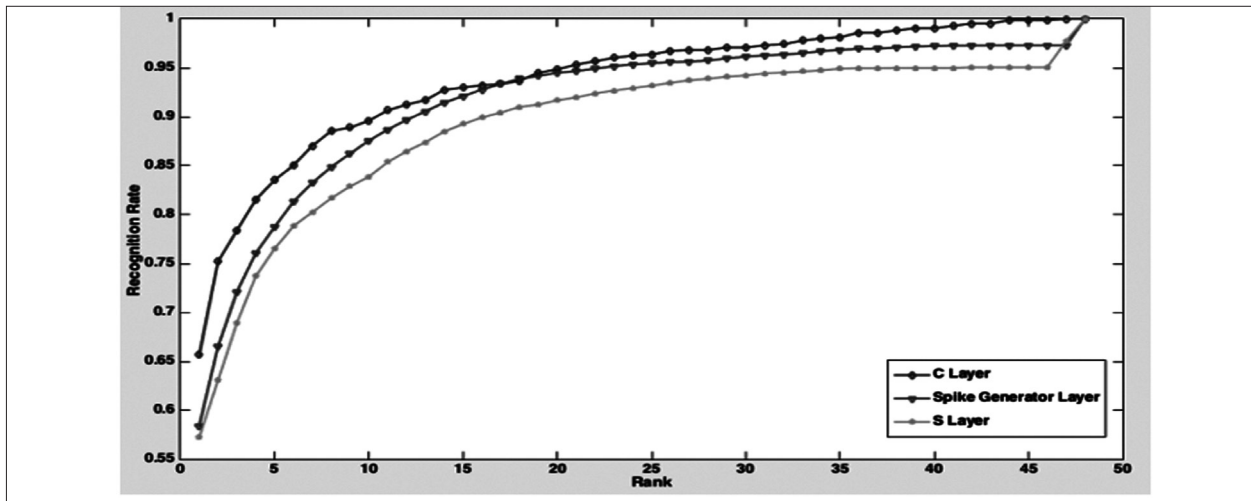
در این تحقیق نرخ‌های شناسایی در رتبه ۱ و رتبه ۵ برای مقایسه نتایج HFS با دیگر روش‌ها انتخاب شده است. جدول ۳ نرخ‌های شناسایی روش‌های DFT، EnDFT، GEnI، GII-EF، GII-SF، GHGI و مدل HFS را در رتبه ۱ و رتبه ۵ نمایش می‌دهد. بالاترین نرخ شناسایی در رتبه ۱ برای مدل پیشنهادی است. در رتبه ۵، بالاترین نرخ شناسایی برای روش GII-SF است و مدل HFS در جایگاه دوم قرار دارد. به‌طور کلی این نتایج نشان می‌دهند که کارایی روش ما در رتبه ۱ بیشتر از روش پایه و آخرین روش‌های ارائه شده در شناسایی گیت می‌باشد.

۴-۴ بررسی اثر ساختار سلسله مراتبی مدل

در این بخش این پرسش را بررسی می‌کنیم که اساسا اعمال ۳ لایه برای به دست آوردن ویژگی با استفاده از ساختاری سلسله مراتبی، نسبت به حالتی که فقط از یک



شکل ۱۱: ب) منحنی CMS روش ارائه شده و روش پایه GEI روی مجموعه داده ترمیم B



شکل ۱۲: الف) نمایش منحنی CMC بر روی مجموعه داده ترمیم B در صورت استفاده از ویژگی‌های لایه ۱، لایه ۲ و لایه ۳

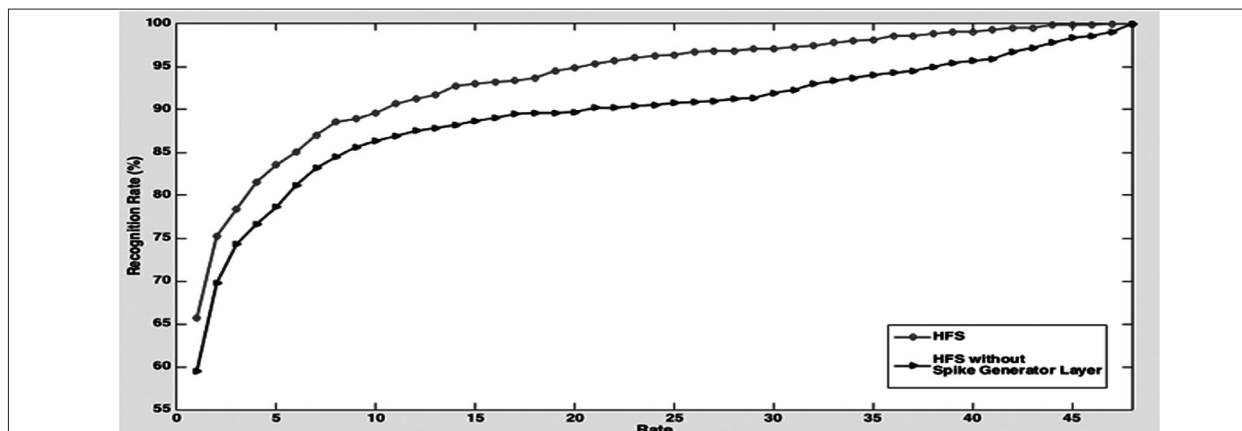
داده ترمیم B مورد بررسی قرار گرفت. همان‌طور که در شکل ۱۲-ب قابل مشاهده است، با اعمال لایه تولید اسپایک، نرخ شناسایی در رتبه ۱ حدود ۶ درصد افزایش یافته است. بنابراین وجود این لایه در دسترسی به ویژگی‌هایی با قدرت تمایز دهندگی بالاتر، موثر می‌باشد.

۵. جمع‌بندی و کارهای آینده

در این تحقیق بر ارائه راه‌حلی برای مسئله شناسایی گیت با استفاده از مدل سلسله مراتبی فازی-اسپایکی، تمرکز شد. مدل پیشنهادی که در بخش ۳ این مقاله، شرح داده شد، به نتایج بهتری بر روی مجموعه داده‌های مورد استفاده، دست یافت. به نظر می‌رسد استفاده از توصیفگر

تمام ویژگی‌ها، PCA اعمال شده است. منحنی CMS این نتایج در شکل ۱۲-الف قابل مشاهده است. همان‌طور که مشخص است، استفاده از ویژگی‌های لایه‌های بالاتر، منجر به نرخ شناسایی بیشتر و منحنی CMS برتر شده است. به طوری که نتایج حاصل از ویژگی‌های لایه سوم، نسبت به نتایج حاصل از ویژگی‌های لایه اول، حدود ۱۰ درصد بهبود پیدا کرده است. این موضوع نشان دهنده مفید بودن ساختار سلسله مراتبی در این مدل است و نشان می‌دهد ویژگی‌ها با عبور از هر لایه قدرت تمایز دهندگی بالاتری پیدا می‌کنند.

در آزمایشی دیگر، تاثیر لایه تولید اسپایک که در این پژوهش به مدل HMAX اضافه شده، با استفاده از مجموعه



شکل ۱۲: (ب) منحنی CMC بر روی مجموعه داده B در صورت استفاده و عدم استفاده از لایه تولید اسپایک

classifier training. *Soft computing*, 14(4), 365-378.

[6] Cunado, D., Nixon, M. S., & Carter, J. N. (1997, March). Using gait as a biometric, via phase-weighted magnitude spectra. In *International Conference on Audio-and Video-Based Biometric Person Authentication* (pp. 93-102). Springer Berlin Heidelberg.

[7] Cunado, D., Nixon, M. S., & Carter, J. N. (2003). Automatic extraction and description of human gait models for recognition purposes. *Computer Vision and Image Understanding*, 90(1), 1-41.

[8] Fukushima, K., Miyake, S., & Ito, T. (1983). Neocognitron: A neural network model for a mechanism of visual pattern recognition. *IEEE Transactions on Systems, Man, and Cybernetics*, (5), 826-834.

[9] Gerstner, W., & Kistler, W. M. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge university press.

[10] Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), 100-108.

[11] Hossain, M. A., Makihara, Y., Wang, J., & Yagi, Y. (2010). Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control. *Pattern Recognition*, 43(6), 2281-2291.

[12] Hubel, D. H. (1995). *Eye, Brain, and Vision* (Scientific American Library, No 22).

[13] Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1), 106-154.

[14] Iwashita, Y., Uchino, K., & Kurazume, R. (2013). Gait-based person identification robust to changes in appearance. *Sensors*, 13(6), 7884-7901.

Flusser J., Suk, T. and Zitová, B. (2009) *Front Matter: Wiley Online Library*.

[15] Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3), 665-685.

[16] Johnson, A. Y., & Bobick, A. F. (2001, June). A multi-view method for gait recognition using static body parameters. In *International Conference on Audio-and Video-Based Biometric Person Authentication* (pp. 301-311). Springer Berlin

بافت LPQ برای استخراج ویژگی‌های مشابه بافت در تصاویر GEI، نسبت به استفاده از پالایه‌های گابور که در مدل HMAX استفاده شده است، منجر به نتایج بهتری می‌شود. بنابراین به نظر می‌رسد می‌توان روی استفاده از توصیفگرهای بافت دیگر به جای پالایه گابور در مدل HMAX، در انواع شناسایی‌ها تحقیقات بیشتری صورت داد. همچنین مشخص شد افزودن لایه تولید اسپایک و دسته‌بند فازی نیز به بهبود نتایج کمک می‌کند. تحقیق در مورد اثر دقیق لایه تولید اسپایک روی داده‌ها و همچنین افزودن چنین لایه‌ای به مدل‌هایی دیگر نظیر CNN و نتوکانیترن، مفید خواهد بود.

مراجع

[1] Abdel-Hamid, O., Mohamed, A. R., Jiang, H., & Penn, G. (2012, March). Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on* (pp. 4277-4280). IEEE.

[2] Arora, P., Hanmandlu, M., & Srivastava, S. (2015). Gait based authentication using gait information image features. *Pattern Recognition Letters*, 68, 336-342.

[3] Arora, P., Srivastava, S., Arora, K., & Bareja, S. (2015). Improved Gait Recognition Using Gradient Histogram Gaussian Image. *Procedia Computer Science*, 58, 408-413.

[4] Bashir, K., Xiang, T., & Gong, S. (2009). Gait recognition using gait entropy image. Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, 19(7), 711-720.

[5] Cetişli, B., & Barkana, A. (2010). Speeding up the scaled conjugate gradient algorithm and its application in neuro-fuzzy

- [26] Rokanujjaman, M., Islam, M. S., Hossain, M. A., Islam, M. R., Makihara, Y., & Yagi, Y. (2015). Effective part-based gait identification using frequency-domain gait entropy features. *Multimedia Tools and Applications*, 74(9), 3099-3120.
- [27] Rüdiger, P. J. F. Estimating Simple and Complex Cell Receptive Fields from Natural Image Stimuli and 2-Photon Imaging Recordings of the V1 in Ferrets.
- [28] Serre, T., & Riesenhuber, M. (2004). Realistic modeling of simple and complex cell tuning in the HMAX model, and implications for invariant object recognition in cortex (No. AI-MEMO-2004-017). MASSACHUSETTS INST OF TECH CAMBRIDGE COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LAB.
- [29] Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., & Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. *IEEE transactions on pattern analysis and machine intelligence*, 29(3).
- [30] Sorzano, C. O. S., Vargas, J., & Montano, A. P. (2014). A survey of dimensionality reduction techniques. *arXiv preprint arXiv:1403.2877*.
- [31] Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE transactions on systems, man, and cybernetics*, (1), 116-132.
- [32] Whytock, T. P., Belyaev, A., & Robertson, N. M. (2013, July). Towards robust gait recognition. In *International Symposium on Visual Computing* (pp. 523-531). Springer Berlin Heidelberg.
- [33] Yam, C., Nixon, M. S., & Carter, J. N. (2002). Gait recognition by walking and running: a model-based approach.
- [34] Yu, S., Tan, D., & Tan, T. (2006, August). A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on* (Vol. 4, pp. 441-444). IEEE.
- Heidelberg.
- [17] Kruger, N., Janssen, P., Kalkan, S., Lappe, M., Leonardis, A., Piater, J.,... & Wiskott, L. (2013). Deep hierarchies in the primate visual cortex: What can we learn for computer vision?. *IEEE transactions on pattern analysis and machine intelligence*, 35(8), 1847-1871.
- [18] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- [19] Little, J., & Boyd, J. (1998). Recognizing people by their gait: the shape of motion. *Videre: Journal of computer vision research*, 1(2), 1-32.
- [20] Makihara, Y., Mannami, H., Tsuji, A., Hossain, M. A., Sugiura, K., Mori, A., & Yagi, Y. (2012). The OU-ISIR gait database comprising the treadmill dataset. *IPSJ Transactions on Computer Vision and Applications*, 4, 53-62.
- [21] Makihara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T., & Yagi, Y. (2006, May). Gait recognition using a view transformation model in the frequency domain. In *European Conference on Computer Vision* (pp. 151-163). Springer Berlin Heidelberg.
- [22] Man, J., & Bhanu, B. (2006). Individual recognition using gait energy image. *IEEE transactions on pattern analysis and machine intelligence*, 28(2), 316-322.
- [23] Nash, J. M., Carter, J. N., & Nixon, M. S. (1997). Dynamic feature extraction via the velocity Hough transform. *Pattern Recognition Letters*, 18(10), 1035-1047.
- [24] Ojansivu, V., & Heikkilä, J. (2008, July). Blur insensitive texture classification using local phase quantization. In *International conference on image and signal processing* (pp. 236-243). Springer Berlin Heidelberg.
- [25] Ojansivu, V., Rahtu, E., & Heikkilä, J. (2008). Rotation invariant blur insensitive texture analysis using local phase quantization. In *Proceedings of 19th International Conference on Pattern Recognition* (Vol. 4).

علوم رایانشی / زمستان ۱۳۹۷

جدیدترین کتاب
از انتشارات انجمن انفورماتیک ایران
منتشر شد!

کار عمیق

برای تهیه کتاب با دفتر انجمن انفورماتیک ایران
تماس بگیرید ۶۶۴۱۲۸۶۱

چاپ سوم