





دانشکده پردیس خوارزمی

رشته مهندسی کامپیوتر، گرایش هوش مصنوعی

رساله دکتری

طبقه‌بندی تصاویر سیستم‌های کپی‌مکانه بر مبنای یادگیری عمیق

نگارنده

سید محمدرضا هاشمی

استاد راهنما

دکتر حمید حسن‌پور

اساتید مشاور

Dr Tao Tan

دکتر احسان اله کوزه‌گر

تابستان ۱۴۰۰

شماره: ۲۴۳۷۹/م.الف
تاریخ: ۱۸/۷/۱۴۰۰
ویرایش:

باسمه تعالی



مدیریت تحصیلات تکمیلی

فرم شماره ۱۱: صورت جلسه نهایی دفاع از رساله دکتری (Ph.D)
(تعیین درجه نمره رساله برای ورودی‌های ۹۵ و سال‌های بعد)

بدینوسیله گواهی می‌شود آقای سید محمدرضا هاشمی دانشجوی دکتری رشته کامپیوتر- گرایش هوش مصنوعی به شماره دانشجویی ۹۶۰۱۵۷۵ ورودی ۹۶ ماه ۷ سال ۱۳۹۶ در تاریخ ۱۴۰۰/۶/۱۳ از رساله نظری / عملی ☐ خود با عنوان: طبقه‌بندی تصاویر سیستم‌وسکوپی مثانه بر مبنای یادگیری عمیق
دفاع و با اخذ نمره ۱۷.۴۶ به درجه : خوب نائل گردید.

الف) درجه عالی: نمره ۲۰-۱۹ ☐ ب) درجه خیلی خوب: نمره ۱۸/۹۹-۱۸ ☐
ج) درجه خوب: نمره ۱۷/۹۹-۱۶ ☒ د) مردود: کمتر از ۱۶ ☐

ردیف	هیئت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱	استاد راهنمای اول	حمید حسن پور	استاد	
۲	استاد راهنمای دوم	-	-	-
۳	استاد مشاور اول	احسان اله کوزه گر	استادیار	
۴	استاد مشاور دوم	تائو تن	استادیار	-
۵	استاد مدعو خارجی	گوروش کیانی	دانشیار	
۶	استاد مدعو داخلی	حسین خسروی	دانشیار	
۷	استاد مدعو داخلی	سید مسعود میررضایی	استادیار	
۸	نماینده تحصیلات تکمیلی دانشگاه	ایمان آقاییان	دانشیار	

مدیر محترم تحصیلات تکمیلی دانشگاه:

ضمن تأیید مراتب فوق مقرر فرمائید اقدامات لازم در خصوص انجام مراحل دانش آموزی آقای محمد رضا هاشمی بعمل آید.

نام و نام خانوادگی رئیس دانشکده:
تاریخ و امضاء و مهر دانشکده:

تقدیر و تشکر

اکنون که به یاری پروردگار و یاری و راهنمایی اساتید بزرگ موفق به پایان این رساله شده‌ام وظیفه خود دانسته که نهایت سپاسگزاری را از تمامی عزیزانی که در این راه به من کمک کرده‌اند را به عمل آورم :

در آغاز از استاد بزرگ و دانشمند جناب آقای پروفسور حمید حسن پور که راهنمایی این رساله را به عهده داشته‌اند کمال تشکر را دارم.

از جناب آقایان دکتر تائو تن و دکتر احسان کوزه گر که اساتید مشاور این رساله بوده‌اند نیز قدردانی می‌نمایم.

از داوران گرامی آقایان دکتر کوروش کیانی، دکتر حسین خسروی، دکتر سید مسعود میررضایی که زحمت داوری و تصحیح این رساله را به عهده داشتند کمال سپاس را دارم.

از پدر و مادر عزیز و مهربانم به خاطر زحماتی که در طول زندگی همواره برای پیروزی و شادکامی من به جان خریدند، تشکر می‌کنم. از همسر عزیزم به خاطر همکاری‌ها و دلگرمی‌هایش متشکرم.

و در پایان این رساله را تقدیم می‌کنم به:

همسر و دختر عزیزم (سیلوانا جان)

که مرا در این راه استوار و ثابت قدم نمودند.

تعهد نامه

اینجانب **سید محمد رضا هاشمی** دانشجوی دوره دکتری رشته کامپیوتر گرایش هوش مصنوعی دانشگاه صنعتی شاهرود نویسنده پایان نامه با موضوع **طبقه‌بندی تصاویر سیستم‌توسکوپی ماثانه بر مبنای یادگیری عمیق** تحت راهنمایی دکتر حمید حسن پور متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان‌نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می‌باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان‌نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا یافته‌های آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاقی انسانی رعایت شده است.

تاریخ: ۱۴۰۰/۰۷/۱۹

امضا دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود. استفاده از اطلاعات و نتایج موجود در پایان‌نامه بدون ذکر مرجع مجاز نمی‌باشد.

چکیده

سرطان مثانه به عنوان نهمین بیماری شایع در جهان شناخته می‌شود و از تصاویر سیستم‌سکوپی برای تشخیص زودهنگام این نوع از سرطان توسط پزشکان و متخصصان استفاده می‌شود. برای تشخیص این نوع از سرطان توسط تصاویر سیستم‌سکوپی با وجود شیوع بسیار بالا در سطح جهان متاسفانه سیستم هوشمندی با دقت قابل قبول در حوزه پردازش تصاویر پزشکی ارائه نشده است. از مهم‌ترین چالش‌های موجود در این نوع از تصاویر می‌توان به عدم وجود مجموعه داده استاندارد با تعداد مناسب، شرایط متفاوت تصویر برداری از لحاظ شدت روشنایی، اندازه و زاویه تصاویر استخراج شده و همچنین پیچیدگی بالای تصاویر که در بسیاری از حالت‌ها امکان تشخیص دقیق توسط متخصصان (پزشکان) وجود ندارد و نیاز به آزمایشات و هزینه بیشتر می‌باشد اشاره نمود. در همین راستا طراحی سیستمی هوشمند براساس استخراج ویژگی‌های مناسب مبتنی بر یادگیری ماشین در این نوع از کاربردها حیاتی می‌باشد.

در این رساله دو روش پیشنهادی کلی برپایه ویژگی‌های سطح پایین و سطح بالای معنایی معرفی شده است. در روش پیشنهادی اول، از شبکه عصبی چندلایه پرسپترون بهبود یافته بر اساس نرخ یادگیری تطبیقی و وزن‌های اولیه استخراج شده بر پایه الگوریتم جستجو تکاملی ژنتیک برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه استفاده شده است. این روش سبب بهبود سرعت همگرایی و دقت طبقه‌بند شده است. برای استخراج ویژگی در این نوع از تصاویر از الگوهای باینری محلی با روش تجزیه و تحلیل مولفه‌های اصلی به منظور کاهش ویژگی استفاده شده است. نتایج بدست آمده از روش پیشنهادی اول نشان‌دهنده عملکرد بهتر نسبت به روش‌های دیگر بر پایه ویژگی‌های سنتی یا سطح پایین بوده است. ولی همچنان این نتایج برای کاربردهای واقعی مناسب نمی‌باشند در همین راستا در ادامه، یک روش نوین برپایه یادگیری عمیق پیچشی برای طبقه‌بندی تصاویر سیستم‌سکوپی ارائه شده است.

با توجه به تعداد محدود تصاویر سیستم‌سکوپی، برای استخراج ویژگی در روش پیشنهادی دوم از یادگیری انتقال بر اساس معماری‌های متفاوت شبکه عصبی پیچشی استفاده شد. روش پیشنهادی به گونه‌ای طراحی شد که نسبت به تغییرات شرایط تصویربرداری مستقل باشد. در رویکرد یادگیری انتقال، آخرین لایه‌های یک شبکه از پیش آموزش دیده، که قبلاً روی یک مجموعه داده بزرگ مانند ImageNet آموزش دیده‌اند، استفاده شد. برای جلوگیری از مشکل ابعاد بالا و همچنین کاهش زمان و پیچیدگی فضای ویژگی، از تعداد ویژگی‌ها با استفاده از روش‌های کاهش بعد کاسته شد. سرانجام، یک طبقه‌بند ترکیبی برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه پیشنهاد شد. در روش پیشنهادی برای ترکیب از طبقه‌بندهای شناخته شده با استفاده از روش آرای اکثریت وزنی استفاده شد. روش پیشنهادی نتایج امیدوار کننده مربوط به طبقه‌بندی این نوع از تصاویر سیستم‌سکوپی را نشان داد. در ادامه برای بهبود روش‌های ارائه شده، از روش بخش‌بندی برای انتخاب نواحی با ارزش از نگاشت‌های ویژگی مربوط به لایه‌های پیچشی استفاده شد. ویژگی‌های سطح بالا و نواحی مربوط به این ویژگی‌ها بر اساس روش بهینه‌سازی ازدحام ذرات بهبود یافته و اتسو استخراج شد. سپس، با استفاده از روش PCA، ویژگی استخراج شده در نواحی کاهش یافت. در آخر، طبقه‌بند ترکیبی مبتنی بر رای گیری اکثریت با بیزین گوسی و ماشین بردار پشتیبان استفاده شد.

روش‌های پیشنهادی بر روی پایگاه داده معرفی شده در این پژوهش از تصاویر پزشکی سیستم‌سکوپی در مرکز پزشکی در کشور هلند با چهار کلاس به صورت کلاس‌های ادرار خونی، توده خوش خیم، توده بدخیم و سالم مورد ارزیابی و تحلیل قرار گرفت. نتایج بدست آمده در این رساله بر اساس معیارهای دقت، صحت، یادآوری (حساسیت)، خاص بودن و F1 به ترتیب برابر با (± 0.00234) ، ۷۲.۵، ۷۰.۵، ۷۱، ۸۹.۵ و ۷۲.۵ درصد بوده است که نشان دهنده بهبود عملکرد روش پیشنهادی نسبت به روش‌های موجود در این کاربرد می‌باشد.

کلمات کلیدی: یادگیری انتقال، طبقه‌بندی تصاویر سیستم‌سکوپی، ویژگی سطح بالا معنایی، طبقه‌بندی ترکیبی، یادگیری عمیق.

لیست مقالات مستخرج از رساله

مجلات

[مقاله ۱].

سید محمدرضا هاشمی، حمید حسن پور، احسان اله کوزه گر، تائو تن. "رده بندی تصاویر سیستم اسکوپ بر اساس یک شبکه عصبی تطبیق پذیر" مجله علمی پژوهشی پدافند الکترونیکی و سایبری، ویژه نامه ترکیبات، رمزنگاری و محاسبات، صفحه ۲۳۱-۲۳۸، سال انتشار: ۱۳۹۸.

[مقاله ۲].

S. M. R. Hashemi, H. Hassanpour, E. Kozegar, and T. Tan, "Cystoscopy Image Classification Using Deep Convolutional Neural Networks," International Journal of Nonlinear Analysis and Applications, vol. 10, pp. 193-215, 2019.

[مقاله ۳].

S. M. R. Hashemi, H. Hassanpour, E. Kozegar, and T. Tan, "Cystoscopic Image Classification Based on Combining MLP and GA," International Journal of Nonlinear Analysis and Applications, vol. 11, pp. 93-105, 2020.

[مقاله ۴].

S. M. R. Hashemi, H. Hassanpour, E. Kozegar and T. Tan, "Cystoscopic image classification by unsupervised feature learning and fusion of classifiers," in IEEE Access, doi: 10.1109/ACCESS.2021.3098510.

[مقاله ۵].

S. M. R. Hashemi, H. Hassanpour, E. Kozegar and T. Tan, "A new unsupervised feature learning for cystoscopic image classification", Medical Physics, (Under Review).

فهرست مطالب

۱	۱ مقدمه‌ای بر تصاویر سیستم‌توسکوپی مثانه
۲	۱ - ۱ مقدمه
۳	۱ - ۲ هدف و فرضیات رساله
۵	۱ - ۳ چالش‌های رساله
۷	۱ - ۴ دستاوردهای رساله
۱۰	۱ - ۵ سازماندهی طرح تحقیق
۱۱	۱ - ۶ جمع‌بندی
۱۳	۲ پیشینه تحقیق
۱۴	۲ - ۱ مقدمه
۱۵	۲ - ۲ مروری بر تکنیک‌های یادگیری عمیق
۱۶	۲ - ۲ - ۱ شبکه‌های عصبی
۱۷	۲ - ۲ - ۲ خودرمزنگارها و خودرمزنگارهای انباشته
۱۷	۲ - ۲ - ۳ ماشین‌های بولتزمن محدود (RBM) و شبکه‌های باور عمیق
۱۸	۲ - ۲ - ۴ شبکه پیچشی
۱۹	۲ - ۲ - ۵ ساختارهای چند جریانی
۲۰	۲ - ۲ - ۶ بخش‌بندی تصاویر پزشکی
۲۱	۲ - ۲ - ۷ طبقه‌بندی تصویر
۲۲	۲ - ۲ - ۸ طبقه‌بندی شیء یا ضایعه
۲۳	۲ - ۳ طبقه‌بندی تصاویر پزشکی بر پایه ویژگی‌ها
۲۷	۲ - ۴ جمع‌بندی
۲۹	۳ روش پیشنهادی در طبقه‌بندی تصاویر سیستم‌توسکوپی مثانه

۳۰	۳ - ۱ مقدمه
۳۱	۳ - ۲ روش پیشنهادی برای ویژگی‌های سطح پایین
۳۲	۳ - ۲ - ۱ استخراج ویژگی LBP و کاهش ابعاد
۳۴	۳ - ۲ - ۲ شبکه عصبی چندلایه پرسپترون
۳۵	۳ - ۲ - ۳ نرخ یادگیری تطبیقی
۳۶	۳ - ۲ - ۴ وزن‌دهی اولیه به کمک الگوریتم ژنتیک
۳۸	۳ - ۳ روش پیشنهادی بر اساس ویژگی سطح بالا بر پایه انتقال یادگیری
۴۳	۳ - ۳ - ۱ یادگیری انتقالی
۴۸	۳ - ۳ - ۲ نمایش ویژگی‌ها
۴۹	۳ - ۳ - ۳ کاهش ویژگی
۵۰	۳ - ۳ - ۴ طبقه‌بندی تصویر
۵۱	۴ - ۳ روش پیشنهادی بر اساس بخش‌بندی نواحی نگاشت ویژگی CNN
۵۴	۴ - ۳ - ۱ بخش‌بندی نگاشت ویژگی
۵۷	۴ - ۳ - ۲ مدل تقسیم تصویر چندسطحی با استفاده از روش اتسو بهینه شده
۵۸	۴ - ۳ - ۳ کاهش ویژگی بر اساس نواحی بخش‌بندی شده
۵۹	۳ - ۵ جمع‌بندی
۶۱	۴ پیاده‌سازی و تحلیل نتایج آزمایش‌ها
۶۲	۴ - ۱ مقدمه
۶۲	۴ - ۲ تنظیمات کلی پیاده‌سازی
۶۲	۴ - ۲ - ۱ پایگاه داده
۶۴	۴ - ۲ - ۲ معیارهای ارزیابی
۶۵	۴ - ۳ آزمایش‌ها و نتایج روش‌های پیشنهادی بر پایه ویژگی‌های سطح پایین
۶۹	۴ - ۳ - ۱ مقایسه روش پیشنهادی بر پایه ویژگی سطح پایین با روش‌های دیگر
۷۲	۴ - ۴ آزمایش‌ها و نتایج روش‌های پیشنهادی بر پایه ویژگی‌های سطح بالا و یادگیری انتقال
۷۲	۴ - ۴ - ۱ تنظیمات مربوط به پارامترهای استفاده در روش پیشنهادی بر اساس ویژگی سطح بالا
۷۳	۴ - ۴ - ۲ انتخاب معماری
۷۵	۴ - ۴ - ۳ مقایسه روش پیشنهادی بر پایه ویژگی‌های استخراج شده از یادگیری انتقال

۷۸	۴ - ۵ آزمایش‌ها و نتایج روش‌های پیشنهادی بر پایه ویژگی‌های سطح بالا و بخش‌بندی نواحی نگاشت ویژگی
۸۰	۴ - ۵ - ۱ نتایج بدست آمده از روش بخش‌بندی نواحی نگاشت ویژگی
۸۲	۴ - ۵ - ۲ تحلیل پیچیدگی زمانی
۸۳	۴ - ۶ مقایسه روش‌های مختلف و ارزیابی نتایج
۸۷	۴ - ۷ جمع‌بندی
۸۹	۵ جمع‌بندی و کارهای آتی
۹۰	۵ - ۱ مقدمه
۹۰	۵ - ۲ جمع‌بندی پژوهش انجام شده در رساله
۹۵	۵ - ۳ پیشنهادهایی برای کارهای آینده
۹۶	۶ منابع

فهرست تصاویر

شکل ۱-۱ نمونتهایی از تصاویر سیستم اسکوپ میثانه، الف) نمونه بدخیم، ب) خوش خیم، ج) خونی و د) سالم	۵
شکل ۲-۱ نمونتهایی از تصاویر موجود در پایگاه داده	۶
شکل ۱-۳ : مراحل استخراج ویژگی در روش پیشنهادی توسط روش LBP	۳۳
شکل ۲-۳ کاهش ویژگی در روش پیشنهادی بر پایه ویژگی سطح پایین	۳۴
شکل ۳-۳ : روش پیشنهادی برای وزن دهی اولیه و ضریب یادگیری تطبیقی	۳۷
شکل ۴-۳ روش پیشنهادی بر پایه یادگیری انتقال و ترکیب طبقه بندی تصاویر سیستم اسکوپ میثانه	۴۲
شکل ۵-۳ نمایش تصویر خروجی لایه های مختلف شبکه GoogLeNet با استفاده از یادگیری انتقالی و با تصویر ورودی OCT از شبکه چشم (نوع AMD)	۴۶
شکل ۶-۳ نمایش تصویر خروجی لایه های مختلف شبکه GoogLeNet با استفاده از یادگیری انتقالی و با تصویر ورودی OCT از شبکه چشم (نوع DME)	۴۷
شکل ۷-۳ استفاده از یادگیری انتقالی در روش پیشنهادی	۴۸
شکل ۸-۳ نمایشی از خروجی لایه های مختلف CNN در تصاویر سیستم اسکوپ میثانه. به ترتیب خروجی لایه کانولوشن اول، سوم و پنجم در هر سطر نمایش داده شده است.	۴۹
شکل ۹-۳ روش پیشنهادی برای طبقه بندی تصاویر سیستم اسکوپ میثانه بر پایه بخش بندی نواحی نگاشته های ویژگی CNN	۵۲
شکل ۱۰-۳ بخش بندی و طبقه بندی مبتنی بر انتخاب ویژگی توسط روش IPSO و اتسو	۵۳
شکل ۱-۴ نمونتهایی از تصاویر پایگاه داده. (a) نمونه ای از تصویر توده بدخیم میثانه، (b) نمونه ای از تصویر توده خوش خیم میثانه، (c) نمونه ای از تصویر خونی میثانه (d) نمونه ای از تصاویر سالم	۶۳
شکل ۲-۴ مراحل انتخاب نمونه در مجموعه آموزش و آزمون توسط روش ShuffleSplit	۶۴
شکل ۳-۴ مقادیر ویژه استخراج بعد از اعمال PCA در تصاویر سیستم اسکوپ میثانه	۶۶
شکل ۴-۴ میزان دقت شبکه با تعداد نورونهای متفاوت در مجموعه اعتبار سنجی	۶۸
شکل ۵-۴ میزان خطای شبکه MLP با نرخ یادگیری ثابت	۶۹
شکل ۶-۴ میزان خطای شبکه MLP با نرخ یادگیری متغیر	۷۰
شکل ۷-۴ میزان خطای شبکه MLP با نرخ یادگیری متغیر و وزندهی اولیه توسط الگوریتم ژنتیک	۷۱

- شکل ۸-۴ نتایج اولیه بر اساس معماری‌های مختلف در طبقه‌بندی تصاویر سیستم اسکوپ میثانه ۷۴
- شکل ۹-۴ یک نمونه خروجی از لایه‌های پیش‌بینی در تصاویر سیستم اسکوپ میثانه برای هر کلاس نرمال،
توده خوش خیم ، خونی و توده بدخیم ۷۹
- شکل ۱۰-۴ مقایسه طبقه‌بندی با معماری‌های مختلف ۸۰
- شکل ۱۱-۴ نمونه ای از نواحی انتخاب شده بر اساس نگاشت ویژگی ۸۱
- شکل ۱۲-۴ مقایسه زمان اجرای الگوریتم پیشنهادی برای مجموعه آموزش و آزمون، در دو حالت
با/بدون Parallel ۸۲
- شکل ۱۳-۴ زمان اجرای الگوریتم در مجموعه آزمون، با/بدون PCA ۸۳

فهرست جداول

جدول ۴-۱	پارامترهای شبکه MLP برای تشخیص بیماری در تصاویر سیستم اسکوپ میانه	۶۷
جدول ۴-۲	پارامترهای پیکربندی الگوریتم ژنتیک برای محاسبه وزن‌های اولیه	۶۸
جدول ۴-۳	میزان تغییر نرخ یادگیری با توجه به خطای اعتبارسنجی	۶۹
جدول ۴-۴	مقایسه روش‌ها برای طبقه‌بندی تصاویر سیستم اسکوپ میانه (میزان دقت)	۷۱
جدول ۴-۵	پارامترها و تعداد ویژگی‌های استخراج شده قبل و بعد از روش کاهش ویژگی	۷۲
جدول ۴-۶	پارامترهای طبقه‌بندی پایه	۷۳
جدول ۴-۷	مقایسه روش‌های پایه در طبقه‌بندی تصاویر سیستم اسکوپ	۷۵
جدول ۴-۸	مقایسه روش پیشنهادی با روش‌های ترکیبی	۷۶
جدول ۴-۹	مقایسه روش پیشنهادی بر اساس چهار معیار ارزیابی صحت (PR)، یادآوری (حساسیت)	
	(RE)، خاص بودن (Spe) و F1 در هر کلاس به صورت جداگانه	۷۶
جدول ۴-۱۰	مقایسه روش‌های مختلف بر اساس معماری ResNet برای استخراج ویژگی	۷۷
جدول ۴-۱۱	بررسی تأثیر کاهش ویژگی در روش پیشنهادی برای طبقه‌بندی تصاویر سیستم اسکوپ	
		۷۸
جدول ۴-۱۲	مقایسه روش‌های طبقه‌بندی بر اساس معیارهای صحت (PR)، یادآوری (حساسیت)	
	(RE)، خاص بودن (Spe) و F1 در هر کلاس به صورت جداگانه با روش پیشنهادی بر پایه بخش‌بندی	
		۸۱
جدول ۴-۱۳	مقایسه روش‌های موجود در طبقه‌بندی تصاویر سیستم اسکوپ میانه	۸۴

واژه‌های اختصاری

علامت اختصاری	واژه اصلی
CNN	Convolutional Neural Network
MLP	Multilayer Perceptron
DNN	Deep Neural Network
SAE	Stacked Autoencoder
DBN	Deep Belief Network
RBM	Restricted Boltzmann Machine
VOI	Volume of Interest
fcNN	fully convolutional neural network
BL	Blue Light
GPU	Graphical Processing Unit
CBIR	Content-Based Image Retrieval
SVM	Support Vector Machine
RNN	Recurrent Neural Networks
DWT	Discrete Wavelet Transform
MLP	Multi-Layer Perceptron
LBP	Local Binary Pattern
PCA	Principal Component Analysis
LR	Logistic Regression
GB	Gaussian Bayes
RF	Random Forest
IPSO	Improved Particle Swarm Optimization
SVC	Support Vector Classifier

۱ مقدمه‌ای بر تصاویر سیستم‌های کپی

مثانه

۱ - ۱ مقدمه

سرطان مثانه با بیش از ۳۳۰/۰۰۰ مورد جدید در سال و بیش از ۱۳۰/۰۰۰ مرگ در سال از نظر شیوع سرطان مرتبه نهم را در جهان دارا است. در ابتدای تشخیص سرطان مثانه در بیماران ۷۰٪ موارد بدون تهاجم به عضله و ۳۰٪ مهاجم به عضله می‌باشند [۱]. در بین بیمارانی که به دلیل مهاجم بودن سرطان به عضله مثانه، تحت عمل رادیکال سیستکتومی قرار گرفته‌اند ۵۷٪ آن‌ها از ابتدا تومور مهاجم به عضله بوده و در ۴۳٪ موارد ابتدا تومور غیر مهاجم به عضله بوده و بعداً پیشرفت کرده و تومور به عضله تهاجم کرده است [۱، ۲]. تقریباً یک‌سوم بیمارانی که تشخیص تومور مهاجم به عضله در آن‌ها گذاشته می‌شود در زمان درمان تومور اولیه متاستازهای تشخیص داده نشده دارند [۳]. ۲۵٪ بیمارانی که تحت رادیکال سیستکتومی قرار گیرند دارای درگیری غدد لنفی می‌باشند که بر اساس پاتولوژی مشخص می‌گردد [۴]. طیف بالینی سرطان مثانه را می‌توان به سه گروه تقسیم‌بندی کرد که در هر کدام پیش‌آگهی و درمان متفاوت است.

- گروه اول شامل تومورهای غیر مهاجم به عضله است که عمدتاً هدف درمان آن‌ها کاهش عود تومور و جلوگیری از پیشرفت تومور به مرحله بالاتر است.
- گروه دوم شامل تومورهای مهاجم به عضله است. نکته مهم در این گروه این است که آیا مثانه باید برداشته شود یا می‌توانیم بدون تأثیر بر بقا، مثانه را حفظ کنیم. نکته دیگر این است که آیا تومور اصلی را فقط درمان کنیم یا اینکه بیمار در خطر بروز متاستاز است و باید درمان سیستمیک نیز علاوه بر درمان موضعی، در نظر گرفته شود.
- گروه سوم بیماران متاستاتیک است. نکته اصلی در این گروه، بهتر کردن کیفیت و در درجه بعد کمیت زندگی است.

این مسئله در حوزه طبقه‌بندی تصاویر سیستوسکوپی مثانه مطرح شده است. تشخیص توده خوش‌خیم و بدخیم از روی تصاویر سیستوسکوپی حتی توسط پزشک با احتمال پایی (۱۰ تا ۲۰

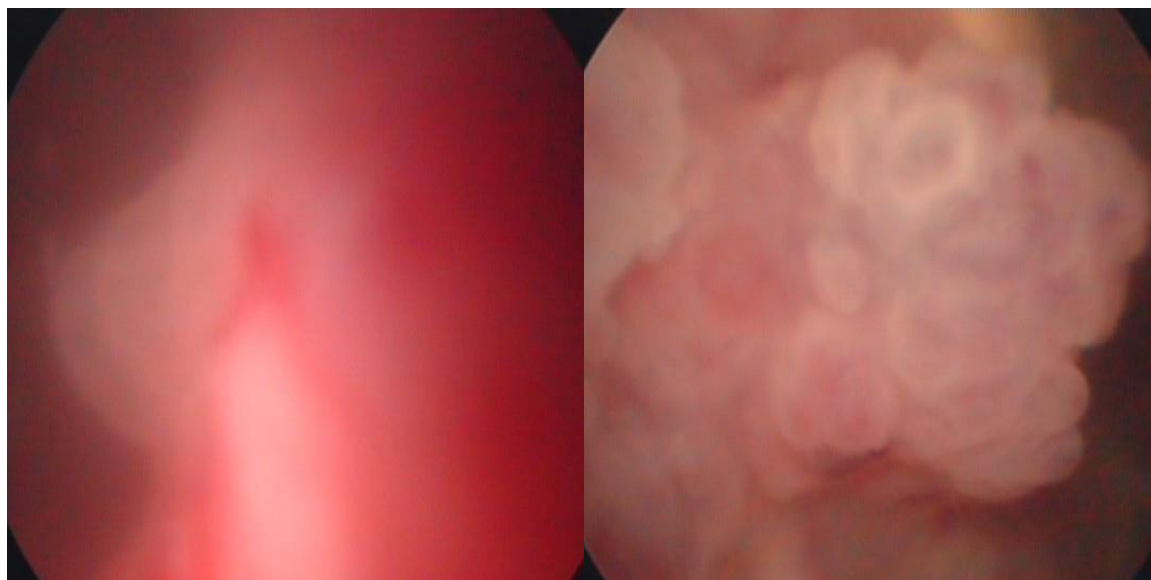
درصد) ممکن است. متاسفانه برای این کاربرد پایگاه داده قابل‌دسترس و استاندارد ی وجود ندارد و باید در گام نخست یک پایگاه داده استاندارد از مراکز درمانی تهیه نمود. در گام بعدی می‌توان از سیستم دسته‌بندی این نوع از تصاویر به‌عنوان مفسر دوم جهت بهبود تشخیص و کمک به پزشک در تصمیم‌گیری و تشخیص بهتر بیماری استفاده کرد.

۱ - ۲ هدف و فرضیات رساله

انگیزه اصلی این پژوهش کمک به متخصصان در بررسی تصاویر سیستم‌توسکوپي جهت تشخیص بیماری‌های مٲانه و طبقه‌بندی این تصاویر است. پزشک با استفاده از خروجی این سیستم می‌تواند نظر دقیق‌تری در مورد بیماری مٲانه هر بیمار بدهد. در واقع با سیستم پیشنهاد شده در این پژوهش می‌توان به اورلوژیست‌ها در تشخیص بیماری‌های مٲانه کمک کرد. هدف بعدی کمک به بهبود آموزش دانشجویان در تشخیص بیماری‌های مٲانه است. هدف اساتید، آشنایی دانشجویان با نشانه‌های بیماری و تحلیل تصاویر پزشکی گرفته شده از مٲانه است. در قدم بعدی پایین آمدن هزینه بیمار را شامل می‌شود و پزشک می‌تواند با استفاده از نرم‌افزار پیشرفت بیماری یا بهبود آن را بررسی نماید. جلوگیری از پیشرفت بیماری با تشخیص سریع‌تر، قبل از انجام پاتولوژی باعث بهبود سریع‌تر این بیماری خواهد شد.

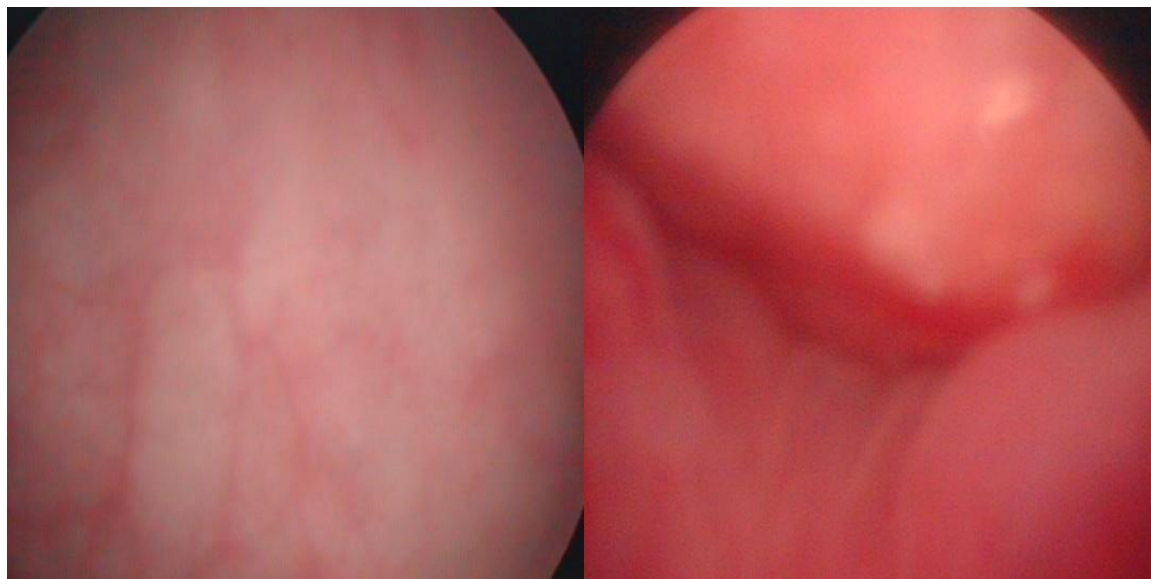
در طول سه دهه اخیر استفاده از روش‌های هوشمند در سامانه تشخیص پزشکی مورد توجه بسیاری از محققان در حوزه پردازش تصویر و یادگیری ماشین قرار گرفته است. ولی سیستم هوشمند موفق برای تشخیص سرطان مٲانه از طریق تصاویر سیستم‌توسکوپي در حوزه پردازش تصاویر پزشکی ارائه نشده است. یک دلیل اساسی برای این مشکل عدم وجود مجموعه داده استاندارد با تعداد مناسب بوده است. برای حل این مشکل باید در گام نخست یک مجموعه داده استاندارد برای این نوع از تصاویر جمع آوری نمود. این تصاویر باید توسط متخصصان و پزشکان این حوزه برچسب‌گذاری شوند. معمولاً در این کاربرد تصاویر به چهار دسته کلی، ادرار خونی، توده خوش خیم، توده بدخیم و سالم طبقه‌بندی

می‌شوند که به یک مسئله طبقه‌بندی تبدیل می‌شود. نمونه‌ها از تصاویر در چهار کلاس متفاوت در شکل ۱-۱ نشان داده شده است. اگر بخواهیم مسئله طبقه‌بندی تصویر سیستم‌توسکوپي را به‌طور دقیق‌تر بیان نمائیم، فرض کنید $X = \{x_1, x_2, \dots, x_N\}$ مجموعه‌ی تمام تصاویر آموزشی و تعداد این تصاویر $|X| = N$ باشد. مجموعه $Y = \{y_1, y_2, \dots, y_N\}$ نیز با همین تعداد اعضاء به عنوان برچسب‌های تصاویر در اختیار است. اگر کل داده‌ها به C دسته مختلف تعلق داشته باشند، آنگاه مقادیر $y_i \in \{1, 2, \dots, C\}$ است. سیستم باید بتواند با بیشترین دقت برچسب داده‌های آزمون را مشخص کند. در این رساله فرض می‌شود که تصاویر استفاده شده از زوایای مختلف با ابعاد متفاوت تهیه شده است.



(ب)

(الف)



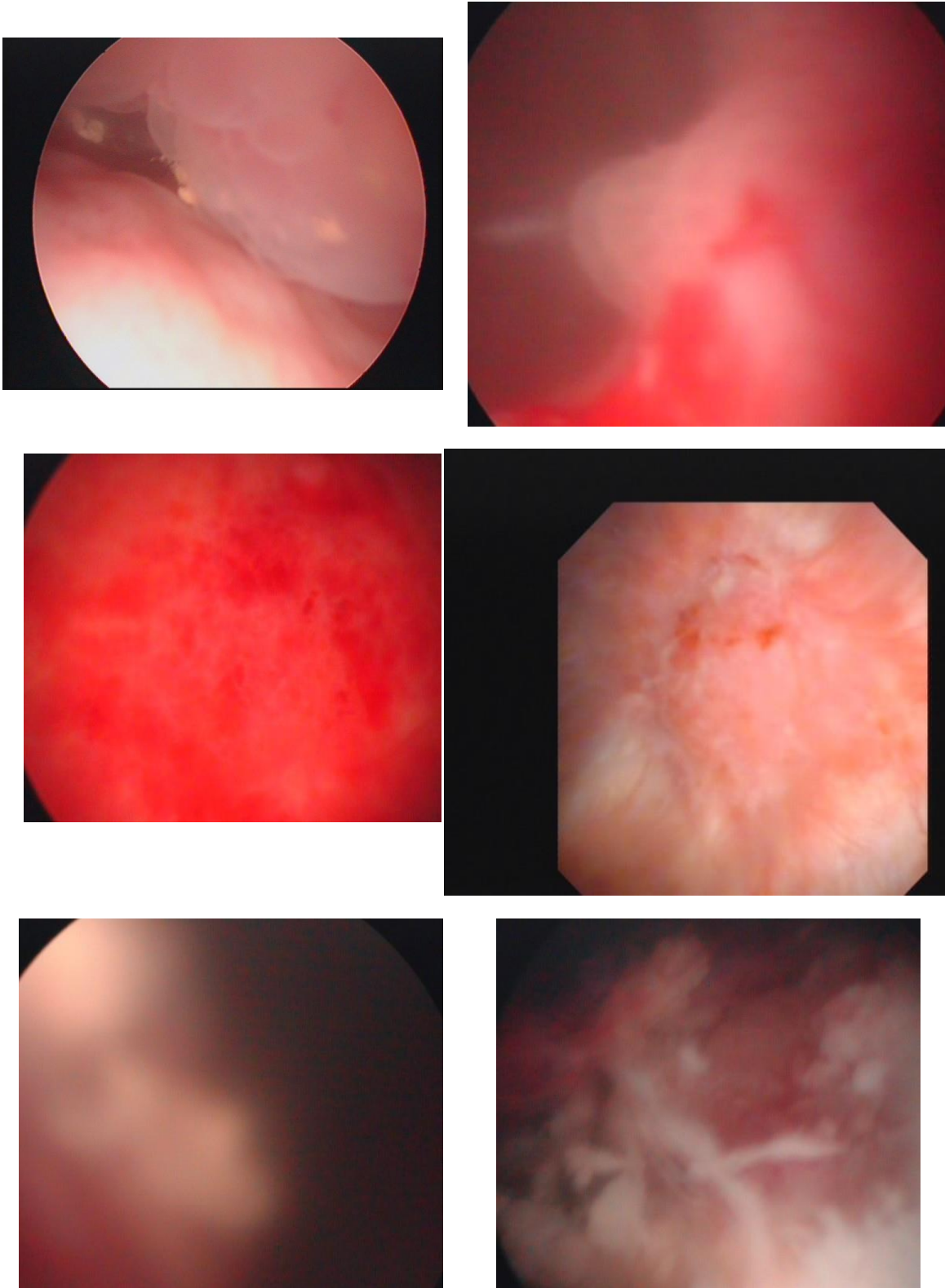
(د)

(ج)

شکل ۱-۱ نمونه‌هایی از تصاویر سیستوسکوپی مثانه، الف) نمونه بدخیم، ب) خوش خیم، ج) خونی و د) سالم

۱ - ۳ چالش‌های رساله

نمونه‌های از تصاویر سیستوسکوپی مثانه در شکل ۱-۲ نشان داده شده است. همانطور که از شکل مشخص می‌باشد کیفیت تصاویر سیستوسکوپی به خاطر شرایط متفاوت تصویربرداری و محیط مثانه یکسان نیست. تصاویر گرفته شده در زاویه‌های متفاوت با اندازه‌های مختلف است و تصاویر سیستوسکوپی از ناحیه‌های متفاوت از مثانه گرفته می‌شود زیرا سیستوسکوپی به معنی آندوسکوپی داخل مثانه است. در این روش با عبور یک دوربین باریک به نام سیستوسکوپ به داخل مجرای ادرار و مثانه، پزشک مجرای ادرار و مثانه را از نظر بیماری‌های احتمالی از قبیل علل خونریزی و تنگی مجرا، بزرگی پروستات، تومور مثانه بررسی می‌کند. با اتصال دوربین به مانیتور همزمان تصاویر مشاهده می‌شود و برای تحلیل دقیق‌تر و بیشتر ذخیره و بعداً بازبینی می‌شود. جمع‌آوری اطلاعات و ایجاد یک پایگاه داده استاندارد به دلیل پیچیدگی بالای این نوع از تصاویر و نظرات متفاوت متخصصین در تصاویر یکسان، با مشکلات زیادی روبروست.



شکل ۱-۲ نمونه‌هایی از تصاویر موجود در پایگاه داده

کیفیت پایین تصاویر و عدم فرمت یکسان تصاویر باعث پیچیدگی سیستم دسته‌بندی تصاویر سیستوسکوپی می‌گردد. دستیابی به داده‌های پزشکی دارای کیفیت بالا چالش‌برانگیز است. از طرفی

به دلیل محدودیت‌های پزشکی و محدودیت زمان و هزینه، جمع‌آوری حجم بالای داده‌ها عملاً امکان پذیر نمی‌باشد. به همین دلیل، این نیاز احساس می‌شود که از روش‌های قدرتمند برپایه‌های مدل‌های از پیش آموزش داده شده به منظور استفاده از دانش موجود در منابع دیگر استفاده کرد. با توجه به موارد بیان شده، می‌توان از چالش‌های اساسی در این رساله به موارد زیر اشاره نمود:

- ۱- نبود مجموعه داده استاندارد و کافی
- ۲- شرایط تصویر برداری متفاوت
- ۳- اندازه و زاویه متفاوت در تصاویر استخراج شده
- ۴- پیچیدگی بالای تصاویر که حتی متخصصان (پزشکان) نمی‌توانند تشخیص دقیق بدهند و نیاز به آزمایشات و هزینه بیشتر دارد.

۱ - ۴ دستاوردهای رساله

در طول سه دهه اخیر، استفاده از روش‌های هوشمند در سامانه‌های تشخیص پزشکی مورد توجه بسیاری از محققان قرار گرفته است [۵، ۶]. ولی برای تشخیص سرطان مthane از طریق تصاویر سیستم‌سکوپی با وجود شیوع بسیار زیاد در سطح جهان فعالیت هوشمندی در حوزه پردازش تصویر پزشکی ارائه نشده است. در روش پیشنهادی اول این رساله برپایه روش‌ها و ویژگی‌های سنتی، از شبکه عصبی چندلایه برای رده بندی تصاویر سیستم‌سکوپی مthane استفاده شده است. یکی از مهمترین مشکلات در الگوریتم‌های شبکه عصبی تعیین میزان نرخ یادگیری می‌باشد اگر نرخ یادگیری کوچک انتخاب شود باعث کاهش سرعت همگرایی و اگر خیلی بزرگ انتخاب شود موجب نوسان و واگرایی می‌شود لذا نیاز به الگوریتمی داریم که بتواند بصورت پویا نرخ یادگیری را تغییر دهد. در این راستا برای بهبود شبکه عصبی چند لایه روشی تطبیقی برای تعیین نرخ یادگیری ارائه شده است. در این روش نرخ یادگیری با استفاده از یک ضریبی بر اساس اختلاف دقت آموزش و اعتبارسنجی با توجه به خروجی خطا مشخص می‌شود. نرخ تغییرات با توجه به میزان تغییرات وزن و خطای خروجی بروز

رسانی می‌شود. روش پیشنهادی بر روی پایگاه داده جمع آوری شده از تصاویر سیستم‌توسکوپي مٲانه مورد ارزیابی قرار گرفته است. نتایج شبیه‌سازی نشان دهنده کاهش خطا و افزایش سرعت همگرایی روش پیشنهادی نسبت به روش‌های دیگر در رده بندی تصاویر سیستم‌توسکوپي می‌باشد. لازم به ذکر است از انجایی تحقیق دیگر بر روی این پایگاه داده انجام نشده بود روش‌های پایه توسط نویسندگان پیاده‌سازی و ارزیابی شد.

در ادامه اشاره شد که، یک روش مرسوم برای طبقه‌بندی تصاویر پزشکی استفاده از ویژگی‌های سطح پایین می‌باشد. در همین راستا، در گام نخست روشی بر پایه الگوهای باینری محلی^۱ LBP با شبکه‌های عصبی پرسپترون ارائه شده است. مراحل اصلی روش پیشنهادی شامل استخراج و کاهش ویژگی، بهبود شبکه عصبی با نرخ یادگیری تطبیقی و به دست آوردن وزن‌های اولیه توسط الگوریتم ژنتیک است. در این روش، یک رویکرد نوین از MLP برای پیش‌بینی و طبقه‌بندی تصاویر سیستم‌توسکوپي با دقت و صحت بالا ارائه شده است.

انتخاب وزن‌های اولیه در بهبود عملکرد شبکه MLP بسیار تاثیرگذار هست. معمولاً از وزن‌هایی به صورت تصادفی برای وزن‌های اولیه استفاده می‌شود. استفاده از وزن‌های اولیه تصادفی در مجموعه داده‌های کوچک می‌تواند اثرات مخربی داشته باشد و آموزش شبکه به خوبی انجام نمی‌شود. معمولاً تصاویر پزشکی شامل مجموعه داده‌های بزرگ نمی‌باشند و تعداد نمونه‌های استخراجی از بیماران کم است. در همین راستا، نباید از وزن‌های اولیه تصادفی برای آموزش شبکه استفاده نمود. در واقع در مجموعه داده‌های کوچک باید وزن‌های اولیه با دقت مناسب‌تری انتخاب شوند. به همین دلیل، یک روش نوین بر پایه الگوریتم ژنتیک برای انتخاب وزن‌های اولیه شبکه MLP ارائه شده است.

نتایج بدست آمده از روش‌های قبل عملکرد بهتری نسبت به روش‌های دیگر بر پایه ویژگی‌های سنتی یا سطح پایین از خود نشان داده است. ولی همچنان این نتایج برای کاربردهای واقعی مناسب

¹ Local Binary Patterns (LBP)

نمی‌باشند در همین راستا، یک روش نوین برپایه یادگیری عمیق پیچشی برای طبقه‌بندی تصاویر سیستم‌سکوپی ارائه شده است. نتایج نشان دهنده بهبود عملکرد روش پیشنهادی نسبت به روش‌های قبل می‌باشد.

در ادامه، یادگیری انتقال بر اساس شبکه CNN برای استخراج ویژگی استفاده شد. روش پیشنهادی به گونه‌ای طراحی شده است که نسبت به تغییرات شرایط تصویربرداری مستقل باشد. این روش با استخراج یک ویژگی سطح بالا از تصاویر سیستم‌سکوپی آغاز می‌شود. از آنجا که اثبات شده است که یادگیری عمیق در استخراج ویژگی با سطوح معنایی مختلف موثر است، این نوع شبکه برای استخراج ویژگی استفاده شده است. به منظور حل مشکل تعداد محدود تصاویر سیستم‌سکوپی، از انتقال یادگیری استفاده می‌شود تا با استفاده از مجموعه آموزش کوچک موجود، یک مدل یادگیری مناسب ساخته شود. در رویکرد یادگیری انتقال، آخرین لایه‌های یک شبکه از پیش آموزش دیده، که قبلاً روی یک مجموعه داده بزرگ مانند ImageNet آموزش دیده‌اند، استفاده شده است. برای جلوگیری از مشکل ابعاد بالا و همچنین کاهش زمان و پیچیدگی فضا، از تعداد ویژگی‌ها با استفاده از روش‌های کاهش بعد کاسته می‌شود. سرانجام، یک طبقه‌بند ترکیبی برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه پیشنهاد می‌شود. در روش پیشنهادی برای ترکیب از طبقه‌بندهای شناخته شده با استفاده از روش آرای اکثریت وزنی استفاده شده است. روش پیشنهادی نتایج امیدوار کننده مربوط به طبقه‌بندی این نوع از تصاویر سیستم‌سکوپی را نشان می‌دهد.

در ادامه بهبود روش ارائه شده، از یادگیری انتقال برای استخراج نواحی مربوط به نگاشت‌های ویژگی^۱ استفاده شده است. آخرین لایه کانولوشن برای انتخاب نگاشت ویژگی استفاده می‌شود. در مرحله بعدی، ویژگی‌های سطح بالا و نواحی مربوط به این ویژگی‌ها بر اساس روش‌های بهینه سازی ازدحام ذرات بهبود یافته (IPSO) و اتسو استخراج می‌شوند. لازم به ذکر است که مقادیر نگاشت ویژگی در

¹ Feature Map

محدوده [۰،۱] نرمال شده‌اند. سپس، با استفاده از روش تجزیه و تحلیل مولفه‌های اصلی^۱ PCA، ویژگی استخراج شده در بخش‌ها کاهش می‌یابد. در آخر، طبقه‌بندی انجام می‌شود. در این روش، از یک روش ترکیبی مبتنی بر رای‌گیری اکثریت با BG و SVM استفاده شده است. نتایج بدست آمده نشان دهنده بهترین عملکرد روش‌های موجود در طبقه‌بندی تصاویر سیستم‌سکوپی می‌باشد.

۱ - ۵ سازماندهی طرح تحقیق

همانطور که در بخش‌های مختلف این فصل بیان شد، طراحی یک سیستم یکپارچه به منظور طبقه‌بندی تصاویر سیستم‌سکوپی این رساله در پنج فصل تهیه شده است. هر فصل به شرح زیر است.

۱. فصل دوم بر کارهای انجام شده در زمینه طبقه‌بندی تصویر پزشکی تمرکز دارد. در ابتدا

روش‌های ارائه شده بر پایه یادگیر عمیق معرفی شده است. در انتها یک مقایسه بین روش-

های موجود در طبقه‌بندی تصاویر ارائه شده است.

۲. در فصل سوم روش پیشنهادی در طبقه‌بندی تصاویر سیستم‌سکوپی معرفی شده است. در

ابتدا روش پیشنهادی بر پایه ویژگی‌های سطح پایین معرفی شده است. در ادامه فصل تمرکز بر

طبقه‌بندی تصاویر بر اساس ویژگی‌های سطح بالا می‌باشد.

۳. در فصل چهارم پیاده‌سازی و آزمایشات انجام شده بر اساس روش‌های پیشنهادی در این

رساله ارائه شده است.

۴. در فصل پنجم نتیجه‌گیری و کارهای آتی در موضوع طبقه‌بندی تصاویر سیستم‌سکوپی بیان

شده است.

¹ Principal Component Analysis (PCA)

۱ - ۶ جمع‌بندی

در دهه‌های اخیر، استفاده از روش‌های هوشمند در سامانه‌های تشخیص پزشکی مورد توجه بسیاری از محققان قرار گرفته است. ولی برای تشخیص سرطان مٲانه از طریق تصاویر سیستم‌توسکوپي با وجود شیوع بسیار زیاد در سطح جهان فعالیت هوشمندی در حوزه پردازش تصویر پزشکی ارائه نشده است. در همین راستا، در این فصل به معرفی فرضیات و اهداف موجود در این کاربرد و رساله پرداخته شده است. در گام بعدی با توجه پیچیدگی‌های موجود در این کاربرد چالش‌های موجود معرفی شد. در انتها راهکارهای ارائه شده در این رساله و دستاوردهای آن به صورت کلی بیان شد.

۲ پيشينه تحقيق

۲ - ۱ مقدمه

در این فصل با معرفی پژوهش‌های پیشین انجام شده در زمینه طبقه‌بندی تصاویر پزشکی، یافته‌های دیگر محققان بررسی می‌گردد. تمرکز اصلی در این فصل استفاده از یادگیری عمیق در طبقه‌بندی تصاویر پزشکی می‌باشد. در همین راستا، برای درک بهتر این موضوع ابتدا روش‌های موجود و چگونگی استفاده از آنها در این حوزه معرفی خواهد شد.

زمانی که امکان اسکن و بارگذاری تصاویر پزشکی به رایانه ایجاد شد، طراحی سیستم‌هایی به‌منظور آنالیز خودکار تصاویر پزشکی مورد توجه محققان این حوزه قرار گرفت. در ابتدا از دهه ۱۹۷۰ تا ۱۹۹۰ تجزیه و تحلیل تصویر پزشکی با استفاده از ویژگی‌های سطح پایین به مانند آشکارسازی لبه‌ها و خط‌ها [۷] و مدل‌سازی ریاضی [۸] به مانند خطوط، دایره‌ها و بیضی‌ها برپایه دستورات شرطی ساده برای یک منظور خاص طراحی می‌شده‌اند. این سیستم‌های خبره به عنوان هوش مصنوعی سنتی توصیف شده‌اند که اغلب به دلیل پیچیدگی بالای تصاویر و قدرت پایین روش‌های استفاده شده، عملکرد مناسبی از خود نشان نمی‌داده‌اند.

در پایان دهه ۱۹۹۰ تکنیک‌های با نظارت که در آن داده‌های آموزشی برای توسعه یک سیستم مورد استفاده قرار می‌گیرند، در آنالیز تصویر پزشکی به طور فزاینده‌ای مورد توجه قرار گرفتند [۵، ۹]. از مهمترین روش‌ها می‌توان به مدل‌های شکل فعال (برای بخش‌بندی)، روش atlas (که در آن اطلس‌هایی که متناسب با داده‌های جدید هستند، داده‌های آموزشی را شکل می‌دهند) و مفهوم استخراج ویژگی‌ها و استفاده از طبقه‌بندی‌های آماری اشاره نمود [۱۰]. این روش‌ها که به عنوان روش‌های مبتنی بر شناسایی الگو می‌باشند، هنوز بسیار محبوب هستند و پایه بسیاری از سیستم‌های تحلیل تصویر پزشکی موفق تجاری را شکل می‌دهد. در این نوع از سیستم‌ها، تمامی مراحل به‌طور کامل توسط کاربر خبره طراحی می‌شود و با استفاده از داده‌های نمونه‌برداری شده، ویژگی‌ها توسط رایانه استخراج و آموزش داده می‌شود.

الگوریتم‌های یادگیری ماشین، مرز تصمیم‌گیری بهینه (طبقه‌بندی) را در فضای ویژگی با ابعاد بالا تعیین می‌کنند. گامی مهم در طراحی چنین الگوریتم‌هایی، استخراج ویژگی‌های متمایز از تصاویر است. این فرآیند تصمیم‌گیری برای انتخاب روش‌های مناسب برای استخراج ویژگی‌های متناسب با تصاویر توسط محقق صورت می‌گیرد. گام بعدی آموزش و یادگیری الگوریتم‌ها بر اساس ویژگی‌های استخراج شده می‌باشد. این مفهوم اساس بسیاری از الگوریتم‌های یادگیری عمیق است. شبکه‌های عمیق از لایه‌های زیادی تشکیل شده‌اند که در آنها، داده‌های ورودی (مثلاً تصویر) به خروجی‌ها (مثلاً بیمار بودن/ نبودن) نگاشت می‌شوند. موفق‌ترین نوع شبکه‌عصبی عمیق برای تجزیه و تحلیل تصویر تا به امروز شبکه‌های عصبی پیچشی^۱ (CNN) است [۱۱]. این نوع از شبکه‌ها شامل بسیاری از لایه‌ها هستند که ورودی‌هایشان را با فیلترهای پیچش در مقیاس کوچک تغییر می‌دهند. پیش از معرفی شبکه‌های CNN، تکنیک‌های مختلفی شامل تحلیل مؤلفه‌های اصلی [۱۲]، خوشه‌بندی تکه‌های تصویر [۱۳]، رویکردهای فرهنگ لغت [۱۴] و بسیاری دیگر برای یادگیری و طبقه‌بندی ویژگی‌ها رایج بودند [۱۵]. ولی در سال‌های اخیر به دلیل پیشرفت‌های سخت‌افزاری و افزایش داده‌های در دسترس و قابل پردازش، در بسیاری از تحقیقات انجام شده در حوزه آنالیز تصاویر پزشکی از تکنیک‌هایی برپایه یادگیری عمیق به عنوان یک رکن اساسی روش‌های پیشنهادی استفاده شده است [۱۱، ۱۶، ۱۷]. در ادامه‌ی این فصل به معرفی روش‌های موجود و مفاهیم پایه در طبقه‌بندی تصاویر پزشکی بخصوص تصاویر سیستوسکوپی مثانه می‌پردازیم.

۲ - ۲ مروری بر تکنیک‌های یادگیری عمیق

در این بخش، مفاهیم یادگیری عمیق را معرفی می‌کنیم که برای آنالیز تصویر پزشکی اعمال شده‌اند.

^۱ Convolutional Neural Network (CNN)

۲-۲-۱ شبکه‌های عصبی

معماری‌های عمیق بر پایه شبکه‌های عصبی هستند و می‌توانند به عنوان تعمیم یک رگرسیون خطی در نظر گرفته شوند [۱۸، ۱۹]. در چنین شبکه‌هایی، فعال‌ساز a از هر نورون نشان‌دهنده ترکیب خطی برخی از ورودی‌ها است. x ورودی از لایه‌های قبل، w و b به ترتیب وزن و بایاس می‌باشند که در ادامه توسط یک تابع سیگموئید ($\sigma(\cdot)$) مقدار خروجی مشابه رابطه ۲-۱ محاسبه می‌شود:

$$a = \sigma(w^T x + b) \quad ۲-۱$$

هنگامی که چندلایه تک جهته (یعنی شبکه‌ای پیشرو) در چنین دور قرار می‌گیرند، مدل به عنوان پرسپترون چندلایه‌ای^۱ (MLP) نامیده می‌شود که لایه‌های میانی معمولاً به عنوان لایه‌های پنهان شناخته می‌شوند که به‌طور مستقیم مشاهده نمی‌شوند (برخلاف لایه‌های ورودی و خروجی). هنگامی که یک شبکه دارای لایه‌های زیادی است، عموماً یک شبکه عصبی عمیق^۲ (DNN) نامیده می‌شود. برای مدتی طولانی، DNN‌ها به دلیل عدم وجود سخت افزار مناسب و تعداد پایین مجموعه داده‌ها، کاربرد چندانی نداشتند ولی در سال ۲۰۰۶ با پیشرفت سخت‌افزار و همچنین حجم مناسب داده، آموزش موثر در این نوع از شبکه‌ها اتفاق افتاد و محبوبیت زیادی را کسب کردند [۲۰-۲۲]. در DNN‌ها دو رویکرد کلی وجود دارد. در رویکرد اول، لایه‌ها توسط روش بدون نظارت آموزش می‌بینند که پس از تنظیم ساختار شبکه می‌توانند به عنوان ابزارهای شناسایی الگوی موثر استفاده شوند. دو ساختار محبوب خودرمنزنگار^۳ (SAE) و شبکه‌های باور عمیق^۴ (DBN) در این رویکرد هستند. در رویکرد دوم، آموزش کل یک شبکه عمیق انتها به انتها با نظارت انجام می‌شود. در ادامه این مفاهیم به اختصار توضیح داده می‌شود.

^۱ Multilayer Perceptron (MLP)

^۲ Deep Neural Network (DNN)

^۳ Stacked Autoencoder (SAE)

^۴ Deep Belief Network (DBN)

۲-۲-۲ خودرمزنگارها و خودرمزنگارهای انباشته

AEها شبکه‌های ساده‌ای هستند که برای بازسازی ورودی x در لایه خروجی x' از طریق یک لایه پنهان h آموزش دیده‌اند. آن‌ها توسط ماتریس‌های وزن $w_{x,h}$ و $w_{h,x'}$ و بایاس‌های $b_{x,h}$ و $b_{h,x'}$ کنترل می‌شوند. اگر لایه پنهان هم اندازه ورودی داشته باشد و هیچ تابع فعال‌ساز غیرخطی اضافه نشود، مدل به سادگی تابع را یاد می‌گیرد. با این حال استفاده از تابع فعال‌ساز غیرخطی برای محاسبه روابط پنهانی بین ورودی‌ها و خروجی‌ها است. رابطه ۲-۲ خروجی این نوع توابع را نشان می‌دهند.

$$h = \sigma(w_{x,h}x + b_{x,h}) \quad 2-2$$

در این رابطه، معمولاً $|h|$ کوچک‌تر از $|x|$ گرفته می‌شود. به این ترتیب، داده‌ها بر روی یک زیر فضای پائینی قرار می‌گیرند که نشان‌دهنده یک ساختار پنهان غالب در ورودی هستند. قیود منظم‌سازی یا پراکندگی را می‌توان برای کشف بهتر ساختار مرتبط به کار برد. ماتریس وزن $w_{h,x}$ اغلب به $w_{x,h}$ می‌رسد. خود رمزنگار نویزگیری [۲۳]، یکی دیگر از راه‌حل‌های پیشنهادی برای جلوگیری از یادگیری یک راه‌حل جزئی است. در این حالت، مدل برای بازسازی ورودی از یک نسخه مخدوش شده (معمولاً نویز) آموزش داده می‌شود.

SAEها (یا AEهای تنک) با قرار دادن لایه‌های خود رمزنگار در بالای یکدیگر شکل می‌گیرند. به‌طور معمول، هر لایه خود رمزنگار به‌طور جداگانه به کمک روش‌های حریصانه آموزش داده می‌شود. بعد از آن شبکه کامل با استفاده از آموزش نظارت شده برای پیش‌بینی برخی از برچسب‌هایی که ورودی را مشخص می‌کنند، تنظیم می‌شود.

۲-۲-۳ ماشین‌های بولتزمن محدود^۱ (RBM) و شبکه‌های باور عمیق

RBMها [۲۴] یک نوع فیلد تصادفی مارکوف هستند که شامل یک لایه ورودی یا لایه قابل مشاهده $x_1, x_2, \dots, x_N$ و یک لایه پنهان $h = h_1, h_2, \dots, h_M$ هستند. ارتباط بین گره‌ها دو سویه است. با توجه

¹ Restricted Boltzmann Machine (RBM)

به یک بردار ورودی x می‌تواند یک نمایه ویژگی نهفته h را به دست آورد. به این ترتیب، RBM یک مدل است به این معنا که می‌توانیم از آن نمونه‌برداری کرده و نقاط داده‌ای جدیدی را ایجاد کنیم که از توزیعی که در آن آموزش دیده است، ایجاد شود. در مقایسه با سیستم‌های فیزیکی، یک تابع انرژی برای یک حالت خاص (x, h) از واحدهای ورودی و نهان به صورت رابطه ۲-۳ تعریف می‌شود.

$$E(x, h) = h^T w x - c^T x - b^T h \quad ۲-۳$$

که در آن c و b بایاس هستند. این دو شامل نگاشت‌های ویژگی با ابعاد متفاوت با ادغام در لایه‌های پیچشی هستند. [۲۵، ۲۶].

۲ - ۲ - ۴ شبکه پیچشی

در مقاله [۲۷] یک شبکه ۲۲ لایه به نام GoogleNet معرفی گردید. مدل آن‌ها از بلوک‌های به اصطلاح آغازین استفاده می‌کند. بلوک‌های آغازین را می‌توان به عنوان شبکه در یک شبکه تفسیر کرد که در آن ورودی به چندین زیر شبکه پیچشی مختلف تقسیم شده است که در انتهای بلوک متصل می‌شود. این ساختار بهترین نوع ساختار عمیقی در سال ۲۰۱۵ بوده که به عنوان برنده چالش ImageNet در آن سال معرفی شده است. در سال بعد شبکه ResNet بر پایه بلوک باقیمانده^۱ به صورت رابطه ۲-۴ تعریف می‌شود:

$$y = x + F(x, w_R) \quad ۲-۴$$

همانطور که می‌توان از این معادله برداشت کرد شبکه فقط باید باقیمانده $F(x, w_R)$ را یاد بگیرد. به این ترتیب، مدل در جهت یادگیری نمایش‌های ساده در هر لایه که به تابع هویت نزدیک هستند، نشان داده می‌شود. تابع $F(x, w_R)$ به عنوان بلوک باقیمانده شناخته می‌شود که انواع مختلف آن در مقالات پیشنهاد شده است. ساختارهای ResNet نیز با بلوک‌های آغازین ترکیب می‌شوند.

^۱ Residual block

۲ - ۲ - ۵ ساختارهای چند جریانی

ساختار CNN پیش فرض می‌تواند منبع اطلاعات چندگانه یا نمایه‌های ورودی را در قالب کانال‌هایی که به لایه ورودی وارد می‌شود، جای دهد. در واقع ورودی‌های متفاوت و چندگانه در ساختار CNN قابل ادغام می‌باشند. در اصل، کانال‌های مختلف می‌توانند در هر نقطه از خط لوله ادغام شوند. تحت شهودی که وظایف مختلف نیاز به روش‌های متنوعی برای همجوشی دارد، ساختار چند جریانی مورد بررسی قرار می‌گیرد. این مدل‌ها گاهی به عنوان ساختار مسیرهای دوگانه نامیده می‌شوند [۲۸]. آنالیز تصویر در مقیاس‌های چندگانه برای چندین دهه انجام شده است که این دانش برای تقسیم‌بندی چند مقیاسی استفاده می‌شود و همچنین محققان این مفهوم را در چندین کاربرد پزشکی با موفقیت استفاده کرده‌اند [۲۳، ۲۹، ۳۰]. یک چالش دیگر در حوزه پزشکی، فرمت‌های ورودی جدید مانند داده‌های سه بعدی است. در کاربردهای اولیه CNN سه عدد تصویر سه بعدی، پیچش‌ها با تقسیم مقادیر علاقه‌مندی^۱ به تکه‌هایی که به اندازه جریان‌های مختلف به یک شبکه تغذیه می‌شوند، پیش‌دستی کردند.

یکی از تحقیقات پیش‌گام از این استراتژی برای تقسیم‌بندی غضروف زانو استفاده کردند [۳۱]. در این مقاله، یک سیستم جدید برای طبقه‌بندی و کسل با ادغام سه CNN دو بعدی پیشنهاد شد که به ترتیب با سطوح xy ، yz و zx تصویر سه بعدی ارتباط دارند. ادعا شده که، نتایج بدست آمده عملکرد بهتری نسبت به روش‌های دیگر از خود نشان داده است.

راهبردی دیگر تغذیه این شبکه با تکه‌های چند زاویه‌ای از فضای سه بعدی در یک مد چند جریانی است که توسط محققان مختلف در زمینه تصویربرداری پزشکی اعمال شده است [۳۲]. در این مقاله از یک سیستم تشخیص جدید برای ندول‌های ریوی با استفاده از چند شبکه پیچشی (ConvNets) پیشنهاد شد، که ویژگی‌های متمایزکننده^۲ به صورت طور خودکار از داده‌های آموزش استخراج شد. این شبکه از ترکیب سه آشکارساز کاندید به طور خاص برای ندول‌های جامد، نیمه جامد و جامد

^۱ Volume of Interest (VOI)

^۲ discriminative

بزرگ طراحی شده است. برای هر کاندید، مجموعه‌ای از تکه‌های دو بعدی از سطوح مختلف استخراج شد. معماری پیشنهادی شامل جریان‌های متعددی از ConvNets دو بعدی است که در آنها خروجی‌ها توسط روش تلفیقی اختصاصی ترکیب شد تا طبقه بندی نهایی به دست آید.

۲ - ۲ - ۶ بخش‌بندی تصاویر پزشکی

بخش‌بندی، کاری معمول در آنالیز تصویر پزشکی است. اگرچه CNNها در فرم اصلی خود، ساختارهای طبقه‌بندی هستند، اما می‌توانند به راحتی در هر پیکسل یا وکسل^۱ در یک تصویر استفاده شوند. روش ساده "پنجره لغزان" این مشکل را دارد که تکه‌های ورودی با پیکسل همسایه همپوشانی زیادی دارد و پیچش‌های بسیار برای همان نقطه محاسبه می‌شوند.

خوشبختانه، پیچش و ضرب نقطه‌ای هر دو عملگرهای خطی هستند و آن‌ها می‌توانند به جای یکدیگر نمایش داده شوند که به کاربرد شبکه‌ها در تصاویر بزرگ‌تر از آن‌هایی که با بازنویسی لایه‌های کاملاً متصل به هم آموزش دیده بودند، آموزش می‌دهد. نتیجه شبکه پیچشی کامل می‌تواند به یک مجموعه از مجموعه‌های جمع‌آوری شده با یک روش کارآمد برای یک تصویر کامل به کاربرد. با این حال، با استفاده از لایه‌های "معتبر" یا لایه‌های تجمع، این امر ممکن است منجر به خروجی با وضوح بسیار کم‌تر از ورودی شود. چندین راه حل برای گریز از این تنزل وضوح پیشنهاد شده است. ساده‌ترین آن‌ها "تغییر و جوش" نامیده می‌شوند [۳۳، ۳۴] که در آن $fCNN^2$ دقیقاً همان مقدار از فاکتور نمونه‌برداری در هر جهت با تغییر یک پیکسل در هر بار اعمال می‌شود. با اتصال نتیجه هر دو، یک نسخه با رزولوشن بالا در خروجی نهایی به دست می‌آید، به جز پیکسل‌هایی که به علت پیچش‌های معتبر از بین رفته‌اند.

در مرجع [۳۵] از ایده $fCNN$ یک گام فراتر رفتند و ساختار U-net را پیشنهاد کردند که شامل یک قسمت "منظم" از شبکه CNN بود که در آن پیچش‌ها برای افزایش اندازه تصویر، انقباضی و

^۱ voxel

^۲ fully Convolutional Neural Network (fCNN)

مسیرهای طولانی استفاده می‌شوند. یک روش مشابه توسط [۳۶] برای داده‌های سه بعدی استفاده شد. در [۳۷] یک پسوند برای چیدمان U-Net پیشنهاد کردند که بلوک‌های باقیمانده و یک لایه اتلاف را به حداقل می‌رساند.

۲-۲-۷ طبقه‌بندی تصویر

طبقه‌بندی تصویر، یکی از اولین حوزه‌هایی بود که در آن یادگیری عمیق، کمک بزرگی به آنالیز تصویر پزشکی کرد. در طبقه‌بندی تصاویر پزشکی معمولاً تصاویر به عنوان ورودی و یک مقدار منفرد (برچسب) به عنوان خروجی (به عنوان مثال، بیماری وجود دارد یا خیر) در نظر گرفته می‌شود. در چنین شرایطی، هر آزمون، تشخیص برچسب یک نمونه است و اندازه مجموعه داده‌ها معمولاً در مقایسه با آن‌هایی که در بینایی کامپیوتری قرار دارند، کوچکتر (به عنوان مثال، صدها/هزارها در مقابل میلیون‌ها نمونه) هستند. در واقع تعداد نمونه‌های مورد نیاز بسیار کمتر از نمونه‌های موجود می‌باشد. بنابراین محبوبیت یادگیری انتقالی برای چنین کاربردهایی تعجب‌آور نیست.

دو استراتژی یادگیری انتقال شناسایی شده‌اند. اولین استراتژی استفاده از یک شبکه از پیش آموزش دیده به عنوان استخراج‌کننده ویژگی است. استراتژی دوم تنظیم یک شبکه از پیش آموزش دیده بر روی داده‌های پزشکی است. راهبرد پیشین این مزیت اضافی را دارد که به کسی نیاز ندارد تا یک شبکه عمیق را آموزش دهد، و اجازه می‌دهد که به صورت موازی مشخصه‌های استخراج شده به راحتی به تحلیل تصویر موجود متصل شوند. هر دو روش راهبردی محبوب هستند و به‌طور گسترده اعمال شده‌اند.

در برخی مقاله‌ها از ساختاری استفاده کردند که ویژگی‌های منحصر به فرد داده‌های پزشکی را به کار می‌برد. در [۳۸] از یک ساختار CNN مانند یک نمودار اتصال مغزی ناشی از تصویربرداری MRI – تانسور (DTI) استفاده کردند. برای انجام این کار، آن‌ها چندین لایه جدید ایجاد کردند که اساس شبکه آن‌ها را تشکیل می‌داد. در این مقاله اتصالاتی به اصطلاح لبه به لبه، لبه به گره، و لایه‌های گره

به گراف را تعریف نمودند. آن‌ها از شبکه خود برای پیش‌بینی رشد مغز استفاده کردند و نشان دادند که آن‌ها از روش‌های موجود در ارزیابی نمرات شناختی و حرکتی برترند. یکی از تکنیک‌های استاندارد در این زمینه روش‌های انتزاعی در طبقه‌بندی می‌باشد. به‌ویژه CNN‌های از پیش آزمایش داده شده در تصاویر طبیعی به طرز شگفت‌آوری نتایج قطعی را نشان داده‌اند و دقت متخصصان انسانی را در برخی کارها به چالش کشیده‌اند. در آخر، نویسندگان نشان داده‌اند که CNN‌ها را می‌توان برای استفاده از ساختار درونی تصاویر پزشکی تطبیق داد.

۲ - ۲ - ۸ طبقه‌بندی شیء یا ضایعه

طبقه‌بندی شیء معمولاً بر طبقه‌بندی یک قسمت کوچک (قبلاً شناسایی‌شده) تصویر پزشکی به دو یا چند طبقه (مانند مثال طبقه‌بندی گره‌ها در CT قفسه سینه) تمرکز می‌کند. برای بسیاری از این وظایف هر دو اطلاعات محلی در مورد ظاهر ضایعه و اطلاعات بافتی سراسری در محل ضایعه برای طبقه‌بندی دقیق مورد نیاز است. این ترکیب معمولاً در ساختارهای یادگیری عمیق عمومی امکان‌پذیر نیست.

در [۳۹] از سه CNN استفاده کردند که هر کدام یک تکه گره در مقیاس مختلف به عنوان ورودی می‌گیرند. سپس خروجی‌های ویژگی سه CNN باهم ادغام می‌شوند تا بردار ویژگی نهایی را شکل دهند. یک رویکرد تا حدی مشابه در [۴۰] پیشنهاد شد که از CNN چند جریانی برای طبقه‌بندی ضایعات پوست استفاده کردند که در آن هر جریان بر روی وضوح متفاوتی از تصویر عمل می‌کند. در [۴۱] پیشنهاد کردند که از ترکیبی از CNN‌ها و RNN‌ها برای درجه‌بندی آب‌مروارید هسته‌ای در تصاویر استفاده شود. در آن فیلترهای CNN پیش از آموزش قرار گرفته است. این ترکیب اجازه پردازش همه اطلاعات متنی بدون توجه به اندازه تصویر را می‌دهد.

استفاده از اطلاعات سه بعدی نیز اغلب برای عملکرد خوب در فعالیت‌های طبقه‌بندی شیء در تصویربرداری پزشکی ضروری است. در [۴۲] از طبیعت سه بعدی MRI با آموزش یک CNN سه

بعدی برای ارزیابی بقا در بیمارانی که از گلیوما^۱ (یک توده بدخیم در بافت سیستم عصبی) درجه بالا رنج می‌برند، بهره برداری کردند.

تقریباً تمام مقالات اخیر استفاده از CNNهای آموزش دیده انتها به انتها را ترجیح می‌دهند. در برخی موارد، ساختارها و رویکردهای دیگر مانند RBMها [۴۳، ۴۴]، SAEها [۴۵] و خودرمنگاره‌های پراکنده پیچشی [۴۶] مورد استفاده قرار می‌گیرند. تفاوت عمده بین CSAE و یک CNN کلاسیک، استفاده از پیش آموزش بدون نظارت با خودرمنگاره‌های پراکنده است.

۲ - ۳ طبقه‌بندی تصاویر پزشکی بر پایه ویژگی‌ها

برای پردازش تصاویر پزشکی در سال‌های اخیر از روش‌هایی برپایه ویژگی‌های سطح پایین و سطح بالا استفاده شده است. در ویژگی‌های سطح پایین از خصیصه‌هایی به مانند تشخیص لبه و مدل‌سازی ریاضی برای تشخیص خطوط، دایره‌ها و اشکال برای تصاویر پزشکی بکار گرفته شده است. سپس خروجی بر اساس جداول شرطی if-else تعیین می‌شوند [۹]. در سال‌های بعد، تکنیک‌های بانظارت بر اساس مجموعه آموزش جمع آوری شده [۴۷] و روش‌های مبتنی بر شناسایی الگو با استفاده از یادگیری ماشین در تجزیه و تحلیل تصاویر پزشکی مورد توجه بسیاری از محققان قرار گرفته است [۴۸].

استخراج ویژگی‌های مناسب نقش اساسی در طبقه بندی تصویر پزشکی دارد [۴۹]. روش‌های استخراج ویژگی به طور کلی به دو دسته تقسیم می‌شوند. دسته اول شامل استخراج ویژگی‌های سنتی است که در آن یک مرحله مجزا برای استخراج ویژگی با تمرکز بر استخراج ویژگی سطح پایین می‌باشد. در دسته دوم، مرحله استخراج ویژگی بر اساس شبکه‌های عصبی عمیق است که به عنوان ویژگی‌های سطح بالا یا ویژگی‌های نوین معرفی می‌شوند. در واقع در این روش‌ها مرحله استخراج

¹ gliomas

ویژگی به عنوان یک مرحله جداگانه در نظر گرفته نمی شود و در آن مرحله آموزش با روشی بدون نظارت ادغام می شود. در این روش ها، چندین لایه برای استخراج ویژگی استفاده می شود. در شبکه های عمیقی که برای استخراج ویژگی استفاده شده است، ویژگی های سطح پایین در لایه های اول استخراج می شوند در حالی که ویژگی های سطح بالا در لایه های بعدی با ترکیب ویژگی های سطح پایین لایه های قبل استخراج می شوند [۹، ۴۸، ۵۰].

نوع متداول شبکه های عصبی عمیق به نام شبکه عصبی پیچشی برای اولین بار در شناسایی ارقام دست نویس استفاده شده است [۵۱]. علی رغم دستاورد قابل توجه در شبکه های مبتنی بر CNN ولی این شبکه ها به دلیل کمبود داده های حجیم برای آموزش، الزامات سخت افزاری به مانند واحد پردازش گرافیکی^۱ (GPU) با چالش های زیادی روبرو بوده است. از سال ۲۰۱۲ به بعد با پیشرفت های سخت افزاری خصوصاً GPU، تحولات جدید در تکنیک های یادگیری عمیق، معرفی معماری های عمیق نوین و تهیه مجموعه داده بزرگ مانند ImageNet عملکرد این شبکه های مبتنی بر CNN بسیار پیشرفت کرده است. از مهمترین معمارها می توان به AlexNet [۵۲]، VGG [۵۳]، GoogleNet [۲۷] و ResNet [۵۴] اشاره نمود که در کاربردهای بسیاری بخصوص در آنالیز تصاویر پزشکی استفاده شده است [۱۱، ۵۵].

امروزه، به دلیل تکامل مدل های یادگیری عمیق مانند CNN، عملکرد طبقه بندی تصاویر به طور قابل توجهی بهبود یافته است. با وجود دستاوردهای مهم در استخراج ویژگی بدون نظارت، تنوع بین طبقاتی و شباهت درون طبقاتی هنوز موضوعات چالشی است. برای رفع این مشکلات، مطالعات مختلفی در طیف وسیعی از وظایف طبقه بندی تصویر از طبقه بندی تصویر چند-شات^۲ گرفته تا طبقه بندی تصویر سنجش از دور انجام شده است [۵۶، ۵۷].

شبکه های عصبی عمیق به دلیل میلیون ها پارامتر (به عنوان مثال وزن) برای یادگیری در مرحله آموزش، به مجموعه های بزرگ آموزشی نیاز دارند. متأسفانه، تعداد محدود تصاویر یک مسئله چالش

^۱ Graphical Processing Unit (GPU)

^۲ few-shot image

برانگیز در زمینه تجزیه و تحلیل تصویر پزشکی است. در چنین مواردی، یادگیری انتقالی^۱ می‌تواند به عنوان یک راه حل بالقوه در نظر گرفته شود. برای یادگیری انتقال دو استراتژی کلی وجود دارد. در استراتژی اول، شبکه‌های از پیش آموزش دیده به عنوان ویژگی‌های استخراج استفاده می‌شود. استراتژی دوم شبکه‌ای را که از قبل آموزش دیده است در یک مجموعه آموزش جدید تنظیم می‌کند. روش‌های زیادی مانند روش بازیابی تصویر مبتنی بر محتوا^۲ (CBIR) اخیراً برای استخراج مفاهیم سطح بالا از تصویربرداری پزشکی استفاده شده است. CBIR روشی برای اکتشاف دانش در مجموعه داده‌های بزرگ و تشخیص وقایع مشابه است، درک اختلالات نادر و بهبود مراقبت از بیمار را فراهم می‌کند. چالش اصلی در روش‌های CBIR استخراج بازنمایی ویژگی‌های موثر از اطلاعات در سطح پیکسل و ارتباط آنها با مفاهیم معنایی^۳ است [۵۸]. توانایی مدل‌های عمیق CNN برای یادگیری ویژگی‌های غنی در سطوح مختلف انتزاعی، جامعه CBIR را گسترش داد [۵۹]. تمام روش‌های فعلی از CNN آموزش دیده برای استخراج توصیفگر ویژگی از تصاویر پزشکی استفاده می‌کنند [۶۰]. آناوی و همکاران [۶۱] از یک شبکه پنج لایه CNN قبلاً آموزش دیده استفاده کردند که در آن از یک لایه نهایی کاملاً متصل^۴ برای استخراج ویژگی‌ها استفاده شده است. آنها خروجی لایه نهایی را به عنوان ورودی ماشین بردار پشتیبانی^۵ (SVM) برای طبقه بندی تصاویر اشعه ایکس آماده کردند. شاه و همکاران [۶۲] برای طبقه بندی تصاویر MRI از پروستات، توصیفگرهای ویژگی CNN را با جنگل درهم‌سازی ترکیب کرد. آنها نشان دادند که این ترکیب، توصیفگرهای ویژگی بهتری را با عملکرد بالاتر ارائه می‌دهد. از آنجا که مطالعات کمی بر روی تصاویر مثانه از تصاویر سیستم اسکوپ متمرکز شده است، مروری بر مطالعات یادگیری عمیق در مهندسی پزشکی ارائه شده است.

¹ transfer learning

² Content-Based Image Retrieval (CBIR)

³ Semantic concepts

⁴ fully connected

⁵ Support Vector Machine (SVM)

برخی از مطالعات برای طبقه بندی اختلالات مغزی از CNN های دو بعدی و سه بعدی استفاده کرده‌اند [۶۳-۶۶]. شی و همکاران [۶۳] شبکه‌های چند جمله‌ای عمیق انباشته چند حالت^۱ را برای طبقه‌بندی اختلالات مغزی با استفاده از طبقه بندی SVM از تصاویر MRI اعمال کرده است. برخی از محققان از CNN و شبکه‌های عصبی بازگشتی^۲ (RNN) برای تقسیم بندی تصویر مغز بر اساس بافت/آناتومی/ضایعه/تومور استفاده کردند. به عنوان مثال، کامنیتساس و همکاران [۲۸] از شبکه کاملاً کانولوشن چند مقیاس سه بعدی برای تقسیم‌بندی تصاویر تومور مغزی استفاده کرد و نشان داد که روش آنها نسبت به روش‌های اخیر برتر است.

اخیراً، تشخیص فتودینامیکی^۳ مبتنی بر سیستم‌سکوپی نور آبی^۴ (BL) به عنوان یک فناوری منحصر به فرد برای افزایش تشخیص سرطان مثانه، به ویژه برای تشخیص ضایعات صاف و کوچک معرفی شد [۶۷، ۶۸]. در همین راستا در مرجع [۶۷]، یک بستر تشخیصی مبتنی بر هوش مصنوعی با استفاده از ۲۱۶ تصویر BL ارائه شد. در این تحقیق از چهار شبکه عصبی کانولوشن پیش آموزش دیده برای پیش بینی بدخیمی^۵ تصویر، تهاجم و درجه‌بندی^۶ استفاده شد. نتایج نشان داد که استفاده از یادگیری عمیق می‌تواند در این نوع از کاربردها عملکرد خوبی داشته باشند. در مرجع [۶۸]، یک مجموعه داده از تصاویر سیستم‌سکوپی ۲ کلاسه با ۲۱۰۲ تصویر معرفی شد که این مجموعه داده شامل ۴۳۱ تصویر با ضایعات تومور و ۱۶۷۱ بافت سالم می‌باشد. برای طبقه‌بندی این نوع از تصاویر از شبکه عصبی پیچشی استفاده شده است.

طبقه‌بندی صحیح تصاویر سیستم‌سکوپی به طور مستقیم با تجربه خبره ارتباط داشت. از آنجا که سرطان مثانه ضایعه‌ای است که فقط با نمونه‌برداری از بافت قابل تأیید است، تشخیص خودکار تومورها در تشخیص مراحل اولیه و دقت آن نقش مهمی دارد. فریتاس و همکاران با فرض تغییرات

¹ multi-modal stacked deep polynomial networks

² Recurrent Neural Networks (RNN)

³ photodynamic

⁴ Blue Light (BL)

⁵ malignancy

⁶ grading

تومور در بافت، روش تجزیه و تحلیل بافت را برای تشخیص تومورهای مثانه پیشنهاد کرد. آنها برای بدست آوردن اطلاعات بافت از تبدیل موجک گسسته^۱ (DWT) استفاده کردند. علاوه بر این، آنها از تقسیم‌بندی خودکار برای بهبود تقویت تومور و جلوگیری از ترکیب تومور با بافت‌های طبیعی استفاده کردند. علاوه بر این، آنها از مازول تقسیم‌بندی مبتنی بر درخت تجزیه موجک برای دور انداختن اطلاعات بافت ضعیف استفاده کردند. همچنین، آنها از پرسپترون چند لایه^۲ (MLP) و SVM با یک روش اعتبار سنجی متقابل^۳ برای طبقه بندی استفاده کردند. بعلاوه، لازم به ذکر است که از فضاهای رنگی مقدار اشباع رنگ (HSV)، قرمز-سبز-آبی و CIE Lab استفاده شده است. آنها ادعا کردند که روش پیشنهادی عملکرد خوبی در شناسایی قابهای تومور مثانه دارد [۶۹].

۲ - ۴ جمع بندی

در این فصل کارهای انجام شده برای طبقه‌بندی و تحلیل تصاویر پزشکی به تفصیل بیان شد. یک فاز مهم در این نوع از سیستم‌ها استخراج ویژگی می‌باشد. در همین راستا در این فصل، روش‌های استخراج ویژگی در تصاویر پزشکی به دو دسته کلی، ویژگی سطح-پایین یا سنتی به مانند لبه و بافت، و سطح-بالا معنایی که بر پایه شبکه عصبی عمیق می‌باشند تقسیم شد. همانطور که بیان شد به دلیل پیچیدگی بالای تصاویر و محدودیت‌های موجود از لحاظ تعداد و نحوه تصویربرداری ویژگی‌های سطح پایین مانند لبه یا شکل‌ها نمی‌تواند به تنهایی عملکرد مناسبی داشته باشد. از این‌رو در این رساله به بررسی روش‌های قدرتمند برپایه یادگیری عمیق پرداخته شده است. در این فصل اشاره شد که به دلیل هزینه بالا و زمانبر بودن تهیه تصاویر سیستم‌سکوپی آموزش شبکه عمیق به دلیل پارامترهای زیاد در کاربرد امکان‌پذیر نمی‌باشد. در همین راستا به تحلیل و بررسی روش‌های بر پایه

¹ discrete wavelet transform (DWT)

² multi-layer perceptron (MLP)

³ cross-validation

یادگیری انتقال بر اساس وزن‌های از پیش‌آموزش داده شده پرداخته شده است. مطالب بیان شده در این فصل نشان می‌دهد که استفاده از روش‌های موجود برپایه ویژگی‌های سطح پایین نمی‌توانند در این نوع از محیط‌ها عملکرد مناسبی از خود نشان دهند. با توجه به مطالب ارائه شده در این فصل، به دست آوردن ویژگی‌های مناسب در تصویر پزشکی می‌تواند در بالا بردن دقت سیستم طبقه‌بندی تصویر بسیار موثر باشد. در همین راستا در فصل بعدی روشی جامع بر اساس این روش‌ها معرفی می‌شود.

۳ روش پیشنهادی در طبقه‌بندی تصاویر سیستم اسکوپ میثانه

۳ - ۱ مقدمه

در این فصل روش‌های پیشنهادی در این رساله معرفی می‌شود. ویژگی‌های معرفی شده به دو دسته کلی ویژگی‌های سطح پایین و سطح بالا تقسیم می‌شود. در روش اول، از ویژگی LBP برای استخراج ویژگی استفاده شده است. الگوریتم LBP یکی از الگوریتم‌های محبوب و پرکاربرد استخراج ویژگی برپایه بافت^۱ در علم بینایی ماشین است. این روش، از هیستوگرام برای استخراج ویژگی‌ها استفاده می‌کند. از آنجایی که الگوی دودویی محلی هم از مشخصه‌های آماری و هم ساختاری بافت استفاده می‌کند یک ابزار قدرتمند برای تحلیل بافت می‌باشد. در اپراتور الگوی دودویی محلی، الگوهای بافت محلی بوسیله مقایسه مقدار پیکسل‌های مجاور با مقدار پیکسل مرکزی استخراج شده و با کدهای دودویی نشان داده می‌شوند. برای طبقه‌بندی از شبکه عصبی MLP با یک نرخ یادگیری تطبیقی و انتخاب وزن‌های اولیه بر اساس الگوریتم ژنتیک استفاده شده است. نتایج بدست آمده نشان دهنده بهبود عملکرد روش پیشنهادی نسبت به روش‌های دیگر بر پایه ویژگی‌های سنتی بوده است. ولی با توجه به پیچیدگی بالای تصاویر ویژگی بافت به تنهایی نمی‌تواند در این کاربرد نتیجه قابل قبولی در دنیا واقعی ارائه دهد. برای استخراج ویژگی‌های سطح بالا، از شبکه عصبی عمیق از پیش‌آموزش دیده شده بر پایه هسته‌های متفاوت در لایه‌های کانولوشن استفاده شده است. در ادامه برای طبقه‌بندی از ترکیب طبقه‌بندهای پایه مانند گوسین بیزین و ماشین بردار پشتیبان استفاده شده است. نتایج نشان دهنده عملکرد بالای روش بر پایه ویژگی‌های سطح بالا می‌باشد. برای انتخاب بهتر ویژگی‌ها از روش بخش‌بندی نواحی مربوط به نگاشت ویژگی استفاده شده است. در نتایج نشان داده شده است که انتخاب ویژگی سطح بالا از نواحی مناسب تصاویر سیستم‌سکوپی سبب بهبود طبقه‌بندی می‌شود. در ادامه، روش‌های پیشنهادی و نقاط قوت و ضعف هریک به تفصیل بررسی می‌شوند.

¹ Texture

۳ - ۲ روش پیشنهادی برای ویژگی‌های سطح پایین

در این بخش، روش پیشنهادی برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه بر پایه ویژگی‌های سطح پایین تشریح می‌شود. مراحل اصلی روش پیشنهادی شامل استخراج و کاهش ویژگی، بهبود شبکه عصبی با نرخ یادگیری تطبیقی و به دست آوردن وزن‌های اولیه توسط الگوریتم ژنتیک است که در ادامه به تشریح این روش می‌پردازیم. در این روش، یک رویکرد نوین از MLP برای پیش‌بینی و طبقه‌بندی تصاویر سیستم‌سکوپی با دقت و صحت بالا ارائه شده است. تعیین نرخ یادگیری برای پایداری فرآیند آموزش بسیار با اهمیت هست. نرخ یادگیری یک هایپرپارامتر است که میزان تغییر مدل در پاسخ به خطای تخمینی در هربار به روز رسانی وزن مدل را کنترل می‌کند. انتخاب مقدار نرخ یادگیری چالش برانگیز است چرا که یک مقدار خیلی کوچک ممکن است منجر به یک فرآیند طولانی آموزش شود که می‌تواند باعث کندی عملکرد ماشین شود، در حالی که یک مقدار خیلی بزرگ ممکن است سبب روند آموزشی ناپایدار شود. پس تعیین نرخ یادگیری بزرگ می‌تواند سیستم را ناپایدار کند در همین راستا باید جهت جلوگیری از ناپایداری، نرخ یادگیری تا اندازه‌ای کوچک شود. باید به این نکته توجه نمود که نرخ یادگیری خیلی کوچک سبب طولانی شدن زمان آموزش می‌شود. در همین راستا در این رساله روشی تطبیقی برای انتخاب مناسب نرخ یادگیری ارائه می‌شود.

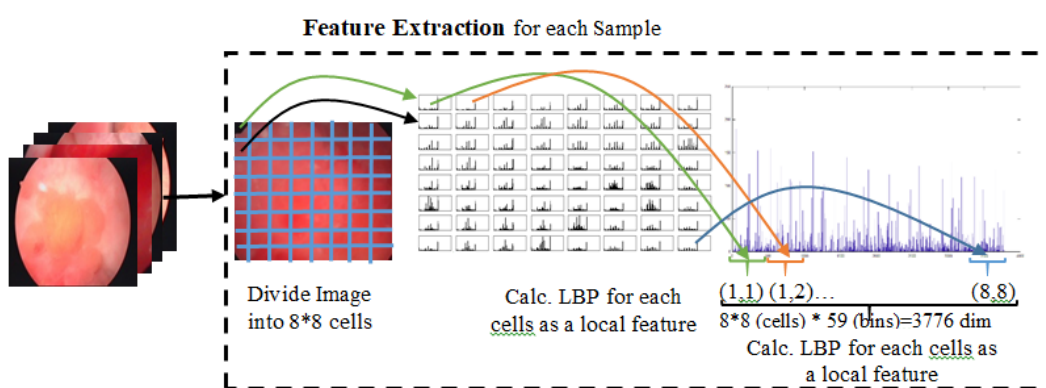
انتخاب وزن‌های اولیه در بهبود عملکرد شبکه MLP بسیار تاثیرگذار هست. معمولاً از مقادیر تصادفی برای وزن‌های اولیه استفاده می‌شود. استفاده از وزن‌های اولیه تصادفی در مجموعه داده‌های کوچک می‌تواند اثرات مخربی داشته باشد و آموزش شبکه به خوبی انجام نمی‌شود. معمولاً تصاویر پزشکی شامل مجموعه داده‌های بزرگ نمی‌باشند و تعداد نمونه‌های استخراجی از بیماران کم است. در همین راستا، نباید از وزن‌های اولیه تصادفی برای آموزش شبکه استفاده نمود. در واقع در مجموعه داده‌های کوچک باید وزن‌های اولیه با دقت مناسب‌تری انتخاب شوند. در همین راستا در این بخش، یک روش نوین بر پایه الگوریتم ژنتیک برای انتخاب وزن‌های اولیه شبکه MLP ارائه شده است.

۳ - ۲ - ۱ استخراج ویژگی LBP و کاهش ابعاد

یکی از گام‌های مهم و اساسی در تشخیص و درمان بیماری‌ها استخراج ویژگی از تصاویر پزشکی است. در این راستا باید از ویژگی‌های مقاوم و قدرتمند استفاده نمود. در این بخش از الگوی باینری محلی (LBP) برای استخراج ویژگی استفاده شده است. الگوی باینری محلی بعلاوه مقاومتش نسبت به تغییرات روشنایی، پیچیدگی محاسباتی کم و توانایی کدگذاری جزئیات یکی از رایج‌ترین توصیفگرها می‌باشد. الگوی باینری محلی مستقل از چرخش، توسعه‌ای دیگر از الگوی محلی باینری است. زمانی که یک تصویر در صفحه می‌چرخد تمام همسایه‌ها حول پیکسل مرکزی در یک جهت خواهند چرخید. این اثر چرخش منجر به مقادیر متفاوتی برای الگوی باینری محلی می‌شود. برای این منظور محققین توسعه‌ی جدیدی از الگوی محلی باینری که مستقل از چرخش می‌باشد را معرفی کرده‌اند [۷۰]. عبارت مستقل از چرخش در اینجا مربوط به تغییرات مکانی وابسته به تغییرات نور و یا اشیا متفاوت نمی‌باشد. از عملگر چرخشی بیتی دایره‌ای به سمت راست برای برطرف کردن اثر نامطلوب چرخش استفاده می‌شود تا با چندین تکرار تمامی کدهای دودویی که با این بیت‌ها می‌تواند تولید کرد را بدست آورد و سپس کمترین مقدار دهنده‌ی را از الگوی باینری بدست آورد. با توجه به ماهیت تصاویر سیستم‌سکوپی این ویژگی می‌تواند عملکرد خوبی در این نوع از تصاویر داشته باشد.

الگوریتم الگوی باینری محلی یکی از قوی‌ترین الگوریتم‌های استخراج ویژگی در علم بینایی ماشین است و همچنین به طور وسیعی در تحقیقات مربوط به تشخیص چهره و بازیابی چهره به کار رفته است. این الگوریتم روشی مؤثر برای توضیح کارایی بافت‌ها است که می‌توان از آن برای اندازه‌گیری استخراج ویژگی‌های بافت‌های مجاور در تصاویر استفاده نمود. از مهم‌ترین مزیت‌های استفاده از این روش می‌توان به مقاوم بودن نسبت به چرخش و مقاوم بودن در تغییرات بالای سطح خاکستری اشاره نمود که می‌توان با استفاده از آن بر مشکلات عدم تعادل در تغییر موقعیت، چرخش و نورافکنی، غلبه نمود. علاوه بر این عملگر LBP دارای محاسبات بسیار ساده‌ای است. با توجه به سرعت بالا و سادگی در محاسبات به راحتی می‌توان از این ویژگی در کاربردهای پزشکی استفاده

نمود. در همین راستا، از این روش برای استخراج ویژگی از تصاویر سیستم اسکوپ می‌ماند استفاده شده که در شکل ۱-۳ نشان داده شده است. همانطور که در شکل نشان داده شده تصاویر ورودی به سلول‌های 8×8 با ابعاد 32×32 تقسیم شد و سپس LBP برای هر سلول به صورت مجزا محاسبه شد. در نهایت این ویژگی‌ها جمع شده است. تعداد ویژگی‌های استخراج شده در این روش بسیار بالا بوده و با توجه به تعداد کم تصاویر سیستم اسکوپ نمی‌توان از شبکه عصبی با تعداد ورودی بالا برای آموزش استفاده نمود در همین راستا از روش تجزیه و تحلیل مؤلفه‌های اصلی برای کاهش ویژگی استفاده شده است.

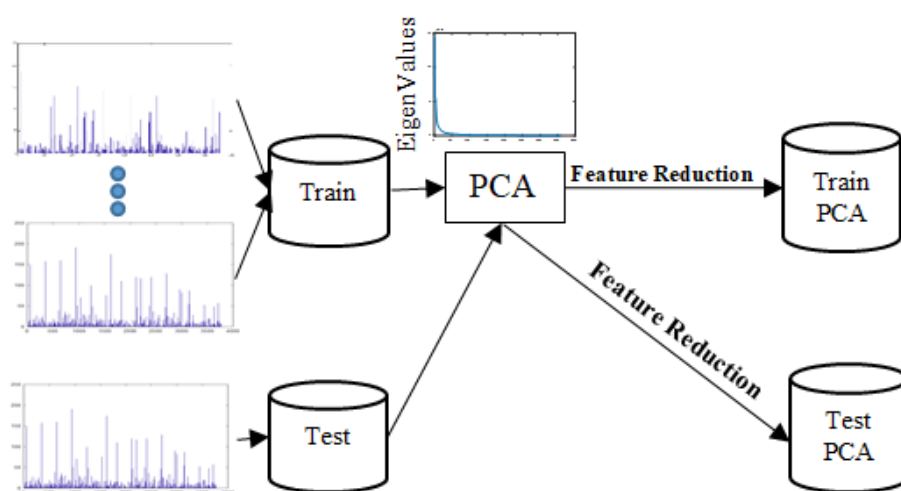


شکل ۱-۳: مراحل استخراج ویژگی در روش پیشنهادی توسط روش LBP

یکی از قدیمی‌ترین و معروف‌ترین روش‌های کلاسیک چند متغیره برای کاهش ویژگی، تحلیل مؤلفه‌های اصلی است. این روش ابتدا به وسیله پیرسون به عنوان وسیله‌ای برای برآوردن صفحات از طریق حداقل مربعات متعامد معرفی شد و سپس به منظور تجزیه و تحلیل ساختارهای ماتریس‌های واریانس-کواریانس و ضریب همبستگی توسعه داده شد [۷۱]. این روش در کاربردهای مانند یافتن ترکیبات خطی با تغییرپذیری نسبی بزرگ یا کوچک، کاهش حجم داده‌ها و تفسیر داده‌ها عملکرد مناسبی از خود نشان داده است. کاهش حجم داده‌ها هدف اصلی این تجزیه را تشکیل می‌دهد که این داده‌ها شامل تعداد زیادی متغیرهای با همبستگی‌های درونی می‌باشند به طریقی که حداکثر اطلاعات موجود در داده‌ها محفوظ بماند. این امر از طریق تبدیل داده‌ها (متغیرها) به متغیرهای جدیدی است

که مؤلفه‌های اصلی نامیده شده و غیر همبسته هستند. این مؤلفه‌های جدید به ترتیبی اولویت‌بندی می‌شوند که تعداد اندکی از آن‌ها اغلب تغییرات موجود در متغیرهای اولیه را با خود به همراه دارند.

در این رساله با توجه به تعداد بالای ویژگی‌های استخراجی از تصاویر پزشکی توسط LBP از این روش برای کاهش ویژگی استفاده می‌شود. همانطور که در شکل ۳-۲ نشان داده شده است، بعد از محاسبه ماتریس ویژه، ویژگی‌های جدید نگاشت شده در فضای PCA بر اساس ارزش و اهمیت آن‌ها در مقادیر ویژه محاسبه می‌شوند.



شکل ۳-۲ کاهش ویژگی در روش پیشنهادی بر پایه ویژگی سطح پایین

۳-۲-۲ شبکه عصبی چندلایه پرسپترون

این الگوریتم جزو الگوریتم‌های یادگیری با ناظر است که اساساً از دو مسیر اصلی تشکیل شده است. مسیر رفت، که در آن بردار ورودی به شبکه عصبی اعمال شده و تأثیر آن از طریق لایه‌های میانی به لایه خروجی انتشار می‌یابد. در این مسیر برای هر ورودی، مقداری تحت عنوان خروجی توسط شبکه محاسبه می‌شود. در این مسیر، پارامترهای شبکه ثابت می‌مانند. مسیر برگشت پس از تولید خروجی در مرحله رفت، اختلاف خروجی مطلوب و خروجی محاسبه شده توسط شبکه تعیین می‌شود. سیگنال‌های خطا در مسیر برگشت از لایه خروجی مجدداً در کل شبکه توزیع می‌شود و پارامترهای شبکه مجدداً تنظیم می‌شوند. فرآیند دوگانه فوق به دفعات تکرار شده تا خروجی شبکه به

خروجی مطلوب نزدیک شود. هنگامی که خطای به دست آمده از آستانه مجاز کمتر شود فرآیند آموزش متوقف می‌شود. فضای فرضیه مورد جستجو در این روش عبارت است از فضای بزرگی که توسط همه مقادیر ممکن برای وزن‌ها تعریف می‌شود. روش گرادیان نزولی سعی می‌کند تا با حداقل کردن خطا به فرضیه مناسبی دست پیدا کند. اما تضمینی برای اینکه این الگوریتم به حداقل مطلق برسد وجود ندارد. معمولاً الگوریتم BP پیش از خاتمه هزاران بار، با همان داده‌های آموزشی تکرار می‌گردد. شروط مختلفی مانند توقف بعد از تکرار به دفعات معین، توقف وقتی که خطا از یک مقدار تعیین شده کمتر شود و توقف وقتی که خطا در مجموعه اعتبارسنجی از قاعده خاصی پیروی نماید را می‌توان برای خاتمه الگوریتم به کار برد.

۳ - ۲ - ۳ نرخ یادگیری تطبیقی

بعد از محاسبه ویژگی از تصاویر مجموعه آموزش و آزمون باید روشی برای طبقه‌بندی تصاویر استفاده نمود. همانطور که در بخش‌های قبل اشاره شد، در این روش از الگوریتم شبکه عصبی چندلایه پرسپترون برای این منظور استفاده می‌شود. هدف طراحی شبکه عصبی چند لایه‌ای است که نرخ یادگیری در آن به صورت متغیر با توجه میزان طبقه‌بندی تصاویر سیستم‌سکوپی تعریف شود. به این صورت که خطای آموزش با اعتبارسنجی مقایسه می‌شود اگر اختلاف این خطاها زیاد باشد، بدین معنی است که تعمیم‌پذیری سیستم پایین است و شبکه احتمالاً به دلیل نرخ بالای یادگیری روی داده‌ها با مشکل روبرو شده است بنابراین این نرخ باید با یک ضریبی کاهش یابد اما در صورتی که این شکاف (اختلاف خطای آموزش و اعتبارسنجی) خیلی کم باشد آنگاه احتمال دارد که به دلیل نرخ یادگیری پایین شبکه به خوبی آموزش ندیده باشد بنابراین باید نرخ یادگیری با یک ضریب کم افزایش یابد پس باید نرخ یادگیری بر اساس اختلاف دقت آموزش و اعتبارسنجی تنظیم شود. در این پژوهش نرخ یادگیری بر اساس ایده یاد شده بر روی شبکه‌های عصبی بررسی شده است. در این روش نرخ تغییرات با توجه به میزان تغییرات وزن بروز رسانی می‌شود که اگر میزان تغییرات زیاد باشد

بروزرسانی بیشتر اتفاق می‌افتد و بالعکس. معمولاً میزان تغییرات نرخ یادگیری در لایه‌ها متفاوت است به طوری که در تکرارهای اول تغییرات وزن بیشتر بوده و به مرور این تغییرات کاهش پیدا می‌کند در اینجا با استفاده از تعریف π بازه برای تخمین خطا و تعیین ضریب نرخ یادگیری این کار انجام می‌شود. بر اساس خطا ایجاد شده ضریب نرخ یادگیری برای π مرحله بروز رسانی می‌شود. فرآیند به این صورت است که اگر خطای حاصل شده بیشتر از α و کمتر از β باشد (منظور از α حد آستانه پایین و β حد آستانه بالا است) در مرحله بعد نرخ یادگیری تغییر می‌کند و به همین صورت برای مقادیر خطاهای حاصل شده اعمال می‌شود که بر اساس مقدار افزایش خطا در هر α ، مقدار نرخ یادگیری افزایش می‌یابد و بر اساس مقدار کاهش خطا در هر α ، مقدار نرخ یادگیری کاهش می‌یابد و اگر خطای حاصل شده بیشتر از یک مقدار π باشد مقدار نرخ یادگیری 0.1 اعمال خواهد شد.

۳ - ۲ - ۴ وزن‌دهی اولیه به کمک الگوریتم ژنتیک

روش‌های تصادفی زیادی برای محاسبه وزن‌های اولیه در شبکه‌های عصبی استفاده می‌شود. وزن‌های تصادفی زمانی که تعداد نمونه‌های آموزش بالا باشد می‌تواند مفید و مؤثر باشد ولی اگر تعداد نمونه‌ها کم باشد معمولاً آموزش با مشکل روبرو می‌شود. در کاربردهای پزشکی معمولاً تعداد پایین تصاویر یک چالش اساسی است. در همین راستا روشی برای محاسبه وزن‌های اولیه ارائه می‌شود. برای رسیدن به ترکیب بهینه از وزن‌های شبکه می‌توان علاوه بر روش‌های معمولی مانند نزول گرادیان از روش‌های جستجو هوشمند مانند جستجوهای اکتشافی، آبکاری فلزات، الگوریتم‌های هوش جمعی و الگوریتم‌های تکاملی نیز استفاده نمود.

در این بخش برای این منظور از الگوریتم ژنتیک به عنوان یک روش تکاملی قدرتمند برای تخمین وزن‌های اولیه استفاده شده است. اولین قدم در طراحی یک سیستم بهینه‌سازی مبتنی بر الگوریتم ژنتیک، تعریف کروموزوم برای مدل کردن راه حل مسئله است. در اینجا، هر کروموزوم در بردارنده کلیه وزن‌های شبکه است. در شکل ۳-۳ مشخصات کروموزوم نشان داده شده است. همانطور

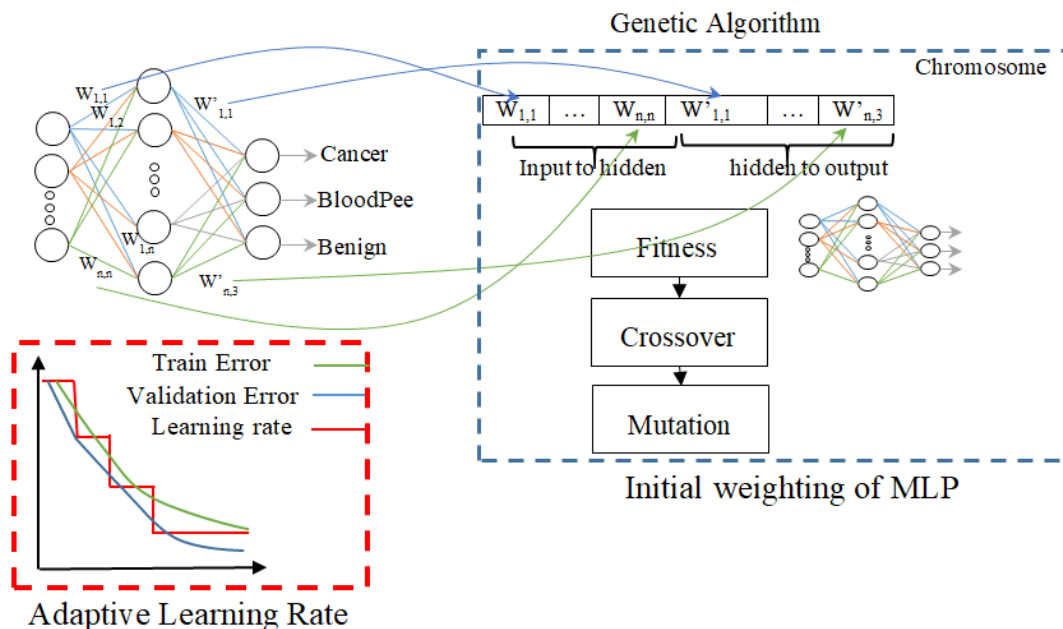
که مشاهده می‌شود یک کروموزوم نمونه که دارای برازندگی برتر باشد، حاوی تمام وزن‌های شبکه خواهد بود و لذا یک راه حل برای مسئله است. در مرحله بعد باید تابع برازش برای مسئله تعریف شود. برای این منظور تابعی با شبه کد زیر (الگوریتم ۳-۱) تعریف می‌شود:

الگوریتم ۳-۱ تابع برازندگی جهت محاسبه برازندگی کروموزوم‌های یک جمعیت

```

for i=1 to PopulationSize
  Decompose a Chromosome to Proper Weight Matrixes
  For j=1 to Number of Pattern
    Compute the Output of the network for the jth pattern
    E=error between desired and network output
    SumSquaredError +=dot(E)
  end
  fitness of ith Chromosome= sqrt (SumSquaredError ) / (NumberofPatterns);
end
    
```

نکته‌ای که در این روش وجود دارد، کند بودن الگوریتم ژنتیک نسبت به گرادیان نزولی است. در واقع این روش با تعداد تکرار پایین اجرا می‌شود تا وزن‌های مناسب استخراج شود و در ادامه از گرادیان نزولی برای محاسبه وزن‌های بهینه استفاده می‌شود.



شکل ۳-۳: روش پیشنهادی برای وزن دهی اولیه و ضریب یادگیری تطبیقی

در این بخش روشی پویا و قدرتمند بر پایه روش‌های هوشمند برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه ارائه شده است. در روش پیشنهادی بخش اول از LBP با تجزیه و تحلیل مؤلفه‌های اصلی برای استخراج و کاهش ویژگی استفاده شده است. در گام بعدی از روش شبکه عصبی پرسپترون چندلایه برای آموزش و ارزیابی طبقه‌بند تصاویر سیستم‌سکوپی مثانه استفاده شده است که برای بهبود عملکرد این روش، استفاده از نرخ یادگیری تطبیقی و به دست آوردن وزن‌های اولیه توسط الگوریتم ژنتیک ارائه شده است. با ارزیابی کارهای انجام شده، استفاده از ویژگی‌های سطح پایین به مانند LBP نمی‌تواند به تنهایی در این نوع از سیستم‌ها با توجه به تعداد پایین تصاویر و تفاوت در نحوه تصویری برداری در طبقه‌بندی تصاویر سیستم‌سکوپی استفاده شود. به همین دلیل در ادامه به معرفی روشی برپایه یادگیری انتقال برای استخراج ویژگی‌ها سطح بالا می‌پردازیم.

۳ - ۲ روش پیشنهادی بر اساس ویژگی سطح بالا بر پایه انتقال یادگیری

در روش پیشنهادی بخش قبل از LBP برای استخراج ویژگی استفاده شده است. در این مرحله می‌خواهیم از انتقال یادگیری بر اساس شبکه CNN برای استخراج ویژگی استفاده نمائیم. روش پیشنهادی به گونه‌ای طراحی شده است که نسبت به تغییرات شرایط تصویربرداری مستقل باشد. این روش با استخراج یک ویژگی سطح بالا از تصاویر سیستم‌سکوپی آغاز می‌شود. از آنجا که اثبات شده است که یادگیری عمیق در استخراج ویژگی با سطوح معنایی مختلف موثر است، این نوع شبکه برای استخراج ویژگی استفاده شده است. به منظور حل مشکل تعداد محدود تصاویر سیستم‌سکوپی، از یادگیری انتقالی استفاده می‌شود تا با استفاده از مجموعه آموزش کوچک موجود، یک مدل یادگیری مناسب ساخته شود. در رویکرد یادگیری انتقالی، آخرین لایه‌های یک شبکه از پیش آموزش دیده، که قبلاً روی یک مجموعه داده بزرگ مانند ImageNet آموزش دیده‌اند، استفاده شده است. برای جلوگیری از مشکل ابعاد بالا و همچنین کاهش زمان و پیچیدگی فضا، از تعداد ویژگی‌ها با استفاده از روش‌های کاهش بعد کاسته می‌شود. سرانجام، یک طبقه‌بند ترکیبی برای طبقه‌بندی تصاویر

سیستم‌سکوپی مثانه پیشنهاد می‌شود. در روش پیشنهادی برای ترکیب از طبقه‌بندی‌های شناخته شده با استفاده از روش آرای اکثریت وزنی استفاده شده است. روش پیشنهادی نتایج امیدوار کننده مربوط به طبقه‌بندی این نوع از تصاویر سیستم‌سکوپی را نشان می‌دهد.

یافته‌های اصلی روش به شرح زیر است:

(۱) روش‌های سنتی یادگیری ماشین قادر به ارائه نتایج قابل قبولی برای استخراج ویژگی با توجه به تعداد محدود تصاویر سیستم‌سکوپی نیستند. انتقال یادگیری، که مبتنی بر یادگیری عمیق است، یکی از موثرترین روش‌ها برای استخراج ویژگی‌ها از مجموعه کوچکی از داده‌های آموزش است. در این روش، آموزش اولیه بر اساس مجموعه ای گسترده از ImageNet انجام شده است و ویژگی‌های مرتبط استخراج می‌شوند. ویژگی‌های استخراج شده از نظر بصری با تصاویر سیستم‌سکوپی متفاوت هستند. از این رو، این ویژگی‌ها برای تصاویر سیستم‌سکوپی قابل استفاده هستند.

(۲) استفاده از یادگیری عمیق در استخراج ویژگی دارای مزایای زیر می‌باشد:

آ. استقلال ورودی: با توجه به ماهیت عمیق یادگیرنده، اندازه تصاویر ورودی در استخراج ویژگی در لایه‌های بالاتر تاثیری ندارد.

ب. مقاومت به چرخش و زاویه: حتی اگر تصاویر در یک محیط ایزوله و مشخص گرفته نشده باشد و تصاویر موقعیت‌های مختلف مربوط به نواحی بیماری گرفته باشد، یادگیری عمیق نتایج امیدوار کننده‌ای را نشان می‌دهد.

ج. مقاومت در برابر تغییرات روشنایی و محیط: ترکیب لایه‌های مختلف برای شبکه‌های عصبی عمیق، اثر تغییر محیط و شرایط روشنایی را کاهش می‌دهد. بنابراین، ویژگی‌های استخراج شده نسبت به محیط در حال تغییر مقاوم ارزیابی می‌شوند.

(۳) انتخاب مناسب معماری در یادگیری انتقال بسیار موثر است. در این روش، معماری‌های مختلف عمیق مانند VGG-16، VGG-19، AlexNet و ResNet بر روی تصاویر

سیستم‌سکوپی ارزیابی می‌شوند. نتایج نشان می‌دهد که معماری VGG-16 نسبت به مجموعه اعتبارسنجی و مجموعه آزمون از سایر معماری‌های رقیب بهتر است. از مهمترین ویژگی این نوع معماری می‌توان به موارد زیر اشاره نمود:

أ. اندازه تمامی فیلترهای ثابت و 3×3 می‌باشد و شیوه استفاده این فیلترها نیز نسبت به فیلترهای 11×11 معماری AlexNet و فیلترهای 7×7 معماری ZF net متفاوت است. دلیل این است که ترکیب دو لایه کانولوشن 3×3 ، همانند میدان تأثیر یک فیلتر 5×5 است. این تکنیک به نوبه خود یک فیلتر با اندازه بزرگتر را شبیه‌سازی می‌کند در حالیکه از مزایای فیلتر با سایر کوچکتر نیز بهره‌مند می‌شود. یکی از تأثیرات این تکنیک کاهش تعداد پارامترها است. همچنین با استفاده از دو لایه کانولوشن می‌توان از دو لایه ReLU استفاده کرد. در تصاویر سیستم‌سکوپی که تعداد تصاویر کم می‌باشد می‌تواند نسبت به معماری‌های عمیق‌تر عملکرد بهتری داشته باشد.

ب. سه لایه کانولوشن پشت سر هم یک میدان تأثیر 7×7 را دارد.

ت. از لایه‌های ReLU بعد از هر لایه کانولوشن استفاده شد و مدل توسط الگوریتم batch gradient descent آموزش داده شد.

ث. تعداد فیلترها بعد از هر لایه maxpool دو برابر می‌شود. این باعث می‌شود که بعد فضایی کاهش و عمق رشد کند.

۴) تعداد و ابعاد بالای ویژگی‌های استخراج شده باعث ایجاد برخی مسائل پیچیده برای مراحل بعدی تجزیه و تحلیل و تشخیص می‌شود. در این روش، PCA برای ارزیابی ویژگی‌ها و کاهش تعداد ویژگی‌ها استفاده شده است. نتایج نشان می‌دهد که کاهش ویژگی دقت طبقه‌بندی را بهبود می‌بخشد و پیچیدگی زمان را کاهش می‌دهد.

۵) انتخاب مناسب تعداد ویژگی‌های انتخاب شده بر عملکرد طبقه‌بندی در این حوزه تأثیر می‌گذارد. هنگامی که ویژگی‌ها توسط بردارهای ویژه PCA استخراج می‌شوند و

طبقه‌بندی‌کننده بر روی مجموعه اعتبار سنجی ارزیابی می‌شود، یک بردار ویژگی از ۱۰۰

ویژگی پس از مرحله نگاشت ویژگی برای طبقه‌بند استفاده می‌شود.

(۶) پس از کاهش ویژگی‌ها توسط PCA، LDA با در نظر گرفتن برچسب‌های کلاس استفاده

می‌شود. استفاده متوالی از LDA پس از ویژگی‌های وزن‌دهی توسط PCA، عملکرد طبقه

بندی تصاویر سیستم‌سکوپی را بهبود می‌بخشد.

(۷) استفاده از یک طبقه‌بند تنها برای کاربردهای بر اساس یادگیری ماشین معمول است در حالی

که روش‌های مختلف به دلیل پیچیدگی و تعداد کم تصاویر سیستم‌سکوپی نتایج متفاوتی را

ارائه می‌دهند. برای دستیابی به بهترین نتیجه، چندین روش تحلیل و ارزیابی می‌شود. در

مرحله بعدی، ترکیبی از طبقه‌بندهای پایه استفاده می‌شوند. آزمایشات نتایج امیدوار کننده‌ای

را به نفع روش پیشنهادی ترکیبی را نشان می‌دهد.

در این بخش، روش پیشنهادی برای طبقه‌بندی تصاویر سیستم‌سکوپی با جزئیات شرح داده شده

است. شکل ۳-۴ روش پیشنهادی را در چند مرحله نشان می‌دهد. انتقال یادگیری برای استخراج

ویژگی‌ها از تصاویر جمع‌آوری شده به دلیل تعداد محدودی از نمونه‌های آموزش و همچنین تعداد

زیادی از پارامترهای قابل آموزش در شبکه‌های عصبی عمیق استفاده می‌شود. در حقیقت، از وزن‌های

تنظیم شده قبلی ImageNet در مرحله یادگیری عمیق استفاده می‌شود.

یک معماری مناسب برای شبکه CNN می‌تواند با استخراج ویژگی‌های معنایی افتراقی بیشتر، دقت

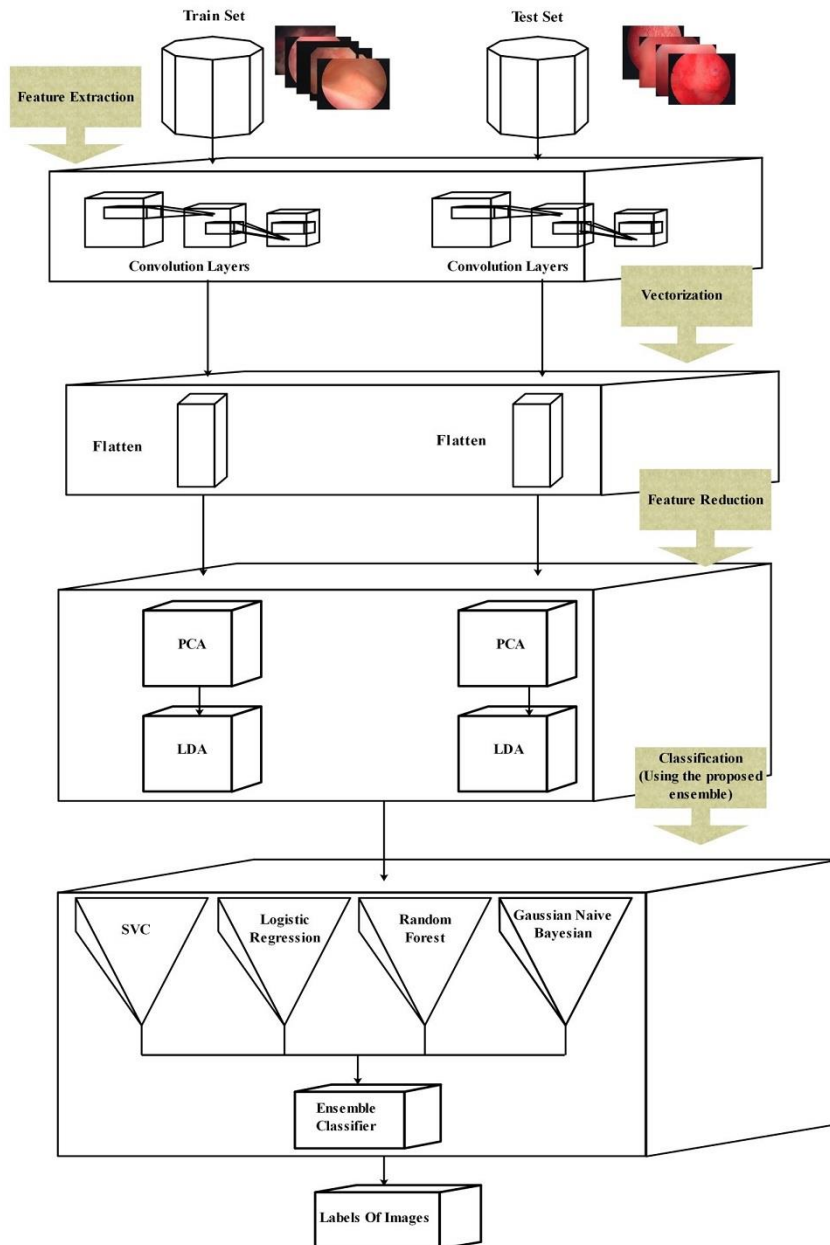
طبقه‌بندی را به میزان قابل توجهی بهبود بخشد. در این مطالعه، چهار معماری شامل VGG-16،

VGG-19، AlexNet و ResNet در مجموعه اعتبارسنجی ارزیابی می‌شوند تا یک معماری نهایی

حاصل شود. سپس، بردار ویژگی با استفاده از ویژگی‌های چند بعدی استخراج شده از لایه نهایی

شبکه عصبی پیچشی ساخته می‌شود. انتخاب لایه‌ها برای نگاشت ویژگی‌ها در نتایج نهایی بسیار موثر

است.



شکل ۳-۴ روش پیشنهادی بر پایه یادگیری انتقال و ترکیب طبقه‌بندها برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه

بنابراین، لایه ویژگی معمولاً پس از مشاهده خروجی‌های مختلف لایه انتخاب می‌شود. به منظور جمع‌آوری ویژگی‌ها، پس از نگاشت ویژگی انتخاب شده، از یک لایه کاملاً متصل استفاده می‌شود. از یک شبکه deconvolution برای ترسیم فضای ویژگی به فضای تصویر استفاده می‌شود. از آنجا که تعداد ویژگی‌های استخراج شده در شبکه‌های CNN زیاد است، روش‌های کاهش ابعاد مانند تجزیه و تحلیل مولفه‌های اصلی (PCA) و تجزیه و تحلیل خطی (LDA) برای بدست آوردن تعداد کمتری از ویژگی‌های افتراقی استفاده می‌شود. PCA پراکندگی و توزیع ویژگی‌ها را بدون در نظر گرفتن

برچسب‌های کلاس تجزیه و تحلیل می‌کند. LDA پراکندگی بین کلاس و درون کلاس را برای انتخاب ویژگی‌ها بر اساس برچسب‌های کلاس تجزیه و تحلیل می‌کند. سرانجام، یک طبقه‌بندی ترکیبی با استفاده از ترکیبی از طبقه‌بندی کننده های پایه شامل ماشین بردار پشتیبان^۱ (SVM)، رگرسیون لجستیک^۲ (LR)، گوسین بیزین^۳ (GB) و جنگل تصادفی^۴ (RF) ساخته می‌شود. برچسب طبقه‌بندی نهایی برای تصویر مثانه از طریق استراتژی رای اکثریت وزنی بدست می‌آید. مراحل مختلف روش پیشنهادی در زیر بخش‌های بعدی توضیح داده شده است.

۳ - ۳ - ۱ یادگیری انتقالی

یادگیری انتقالی، یک مسئله مهم در یادگیری ماشین است که بر ذخیره‌سازی دانش حل یک مسئله، و استفاده از آن برای حل یک مسئله متفاوت اما مرتبط با آن تمرکز دارد. در عمل، تعداد کمی از افراد هستند که یک شبکه عصبی پیچشی را از ابتدا آموزش می‌دهند، زیرا مجموعه داده‌ها به اندازه کافی در دسترس نیستند. بنابراین، استفاده از وزن‌های شبکه از پیش آموزش دیده شده، برای راه‌اندازی یا استخراج ویژگی‌های ثابت، به حل بیشتر مسائل متنوع کمک می‌کند. شبکه‌های بسیار عمیق برای آموزش پرهزینه هستند. مدل‌های پیچیده‌تر، زمان زیادی را برای آموزش، با استفاده از صدها دستگاه با پردازنده‌های گران قیمت نیاز دارند.

تعریف توپولوژی، روش آموزش، تعیین و تنظیم پارامترها برای یادگیری عمیق می‌باشد که تئوری زیادی برای آن وجود ندارد. این شبکه‌ها با میلیون‌ها پارامتر و داده‌های کم، معمولاً با مشکل آموزش دقیق روبرو می‌شوند. در چنین شرایطی یادگیری انتقالی می‌تواند تاثیر بسزای در بهبود یادگیری داشته باشد.

¹ Support Vector Machine (SVM)

² Logistic Regression (LR)

³ Gaussian Bayes (GB)

⁴ Random Forest (RF)

شبکه‌های عصبی پیچشی سعی می‌کنند لبه‌ها را در لایه‌های قبلی، شکل‌ها را در لایه‌های میانی و برخی از ویژگی‌های خاص داده‌های سطح بالا را در لایه‌های بعدی تشخیص دهند. این شبکه‌های آموزش‌دیده، به‌طور کلی در حل مسائل گوناگون بینایی کامپیوتر، بسیار مفید هستند.

ویژگی‌های شبکه پیچشی، در لایه‌های اولیه، به‌صورت عمومی و کلی‌تر هستند و ویژگی‌های خاص مجموعه داده‌های ورودی، بیشتر در لایه‌های بعدی وجود دارند. بسته به بزرگی یا کوچکی مجموعه داده‌های ورودی و میزان شباهت آن‌ها با مجموعه داده‌هایی (اصلی) که شبکه پیچشی، از پیش آموزش‌دیده است، چهار استراتژی برای استفاده از یادگیری انتقالی وجود دارد [۷۲-۷۴].

۱- مجموعه داده‌های جدید، کوچک و شبیه به مجموعه داده‌های اصلی است: اگر سعی شود که کل شبکه را آموزش داد، مشکل بیش برآزش رخ می‌دهد. از آنجاکه داده‌ها شبیه به داده‌های اصلی هستند، انتظار می‌رود که ویژگی‌های سطح بالا در شبکه پیچشی نیز به این مجموعه داده مربوط باشد. از این‌رو، بهترین ایده، ممکن است آموزش یک طبقه‌بند خطی در کدهای شبکه پیچشی باشد. بنابراین در این روش، آموزش تمام لایه‌ها را متوقف کرده و فقط طبقه بند نهایی را برای طبقه‌بندی خاصی که موردنظر است، آموزش داده می‌شود.

۲- مجموعه داده‌های جدید، بزرگ و شبیه به مجموعه داده‌های اصلی است: از آنجایی که داده‌های بیشتری وجود دارد، می‌توان اطمینان بیشتری داشته داشت که حتی اگر تمام شبکه، تحت آموزش قرار گیرند، باز هم دچار بیش برآزش نمی‌شود.

۳- مجموعه داده‌های جدید، کوچک است اما بسیار متفاوت از مجموعه داده‌های اصلی است: از آنجاکه مجموعه داده بسیار کوچک است، ممکن است هدف استخراج ویژگی‌های موجود در لایه قبل باشد و یک طبقه‌بند را در بالای آن ایجاد شود. در این راستا می‌توان، چندلایه با اتصال کامل را اضافه کرد و لایه‌های بالای شبکه عصبی را آموزش داد.

۴- مجموعه داده‌های جدید، بزرگ و بسیار متفاوت از مجموعه داده‌های اصلی است: از آنجاکه مجموعه داده بزرگ است، می‌توان شبکه خود را طراحی کرده یا از شبکه‌های موجود استفاده کنید و یا شبکه را با استفاده از مقادیر اولیه تصادفی، آموزش داد یا از وزن‌های شبکه از پیش آموزش‌دیده شده استفاده شود.

شکل ۳-۵ و شکل ۳-۶، نحوه استخراج برخی از ویژگی‌ها از لایه‌های مختلف شبکه‌عصبی پیچشی گوگل-نت را در تصاویر شبکه چشم نشان می‌دهد. تصویر ورودی، مربوط به یک بیمار مبتلا به بیماری انحطاط ماکولای مرتبط با سن و یک بیمار مبتلا به ماکولای دیابتی است که از تصاویر تومورگرافی انسجام نوری^۱ (OCT) شبکه چشم، با اندازه $3 \times 224 \times 224$ به دست آمده است. لایه‌های پیچش، به‌تنهایی و بسته به ساختار و پارامترهای تعیین‌شده، خروجی‌های زیادی تولید می‌کنند بنابراین، بهترین پاسخ‌ها در هر لایه را می‌توان به‌طور خودکار شناسایی کرد.

لایه اول، فیلتری را شناسایی می‌کند که بتواند منحنی‌ها را اندازه‌گیری کند که با توجه به شکل، برای شناسایی تغییرات در لایه اپیتلیوم رنگ‌دانه شبکه به همراه اطلاعات ظریف دیگر موفق بوده است. لایه ReLU، فقط پاسخ مثبت را حفظ می‌کند و لایه ادغام، با اختصاص حداکثر پاسخ، به تمام پیکسل‌های همسایه، موجب کاهش اندازه تصویر می‌شود. لایه $LRN1^2$ پاسخ نرمال شده را نشان می‌دهد. پاسخ‌های لایه Inception-7، Inception-8 و Inception-9 که ویژگی‌های استخراج‌شده دیگری را نشان می‌دهند، به صورت بصری قابل تفسیر نیستند اما با تصاویر مشابه به دست آمده از بیماری ادما ماکولای دیابتی متفاوت است.

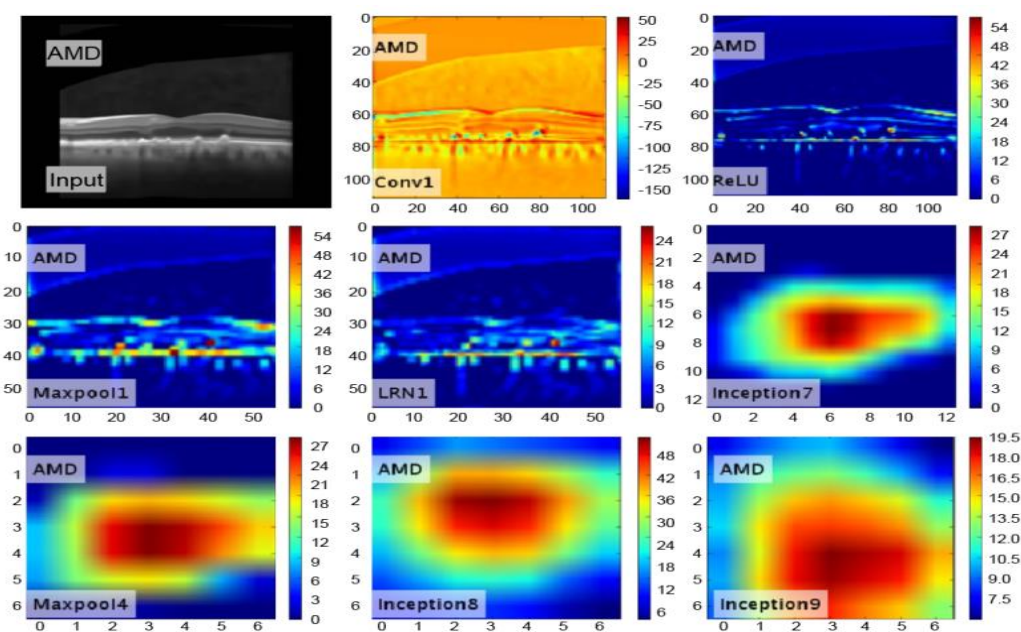
در مقایسه با موارد بیماری انحطاط ماکولای مرتبط با سن، پاسخ لایه اول، نمی‌تواند نشانه‌های برجسته‌ای برای شناسایی رسوبات مایع در شبکه را نشان دهد. با این حال، با حرکت به سمت لایه‌های نهایی، پاسخ‌های Inception-7، Inception-8 و Inception-9 الگویی را به نمایش می‌گذارد

¹ Optical Coherence Tomography (OCT)

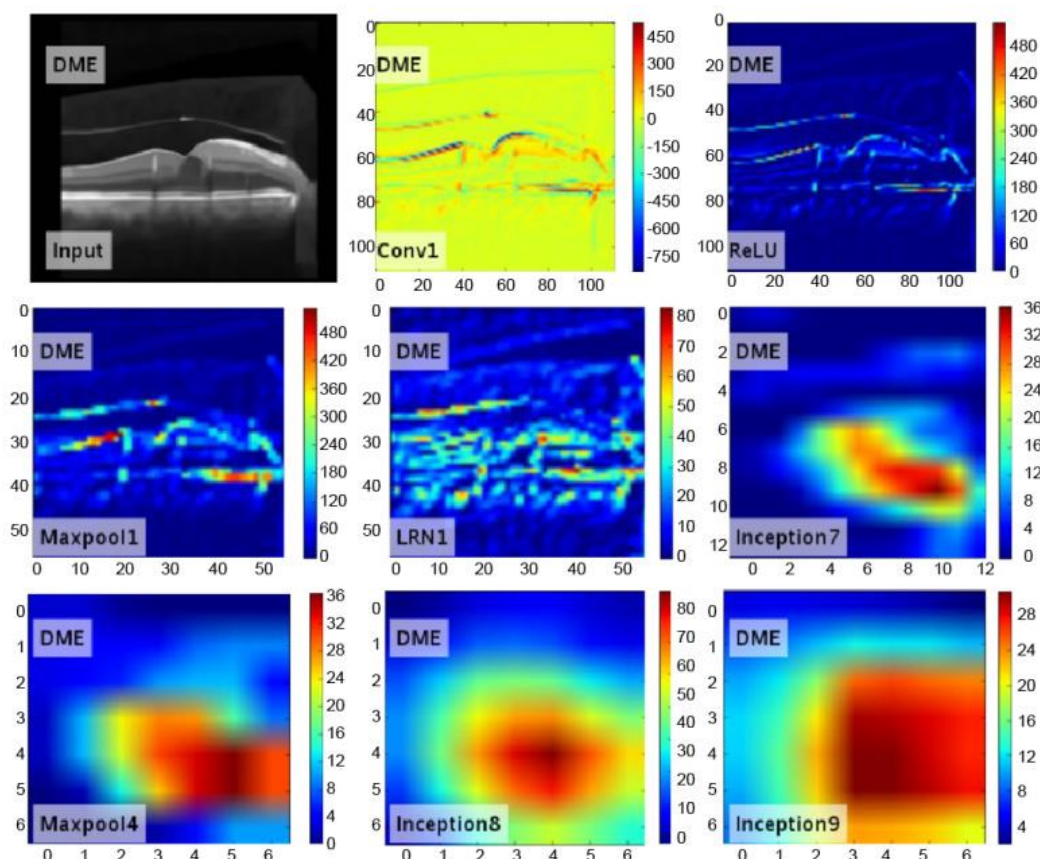
² Local Response Normalization

که نتیجه پاسخ‌های لایه‌های قبلی است و همان‌گونه که مشاهده می‌شود می‌تواند نشانه‌هایی برای تشخیص بیماری انحطاط ماکولای مرتبط با سن باشد.

یکی از دلایل شناسایی تغییر شکل لایه‌ها لایه‌های شبکه چشم به‌جای رسوبات مایع، توسط شبکه عصبی پیچشی می‌تواند این باشد که "تغییر شکل لایه (آتروفی جغرافیایی)"، یک عامل رایج در هر دو نوع بیماری است. در مورد بیماری انحطاط ماکولای مرتبط با سن، بلوردان در زیر لایه اپیتلیوم رنگ‌دانه شبکه رسوب می‌کند که بسیار بازتابنده است و لایه اول، به‌جای بلوردان، به لایه لایه‌های درونی کم است و از این‌رو، لایه بالای شبکه، به علت کنتراست بیشتر آن، پاسخ بهتری می‌دهد. به طور کلی می‌توان گفت که تمام پاسخ‌ها به واکنش‌های پویای مشترک تبدیل نمی‌شوند، زیرا تعداد کمی از پاسخ‌ها یک دامنه تغییرات کم دارند و تغییر آن، منجر به از دست دادن کنتراست بینایی می‌شود.



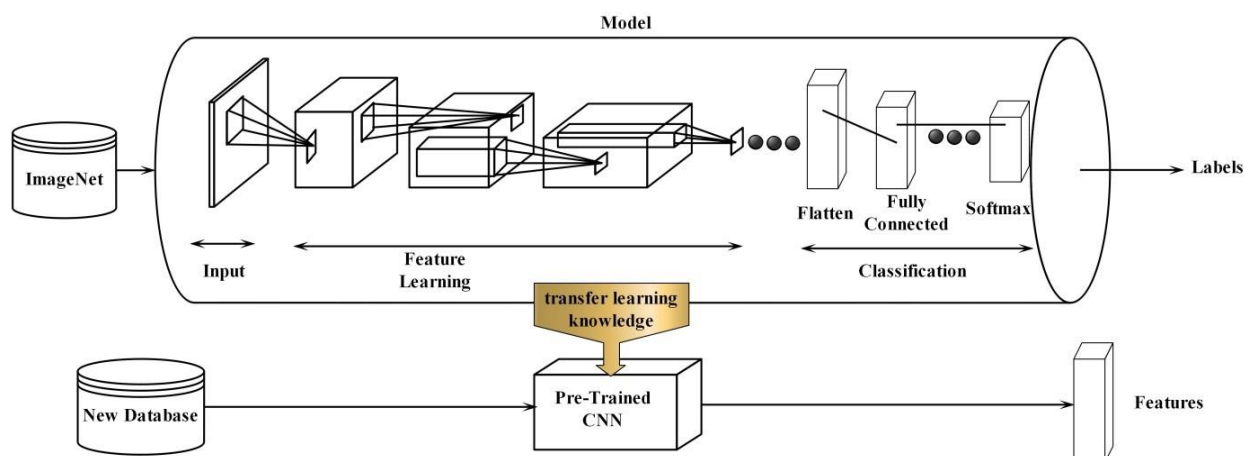
شکل ۳-۵ نمایش تصویر خروجی لایه‌های مختلف شبکه GoogLeNet با استفاده از یادگیری انتقالی و با تصویر ورودی OCT از شبکه چشم (نوع AMD)



شکل ۳-۶ نمایش تصویر خروجی لایه‌های مختلف شبکه GoogLeNet با استفاده از یادگیری انتقالی و با تصویر ورودی OCT از شبکه چشم (نوع DME)

همانطو که بیان شد، هدف اصلی از یادگیری انتقال، نگاشت از یک مدل از قبل آموزش دیده به یک مدل جدید در یک دامنه جدید است. یکی از مزایای اصلی یادگیری انتقال، صرفه جویی در وقت آموزش است. علاوه بر این، آموزش یک شبکه عصبی عمیق از ابتدا به مقدار زیادی داده نیاز دارد که معمولاً ارائه نمی‌شود. در این موارد، یادگیری انتقالی به میزان قابل توجهی نیاز به تعداد نمونه‌های آموزشی را کاهش می‌دهد. در این رساله، از آنجا که تعداد تصاویر سیستوسکوپی از مثانه محدود است، برای اصلاح مشکل یادگیری انتقالی اتخاذ می‌شود. در واقع از سومین استراتژی معرفی شده برای یادگیری انتقال استفاده شده است. یک مجموعه داده بزرگ به نام ImageNet برای بدست آوردن الگوهای اولیه انتخاب شده است. شکل ۳-۷ استفاده از یادگیری انتقالی را در روش پیشنهادی نشان می‌دهد. طراحی CNN عملکرد فاز استخراج ویژگی را بسیار تحت تأثیر قرار می‌دهد. برای دستیابی به بهترین طراحی معماری، یک مجموعه اعتبار سنجی برای ارزیابی معماری‌ها از جمله

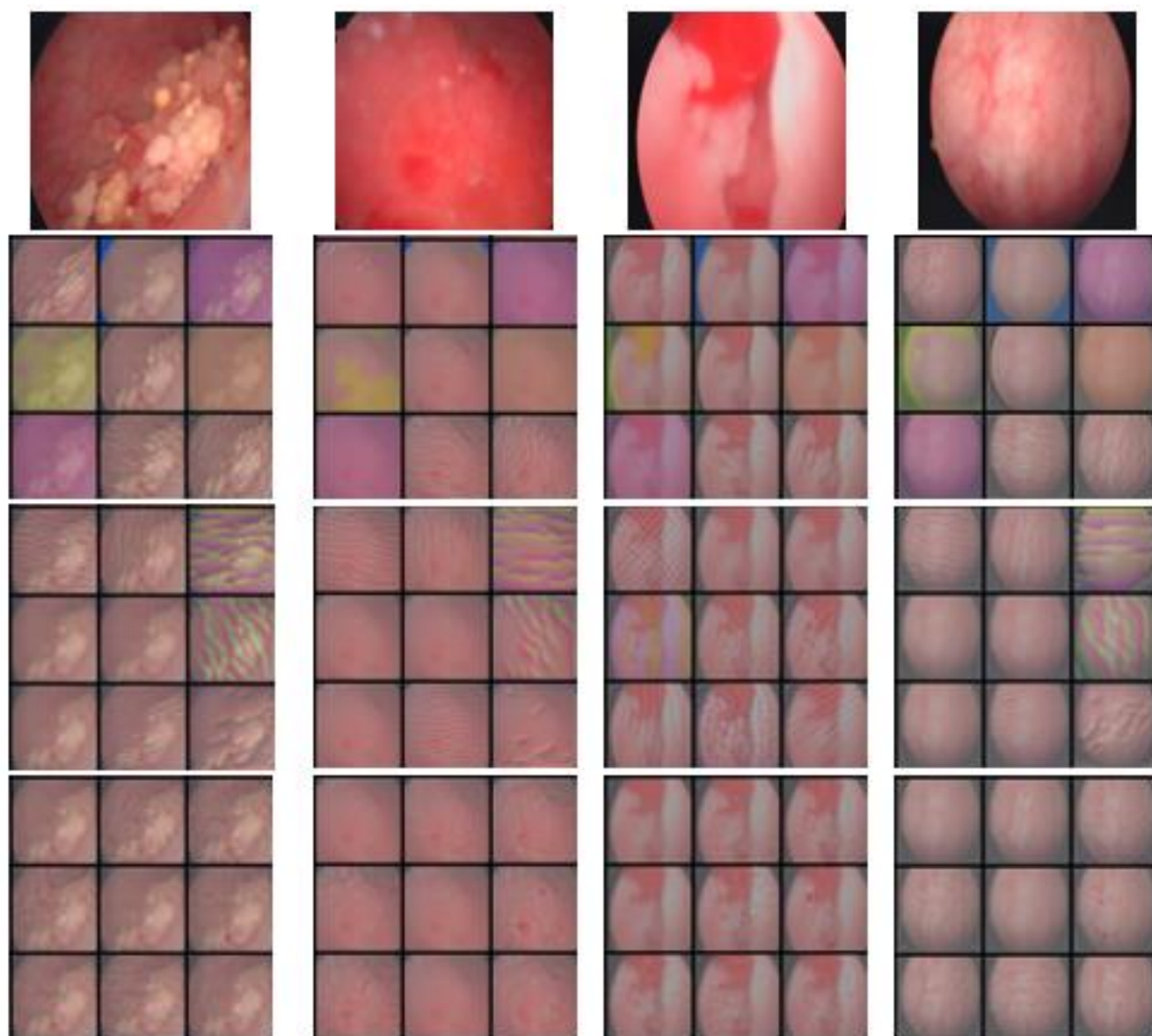
ResNet، AlexNet و VGG استفاده می‌شود. این معماری‌ها همان شبکه‌های از قبل آموزش دیده در مجموعه داده‌های ImageNet هستند. بر اساس نتایج تجربی، معماری VGG از سایر عملکردها بهتر است.



شکل ۳-۷ استفاده از یادگیری انتقالی در روش پیشنهادی

۳-۳-۲ نمایش ویژگی‌ها

قابل ذکر است که خروجی لایه‌های کانولوشن در فضای ویژگی است. از آنجا که از فضای ویژگی نمی‌توان برای نمایش استفاده کرد، از شبکه deconvolution برای تجسم ویژگی‌ها استفاده می‌شود. شبکه فضای ویژگی را به فضای تصویر ترسیم می‌کند. نگاشت ویژگی در سطوح مختلف تجزیه و تحلیل تجسم ویژگی‌های استخراج شده بسیار موثر است. به عنوان مثال، برای انتخاب بصری بهترین لایه، نگاشت‌های ویژگی معماری VGG-16 در شکل ۳-۸ نشان داده شده است. در این شکل، چندین هسته از چندین لایه نیز نشان داده شده است. همانطور که در تصاویر استخراجی مشخص می‌باشد هسته‌های مختلف نواحی مربوط به ضایعات و نواحی برجسته را در لایه‌های متفاوت کانولوشن استخراج می‌کند و نواحی هموار که اطلاعات خاصی را نشان نمی‌دهد دارای ارزش پایین‌تری بوده است.



شکل ۳-۸ نمایشی از خروجی لایه‌های مختلف CNN در تصاویر سیستم‌توسکوپی مثانه. به ترتیب خروجی لایه کانولوشن اول، سوم و پنجم در هر سطر نمایش داده شده است.

۳ - ۳ - ۳ کاهش ویژگی

مرحله استخراج ویژگی مشکلات مربوط به تعداد بالای ویژگی‌های استخراج شده را به همراه دارد. لایه‌های پیچشی به دلیل استفاده از هسته‌های مختلف دارای ابعاد بالایی هستند. برای حل مسئله، هنگامی که ماتریس ویژگی به صورت ترتیبی به بردار ویژگی نگاشت شود، با اعمال PCA و LDA، ویژگی‌های موثر تولید می‌شوند. نگاشت ویژگی جدیدی از ویژگی‌ها بر اساس واریانس داده‌ها در روش

PCA ارائه شده است. اجزای مجموعه داده با حداکثر تأثیر بر واریانس در مکانیزم نگاشت ویژگی نگهداری می‌شوند. تعداد ویژگی‌های موثر توسط رابطه ۱-۳ محاسبه می‌شود.

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{j=1}^f \lambda_j} \geq \text{ThresholdPCA} \quad 1-3$$

که در آن k تعداد ویژگی‌های انتخاب شده است، f تعداد همه ویژگی‌ها را نشان می‌دهد و λ مقدار ویژه^۱ را نشان می‌دهد. ویژگی‌های استخراج شده توسط PCA به عنوان ورودی‌های LDA تغذیه می‌شوند. LDA یک رویکرد آماری کلاسیک است که برای یافتن بهترین ترکیب خطی از ویژگی‌ها برای تشخیص دو یا چند کلاس استفاده می‌شود. LDA افتراق بین طبقات را مدل می‌کند در حالی که PCA افتراق بین آنها را نادیده می‌گیرد. LDA فاصله درون کلاس را به حداقل می‌رساند و فاصله بین کلاس را به حداکثر می‌رساند.

۳-۳-۴ طبقه‌بندی تصویر

پس از کاهش ابعاد و به دست آوردن بردار ویژگی نهایی، از طبقه‌بند ترکیبی برای طبقه‌بندی تصاویر سیستم‌سکوپی استفاده می‌شود. طبقه‌بند ترکیبی به روشی اطلاق می‌شود که در آن تعدادی طبقه‌بند پایه برای ساخت یک طبقه‌بند دقیق‌تر و مقاوم‌تر در مقایسه با هر طبقه‌بند پایه با هم ترکیب می‌شوند. رای‌گیری اکثریت یکی از رایج‌ترین تکنیک‌های ادغام است. همانطور که در شکل ۳-۴ نشان داده شده است، طبقه‌بندهای پایه پیشنهادی SVM، LR، BG و RF هستند. قدرت طبقه‌بند ترکیبی به تنوع^۲ طبقه‌بندهای پایه بستگی دارد. انتظار می‌رود نتایج خروجی طبقه‌بندها، داده‌های نمونه را در حداکثر دامنه ممکن پوشش دهد. استفاده از طبقه‌بندی با نتایج یکسان در طبقه‌بندی ترکیبی توصیه نمی‌شود. بنابراین، انتخاب یک طبقه‌بند مناسب همراه با یک مکانیزم ترکیبی استراتژیک، برای کاهش جمعی خطای طبقه‌بندی در یک طبقه‌بند ترکیبی مورد نیاز است. تنوع طبقه‌بندها می‌تواند به حداکثر

¹ eigenvalue

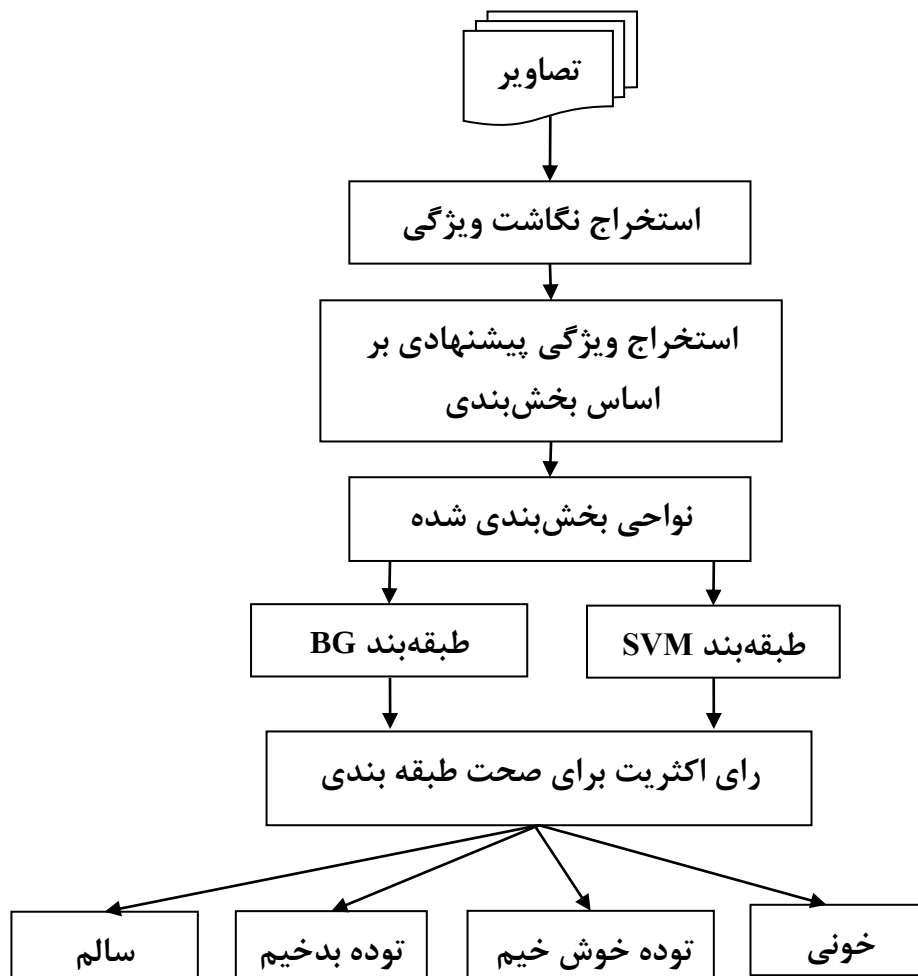
² diversity

پوشش داده‌ها و بهبود تعمیم طبقه‌بندی کمک کند. در SVM، بردارهای پشتیبان اجزای اصلی هستند که از طریق بهینه‌سازی محدب بدست می‌آیند. SVM مرز طبقه‌بندی را در حداکثر فاصله در کلاس‌ها مشخص می‌کند. رگرسیون لجستیک یک روش رگرسیون آماری برای متغیرهای وابسته است. مدل توسعه یافته یک توسعه خطی کلی برای تابع پیوند logit است و تابع خطا توزیع چند جمله‌ای را دنبال می‌کند. این روش ابتدا برای احتمال وقوع بیماری در کاربردهای پزشکی معرفی شد. طبقه‌بند بیزین الگوریتم منحصر به فردی ندارد اما خانواده‌ای از الگوریتم‌ها را با فرض استقلال بین ویژگی‌ها و متغیرها ارائه می‌دهد. بیشتر مدل‌های بیزین عملکرد اکتساب را به حداکثر می‌رسانند. یکی از مزایای طبقه‌بند بیزین برآورد پارامترهای مدل با مجموعه کوچکی از داده‌های آموزشی است. جنگل‌های تصادفی یک روش یادگیری گروهی برای طبقه‌بندی، رگرسیون و سایر کارهایی است که با ساختن تعداد زیادی درخت تصمیم در زمان آموزش کار می‌کند. جنگل تصادفی برای درختان تصمیم‌گیری که دارای اطلاعات آموزشی هستند مناسب است. در روش پیشنهادی، طبقه‌بندی‌ها با رای‌گیری بر اساس تکرارهای برچسب ترکیب می‌شوند. همانطور که در بالا ذکر شد، چهار طبقه‌بند مختلف از جمله SVM، LR، BG و RF به عنوان طبقه‌بند پایه انتخاب می‌شوند. دلیل اصلی این انتخاب "تنوع" است، یعنی مولفه اصلی در ساخت طبقه‌بندهای ترکیبی. همه این طبقه‌بندی‌ها به روشی کاملاً متفاوت آموزش دیده‌اند که باعث افزایش تنوع سطح طبقه‌بندی و همچنین تعمیم‌پذیری ترکیب می‌شود.

۳ - ۴ روش پیشنهادی بر اساس بخش‌بندی نواحی نگاشت ویژگی CNN

روش پیشنهادی در شکل ۳-۹ نشان داده شده است. همانطور که در شکل مشخص می‌باشد به مانند روش قبل از یادگیری انتقال برای استخراج نواحی مربوط به نگاشت‌های ویژگی استفاده شده است. آخرین لایه کانولوشن برای انتخاب نگاشت ویژگی استفاده می‌شود. در مرحله بعدی، ویژگی‌های سطح

بالا و نواحی مربوط به این ویژگی‌ها بر اساس روش‌های بهینه‌سازی ازدحام ذرات بهبود یافته^۱



شکل ۳-۹ روش پیشنهادی برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه بر پایه بخش‌بندی نواحی نگاشته‌های ویژگی CNN (IPSO) و اتسو^۲ استخراج می‌شوند. لازم به ذکر است که مقادیر نگاشت ویژگی در محدوده [۰,۱]

نرمال شده‌اند. سپس، با استفاده از روش PCA، ویژگی استخراج شده در بخش‌ها کاهش می‌یابد.

در آخر، طبقه‌بندی انجام می‌شود. در این روش، از یک روش ترکیبی مبتنی بر رای گیری اکثریت با BG و SVM استفاده شده است.

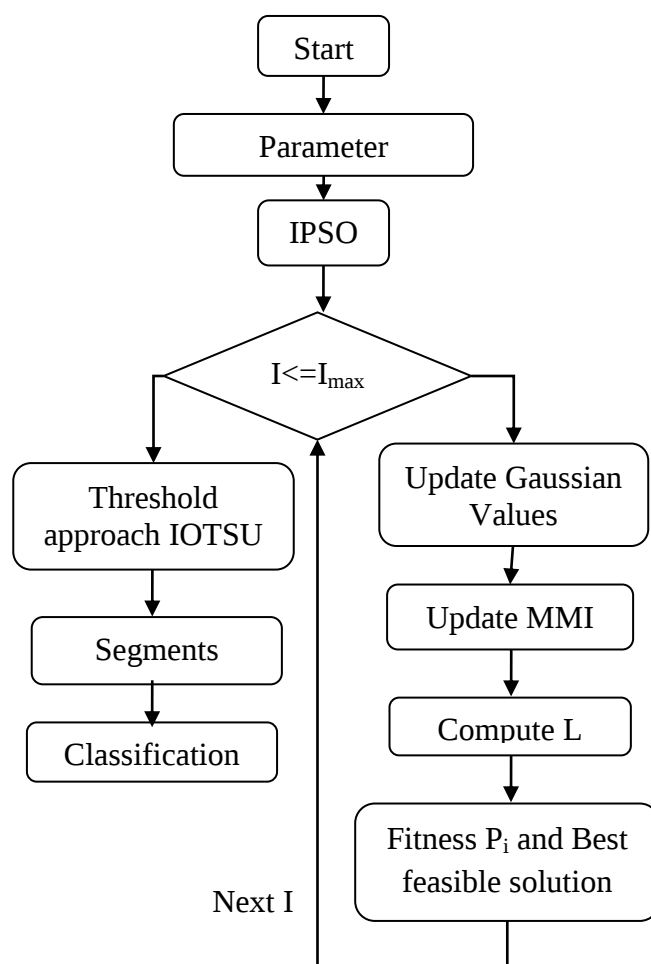
استخراج ویژگی یکی از مهمترین مراحل در تشخیص و درمان بیماری‌ها است. باید از ویژگی‌های قوی و قدرتمندی در این زمینه استفاده شود. در این بخش، به مانند روش قبل از یادگیری انتقال برای

^۱ Improved Particle Swarm Optimization (IPSO)

^۲ Otsu

استخراج ویژگی استفاده شده است. از ویژگی‌های استخراج شده توسط این روش برای مراحل بعدی که شامل بخش‌بندی^۱ است به عنوان ورودی استفاده می‌شود.

در این رساله نگاشت‌های ویژگی از لایه‌های پیچشی بدست می‌آیند. این نگاشت‌های شامل ویژگی‌های سطح بالا از تصاویر ورودی می‌باشند که از شبکه عصبی عمیق توسط یادگیری انتقال استخراج شده‌اند. برای استخراج ویژگی معنایی مناسب‌تر نگاشت‌های ویژگی بر اساس روش IPSO و اتسو بخش‌بندی می‌شوند. در واقع در ادامه برای کاهش ویژگی و دسته‌بندی تصاویر از ویژگی مربوط به نواحی انتخاب شده از نگاشت‌های ویژگی لایه‌های پیچشی استفاده شده است.



شکل ۳-۱۰ بخش‌بندی و طبقه‌بندی مبتنی بر انتخاب ویژگی توسط روش IPSO و اتسو

¹ segmentation

۳ - ۴ - ۱ بخش‌بندی نگاشت ویژگی

فرایند بخش‌بندی تصویر برای تبدیل تصویر به نواحی کوچکتر با ویژگی‌های همگن مشترک است که هر زیرناحیه باید دارای پیکسل‌های همگن باشد. بنابراین، برخی از رویکردهای بخش‌بندی استاندارد را می‌توان در انواع تصاویر سیستم‌سکوپی مثانه استفاده کرد. برای تجزیه و تحلیل و تفسیر تصاویر، از بخش‌بندی تصویر اصلی و استاندارد استفاده می‌شود. برای تقسیم تصویر ورودی به بخش‌های کوچکتر، از روش تقسیم‌بندی آستانه‌ای استفاده می‌شود. مقادیر یک یا چند برای تشخیص مرز در نظر گرفته می‌شود. از این رو، انتخاب مقدار آستانه مناسب با استفاده از قله‌های هیستوگرام کاملاً پیچیده است. اگر تعداد آستانه‌ها افزایش یابد، آستانه‌گذاری پیچیده‌تر می‌شود. بنابراین، بیشترین تحقیق در زمینه پردازش تصویر پزشکی، آستانه چند سطحی است.

با مقادیر آستانه چندگانه، می‌توان یک تصویر را در صورت آستانه چندتایی بخش‌بندی کرد. در فرآیند تجزیه و تحلیل تصویر، رویکردهای آستانه چند سطحی نقش مهمی دارند. تکنیک اتسو مقادیر آستانه بهینه را نشان می‌دهد که حداکثر واریانس را بین خوشه‌های مختلف نشان می‌دهد. در اینجا، با افزایش تعداد آستانه‌ها، زمان محاسبه به سرعت افزایش می‌یابد. بنابراین، برای تولید راه حل بهینه در یک زمان خاص، یک روش بهینه‌سازی لازم است. در مدل استفاده شده همانطور که در شکل ۳-۱۰ نشان داده شده است، با استفاده از روش بهینه‌سازی ازدحام ذرات بهبود یافته، معیاری برای انتخاب نواحی مربوط به ویژگی مناسب معرفی شده است. در واقع در این روش از IPSO برای انتخاب مقادیری که اتسو بهترین عملکرد را برای بخش‌بندی دارد استفاده می‌شود. در ابتدا، مقادیر نگاشت ویژگی با استفاده از مدل IPSO فیلتر می‌شوند تا ویژگی‌های ضروری شناسایی شوند. IPSO پیشنهادی تا زمانی که همه ویژگی‌های نگاشت ویژگی برای فرایند بخش‌بندی بر اساس آستانه استخراج نشود، تکرار می‌شود.

در IPSO، برای استخراج ویژگی‌های فیلتر شده در تصاویر سیستم‌سکوپی مثانه، یک تابع آشفته^۱ تصادفی گوسی و معیار متقابل اصلاح شده، بهینه می‌شود. "L" معیار شباهتی برای اندازه‌گیری شباهت بین هر ذره به بهترین ذره محلی را نشان می‌دهد. بهترین ذرات محلی و جهانی برای یافتن ویژگی‌های عملی در تصویر پس از محاسبه اندازه‌گیری اطلاعات متقابل اصلاح شده^۲ MMI و L استفاده می‌شود. در IPSO، I و Imax به ترتیب تکرار جریان و حداکثر تکرار هستند. ویژگی‌های اساسی برای یافتن نواحی تقسیم شده با استفاده از روش بهبود یافته اتسو و پس از اتمام رویکرد IPSO مشخص شده است.

در پایان، برای یافتن بهینه‌ترین نواحی در نواحی بیش از حد تقسیم شده، این نواحی با استفاده از مدل SVM غیرخطی طبقه‌بندی می‌شوند. با استفاده از روش IPSO در الگوریتم ۲-۳، تصاویر سیستم‌سکوپی مثانه برای استخراج ویژگی‌های پنهان پردازش می‌شوند. در این روش، با استفاده از اطلاعات متقابل و معیارهای آشفته گوسی برای روند استخراج ویژگی، بهینه‌سازی ازدحام ذرات ترکیبی بهبود می‌یابد. در مدل‌های بهینه‌سازی ازدحام محلی، مقدار آستانه به موقعیت پیکسل بستگی دارد. علاوه بر این، رویکردهای آستانه گذاری به دو رویکرد پارامتری و غیرپارامتری تقسیم می‌شوند. در رویکردهای پارامتریک، باید یک توزیع آماری از پیش تعریف شده برای هر کلاس وجود داشته باشد، در حالی که برای رویکردهای غیرپارامتری، توزیعی در نظر گرفته نمی‌شود. بهینه‌سازی یک معیار واحد یکی از شاخص‌های اصلی رویکردهای غیرپارامتری است.

در روش IPSO، در ابتدا گروه‌ای از ذرات به صورت تصادفی به وجود می‌آیند و برای یافتن راه‌حل بهینه از به‌روز رسانی نسل‌ها استفاده می‌شود. در هر گام، هر ذره با استفاده از بهترین مقدار به روز می‌شود. این مقدار مربوط به بهترین موقعیتی است که تاکنون ذره موفق به رسیدن به آن شده‌است.

¹ chaotic

² Modified Mutual Information

موقعیت مذکور شناخته و نگهداری می‌شود که این بهترین مقدار نوستالژی آن ذره نیز گفته می‌شود که آن را با P_{best} نمایش داده می‌شود.

روش‌های آستانه گذاری را می‌توان با توجه به نوع اطلاعات مورد استفاده طبقه بندی کرد. ویژگی‌ها و محتویات تصویر از جمله عواملی هستند که نتایج یک رویکرد آستانه خاص به آنها بستگی دارد. رویکردهای مختلف وقتی بر روی یک تصویر خاص در تصاویر دنیای واقعی پیاده سازی می‌شوند، نتایج متفاوتی ایجاد می‌کنند. رویکرد قابل توجهی برای نشان دادن عملکرد مناسب در انواع تصاویر وجود ندارد.

الگوریتم ۲-۳ الگوریتم پیشنهادی برای استخراج نواحی بهینه برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه

1. Read Feature Maps of Convolutional layers as $I(x, y)$
2. Randomly initialize the search agents $P_i (i = 1, 2, \dots, n)$ in the search space S .
3. $K=1$
4. Calculate the optimized fitness of each agent to find the best feasible practical solution P_{best} .

5. For I in $[1, 2, \dots, I_{max}]$ do

a. Initialize Gaussian chaotic values using

$$GuassiaChaotic(GC_i) =$$

$$\max \left\{ \mu \cdot \varphi_j^k (1 - \varphi_j^k), \frac{1}{\sqrt{2\pi\sigma_{I(x,y)}}} e^{-\frac{(1 - \mu_{I(x,y)})^2}{\sigma_{I(x,y)}^2}} \right\}$$

$$K = 1, 2, \dots, n$$

$$\varphi_j^k \in (0, 1)$$

b. Compute distance between φ_j^k and P_{best} as $\vec{L} = \left| \vec{H}_i \otimes \vec{P}_{best} - \vec{P}_i \right|$; if $Sim(P_{best}, P_i) < 0.5$

$$\vec{L} = \left| \vec{H}_i - MMI \right|; \text{ if } Sim(P_{best}, P_i) \geq 0.5$$

$$ModifiedMutualInformation(MMI) = -P_i \log \left(\sqrt[3]{P_i} \right)$$

$$H_i = 2^* \left(\vec{C}_i - \vec{C}_j \right)$$

$$\text{Where } \vec{C}_i, \vec{C}_j \in GC_k$$

$$\vec{P}_{I+1} = \vec{L} \otimes \exp(s.I) \otimes \cosine(2.\pi.s)$$

$$+ \lambda . \vec{P}_{best}$$

s is shape parameter, I is random number taken from GC.

End For

6. Compute the fitness value of the agent to find P_{best}

۳-۴-۲ مدل تقسیم تصویر چندسطحی با استفاده از روش اتسو بهینه شده

نگاشت ویژگی با $I(x, y)$ با سطوح $0, 1, \dots, \rho-1$ نشان داده می‌شود. هیستوگرام ورودی I به صورت $F = \{f_0, f_1, \dots, f_{\rho-1}\}$ نشان داده می‌شود. که در آن $f_0, f_1, \dots, f_{\rho-1}$ نشان دهنده تعداد تکرار هر سطح در ورودی است. $I_{\max}$ ، μ_I و σ_I به ترتیب نشان دهنده مقدار ویژگی با بیشترین سطوح، میانگین و واریانس مقادیر نگاشت ویژگی می‌باشند. مجموع تعداد پیکسل‌ها در نگاشت ویژگی با

$$N = \sum_{i=1}^{\rho} f_i \quad \text{محاسبه می‌شود. مقدار سطح احتمال تأمین به صورت}$$

$$\text{Pr ob}_i = \frac{f_i}{N}; \quad \text{Pr ob}_i > 0, \quad \sum \text{Pr ob}_i = 1 \quad \text{محاسبه می‌شود. روش مورد استفاده انتشار روش}$$

اتسو است. این روش ورودی را به $T+1$ ناحیه به صورت $S_k = \{S_0, S_1, \dots, S_k\}$ تقسیم می‌کند. T از مجموعه آستانه‌های $T = \{t[r], t[r+1], \dots, t[r+1]-1\}$ انتخاب می‌شود. مجموعه مقادیر گرفته شده از T نتیجه هر بخش تقسیم شده است. برای هر بخش S احتمال وزنی و مقادیر میانگین وزنی η به

صورت رابطه ۳-۲ و ۳-۳ محاسبه می‌شود.

$$WS_k = \sum_{i \in S_k} \text{Pr Ob}_i \quad 2-3$$

$$\eta = \sum_{i \in S_k} \frac{\text{Pr Ob}_i \cdot \max\{i, g * \log(i)\}}{WS_k}; \quad g \in GC \quad 3-3$$

مقادیر میانگین وزنی μ_{WI} و میانگین واریانس بین کلاسی σ_{VS}^2 بین نواحی تمام نگاشت ویژگی توسط

رابطه‌های ۳-۴، ۳-۵ و ۳-۶ محاسبه می‌شود.

$$\mu_{WI} = \sum_{i=0}^{\rho} i * \text{Pr Ob}_i \quad 4-3$$

$$\sigma_{VS}^2 = \sum_{i=0}^{\rho} WS_k \cdot \eta^2 - \mu_{WI}^2 \quad 5-3$$

$$(\sigma_{VS}^2)^* = \sum \max\{\sigma_{VS}^2(S_i), \sigma_{VS}^2(S_j)\} \quad 6-3$$

در این مدل، با به حداکثر رساندن واریانس بین کلاس‌ها، آستانه‌های بهینه برای هر سطح از ناحیه تقسیم شده به صورت رابطه ۷-۳ تعریف می‌شوند.

$$\{t^*[1], t^*[2], \dots, t^*[n]\} = \arg \max \{\sigma_{vs}^2(t[1], t[2], \dots, t[T])\} \quad 7-3$$

در این الگوریتم، برای انجام بخش‌بندی روی نگاشت ویژگی، از روش آستانه ورودی خودکار استفاده شده است. همچنین برای تخمین سطح نویز در فرآیند بخش‌بندی، از پردازش تغییر کلاس بین اشیاء پیش‌زمینه و پس‌زمینه استفاده می‌شود.

در مرحله بعدی، از روش PCA برای کاهش ویژگی تعیین شده توسط نواحی انتخاب شده استفاده می‌شود و هر ناحیه با روش‌های بیزین و SVM طبقه‌بندی می‌شود. در ادامه، با استفاده از روش رأی دادن، برچسب انتخاب می‌شود.

۳-۴-۳ کاهش ویژگی بر اساس نواحی بخش‌بندی شده

مشکل تعداد زیاد ویژگی‌ها، به دلیل استخراج ویژگی از شبکه‌های پیچشی بوجود می‌آید. معمولاً ابعاد بالای لایه‌های پیچشی به دلیل استفاده از هسته‌های مختلف است. در اینجا می‌توان از روش کاهش ویژگی استفاده کرد. در حالت کلی، به دلیل پراکندگی و توزیع ویژگی‌ها در بین کلاس‌ها انتخاب ویژگی موثر و مناسبتر می‌تواند در بهبود عملکرد سیستم‌های طبقه‌بند بسیار موثر باشد. در این رساله برای انتخاب بهترین ویژگی، از روش تحلیل مولفه اصلی استفاده شد. در واقع در این روش، یک نقشه برداری از ویژگی‌های جدید بر اساس واریانس داده معرفی می‌شود و اجزای مجموعه داده را که بیشترین تأثیر را بر واریانس دارند حفظ می‌کند. در این رساله، از روش PCA برای انتخاب ویژگی موثر بر اساس مقادیر ویژه در نواحی مختلف مرتبط به نگاشت ویژگی استفاده شد. در واقع، قبل از اینکه هر ناحیه با روش طبقه‌بندی برچسب گذاری شود، PCA برای نواحی استخراج شده برای هر ناحیه اعمال می‌شود.

در نهایت لازم به ذکر است به ازای نواحی استخراج شده از طبقه‌بندی ترکیبی بر پایه SVM و BG استفاده می‌شود.

۳ - ۵ جمع‌بندی

در این فصل روش‌های پیشنهادی رساله با جزئیات بیان شد. در ابتدا روش پیشنهادی بر پایه ویژگی سطح پایین LBP و ترکیب آن با روش‌های کاهش ویژگی و طبقه‌بند MLP بهبود یافته با الگوریتم ژنتیک و نرخ یادگیری تطبیقی معرفی شد. با توجه به پیچیدگی بالای تصاویر سیستم‌سکوپی این نوع ویژگی‌ها به تنهایی نمی‌تواند در این کاربرد استفاده شوند. در همین راستا در ادامه به معرفی ویژگی‌های سطح بالا بر پایه یادگیری انتقال و معمارهای مختلف CNN پرداخته شده است. با توجه به تعداد محدود تصاویر سیستم‌سکوپی مثانه یادگیری انتقال عملکرد مناسبی در روش پیشنهادی دیگر این رساله از خود نشان داده است. در ادامه به تحلیل نواحی کلیدی در نگاشت‌های ویژگی استخراج شده از لایه‌های متفاوت شبکه CNN پرداخته شده است. نتایج نشان دهنده این نکته می‌باشد که استفاده از ویژگی‌های سطح بالا در نواحی کلیدی سبب بهبود عملکرد روش پیشنهادی در این کاربرد می‌شود.

۴ پیاده‌سازی و تحلیل نتایج آزمایش‌ها

۴ - ۱ مقدمه

در این فصل مراحل پیاده‌سازی و تحلیل نتایج روش‌های پیشنهادی ارائه می‌شود. همان‌طور که در فصل قبل اشاره شد، دو روش کلی بر پایه ویژگی‌های سطح پایین و سطح بالا در این رساله پیشنهاد شده است. در همین راستا، ابتدا مجموعه داده تصاویر سیستم‌سکوپی مثانه به همراه معیارهای ارزیابی و همچنین سخت افزار و شرایط آزمایشات انجام شده بیان می‌شود. در گام بعدی روش پیشنهادی برای استخراج ویژگی‌های سطح پایین به همراه تحلیل نتایج ارائه می‌شود. در این مرحله نقاط قوت و ضعف روش بیان می‌شود. در گام بعدی روش پیشنهادی بر پایه استخراج ویژگی سطح بالا بر اساس یادگیری عمیق و بخش‌بندی نواحی در نگاشت ویژگی لایه‌های عمیق پیچشی به تفصیل بیان می‌شود. و پیاده‌سازی و تحلیل نتایج ارائه می‌شود.

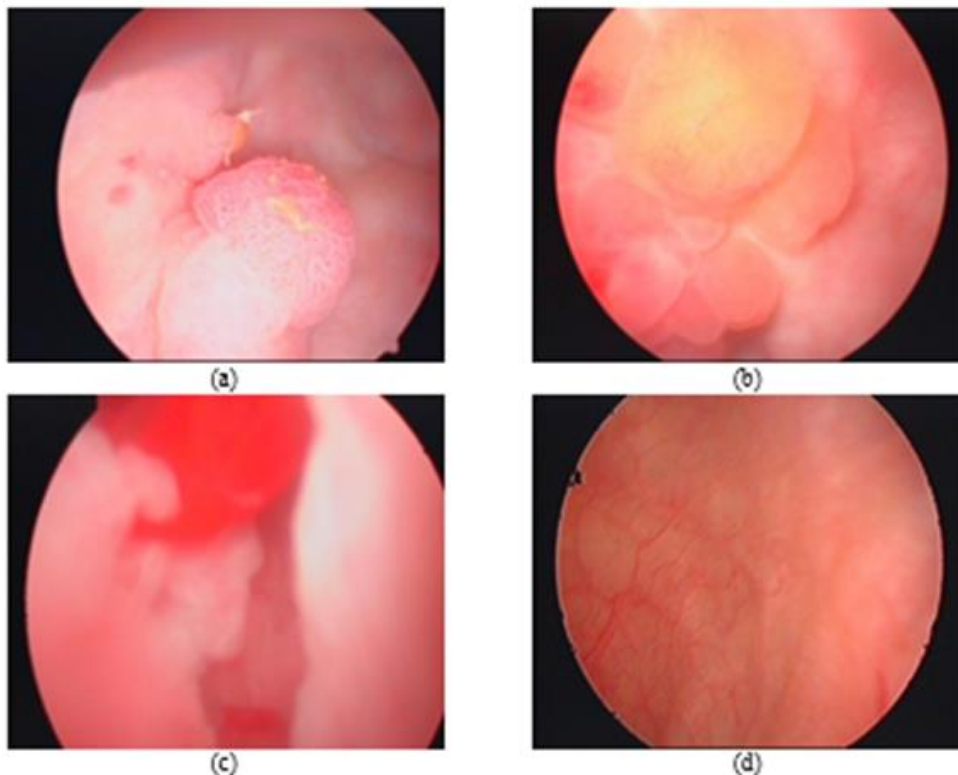
۴ - ۲ تنظیمات کلی پیاده‌سازی

در این بخش به معرفی مجموعه داده و معیارهای ارزیابی استفاده شده در این رساله پرداخته می‌شود. لازم به ذکر تمامی روش‌های ارائه شده و تحلیل نتایج بر روی مجموعه داده یکسان و کامپیوتری با پردازنده Core (TM) i7 M620، حافظه 4GB با کارت گرافیک T4 با زبان برنامه‌نویسی پایتون و فریم‌ورک Keras بوده است.

۴ - ۲ - ۱ پایگاه داده

پایگاه داده معرفی شده در این پژوهش از تصاویر پزشکی سیستم‌سکوپی در مرکز پزشکی در کشور هلند تهیه شده که دارای چهار کلاس به صورت کلاس‌های خونی، توده خوش‌خیم، توده بدخیم و

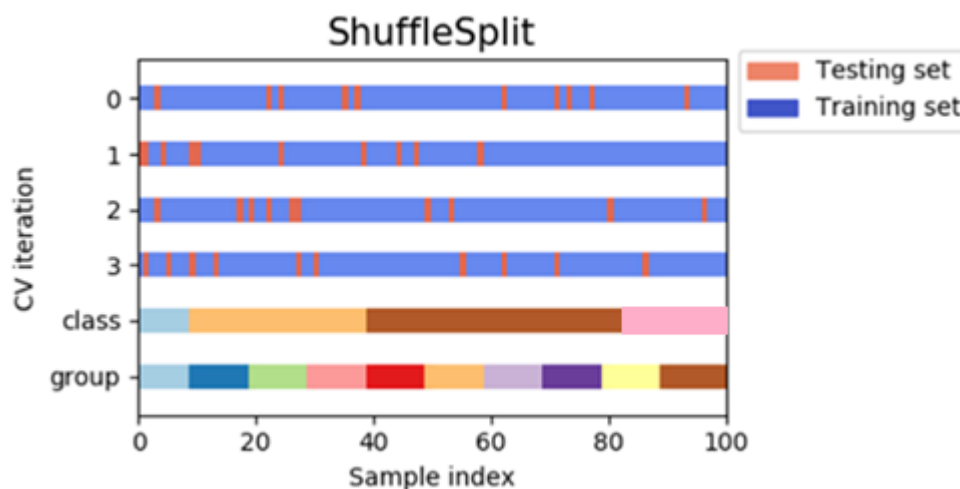
سالم می‌باشد. تصاویر پایگاه داده از زاویه‌های متفاوت می‌باشد. تصاویر از ناحیه‌های مختلف مثانه توسط سیستم سیستوسکوپی گرفته شده است. اندازه تصاویر و کیفیت آنها به علت شرایط مختلف متفاوت می‌باشد. اندازه تصاویر پایگاه داده اولیه به اندازه 540×476 تغییر داده شده است. در شکل ۱-۴ نمونه‌های از تصاویر این پایگاه داده را نشان دهد.



شکل ۱-۴ نمونه‌هایی از تصاویر پایگاه داده. (a) نمونه‌ای از تصویر توده بدخیم مثانه، (b) نمونه‌ای از تصویر توده خوش خیم مثانه، (c) نمونه‌ای از تصویر خونی مثانه (d) نمونه‌ای از تصاویر سالم

تعداد تصاویر برای هر کلاس ۱۸۰ تصویر می‌باشد. همه آزمایشات با استفاده از اعتبارسنجی متقابل^۱ k-fold که k برابر ۱۰ در نظر گرفته شده، انجام شده است. شکل ۲-۴ مراحل ساخت مجموعه آموزش و آزمون را نشان می‌دهد. در این شکل گروه‌های مختلف برای ۱۰ بار اجرا با رنگ‌های متفاوت نشان داده شده است. برای هر تکرار، ۱۰٪ از مجموعه داده به عنوان مجموعه آزمون و بقیه برای مجموعه آموزش در نظر گرفته شد. برای ایجاد یک مجموعه اعتبارسنجی برای هر تکرار، ۱۰٪ از تصاویر آموزش به طور تصادفی به عنوان مجموعه اعتبارسنجی انتخاب شدند.

^۱ cross validation



شکل ۴-۲ مراحل انتخاب نمونه در مجموعه آموزش و آزمون توسط روش ShuffleSplit

۴ - ۲ - ۲ معیارهای ارزیابی

برای نشان دادن عملکرد روش‌های پیشنهادی از معیارهای کمی دقت^۱ (ACC)، صحت^۲ (PR)، یادآوری (حساسیت)^۳ (RE)، خاص بودن (Spe)^۴ و F1 استفاده می‌شود.

به طور کلی، معیار دقت (ACC) نشان می‌دهد که یک مدل چطور می‌تواند برچسب خروجی را با دقت پیش بینی کند. معیار دقت در رابطه ۴-۱ نشان داده شده است. لازم به ذکر است که از میانگین و واریانس در k تکرار برای محاسبه دقت برای k-fold اعتبارسنجی متقابل استفاده می‌شود. این معیار سطح آموزش مدل و کارایی مدل را بررسی می‌کند اما اطلاعات دقیق در مورد کارایی مدل ارائه نمی‌دهد.

$$Accuracy = \frac{TP + TN}{total\ example} \quad ۴-۱$$

معیار صحت در رابطه ۴-۲ نشان داده شده است. این معیار برای مقادیر با مثبت کاذب بالا مناسب است.

^۱ Accuracy
^۲ Precision
^۳ Recall (sensitivity)
^۴ Specificity

$$PR = \frac{TP}{TP + FP} \quad ۲-۴$$

معیار یادآوری (حساسیت) در رابطه ۳-۴ نشان داده شده است. این معیار برای مقادیر با منفی‌های کاذب بالا مناسب است.

$$RE = \frac{TP}{TP + FN} \quad ۳-۴$$

معیار خاص‌بودن در رابطه ۴-۴ نشان داده شده است.

$$Spe = \frac{TN}{TP + FN} \quad ۴-۴$$

معیار F1 در رابطه ۵-۴ نشان داده شده است. این معیار هم معیارهای صحت و هم یادآوری را در بر می‌گیرد. معیار F1 در بهترین و بدترین حالت به ترتیب به ۱ و ۰ نزدیک می‌شوند.

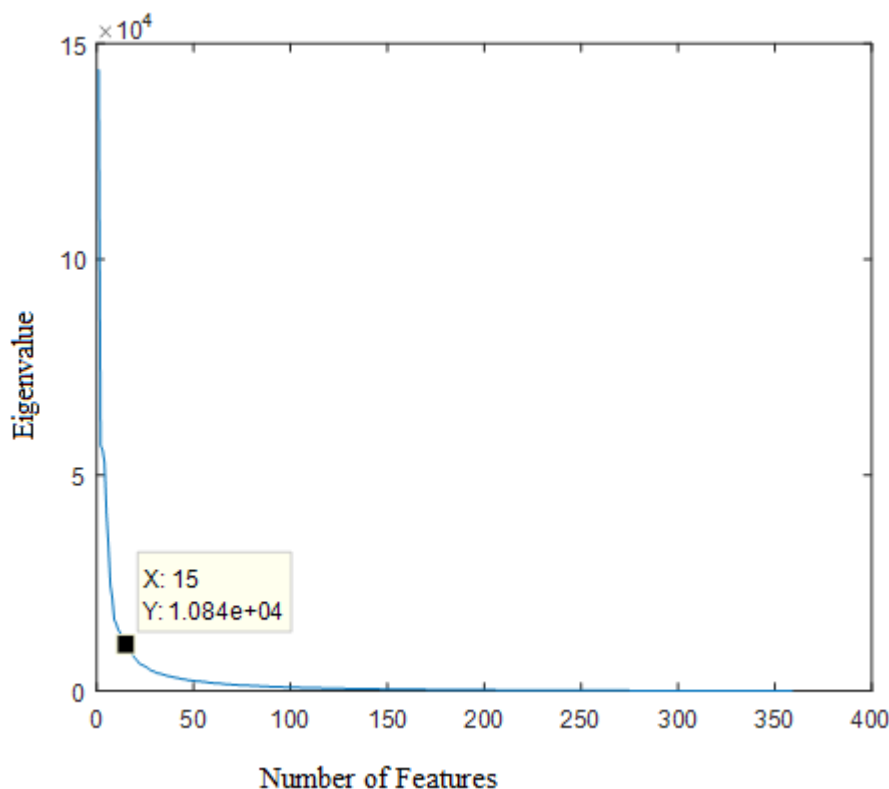
$$F1 = \frac{2 * RE * PR}{PR + RE} \quad ۵-۴$$

در روابط بالا منظور از TP تعداد تصاویری است که به درستی توسط طبقه‌بند به طبقه C_i انتساب یافته و FN یعنی تعداد تصاویر متعلق به سایر طبقه‌ها که به اشتباه توسط طبقه‌بند به طبقه C_i منسوب شده‌اند. FP تعداد تصاویر متعلق به طبقه C_i که توسط طبقه‌بند به سایر طبقه‌ها تعلق گرفته است. معیار TN تعداد تصاویری است که متعلق به طبقه C_i نبوده‌اند و توسط طبقه‌بند به این طبقه اختصاص داده نشده‌اند.

۴ - ۳ آزمایش‌ها و نتایج روش‌های پیشنهادی بر پایه ویژگی‌های سطح پایین

همانطور که در روش پیشنهادی بیان شد، روش ارائه شده در این بخش شامل سه مرحله مهم، استخراج ویژگی از تصاویر سیستم‌سکوپی و کاهش ویژگی، وزن‌دهی اولیه توسط الگوریتم ژنتیک و نرخ یادگیری تطبیقی می‌باشد. در این بخش ساختار استفاده شده برای این منظور معرفی می‌شود. در این روش از LBP برای استخراج ویژگی در تصاویر سیستم‌سکوپی ماثانه استفاده شده است. برای استخراج این ویژگی، تصاویر ورودی به سلول‌های 8×8 با ابعاد 32×32 تقسیم شده است در گام بعدی LBP

برای هر ناحیه محاسبه می‌شود. تعداد bin‌های خروجی ۵۹ در نظر گرفته شده است. در واقع بعد از اعمال این روش $۳۷۷۶ = ۵۹ * ۶۴$ ویژگی به ازای هر تصویر استخراج می‌شود. در گام بعدی با توجه به تعداد بالای ویژگی از روش PCA برای کاهش ویژگی استفاده شده است. جهت محاسبه تعداد ویژگی از نمودار مقادیر ویژه استفاده شده است. همانطور که در شکل ۳-۴ نشان داده شده است بعد از ۱۵ ویژگی اول با ارزش در PCA ارزش ویژگی‌های دیگر بسیار پایین بوده است در همین راستا فقط از این تعداد ویژگی برای مرحله بعدی استفاده می‌شود. در واقع نمونه‌های استخراجی با این تعداد ویژگی به فضای PCA نگاشت می‌شوند. برای انتخاب تعداد ویژگی بر اساس مقادیر ویژگی از رابطه ۱-۳ استفاده شده است.



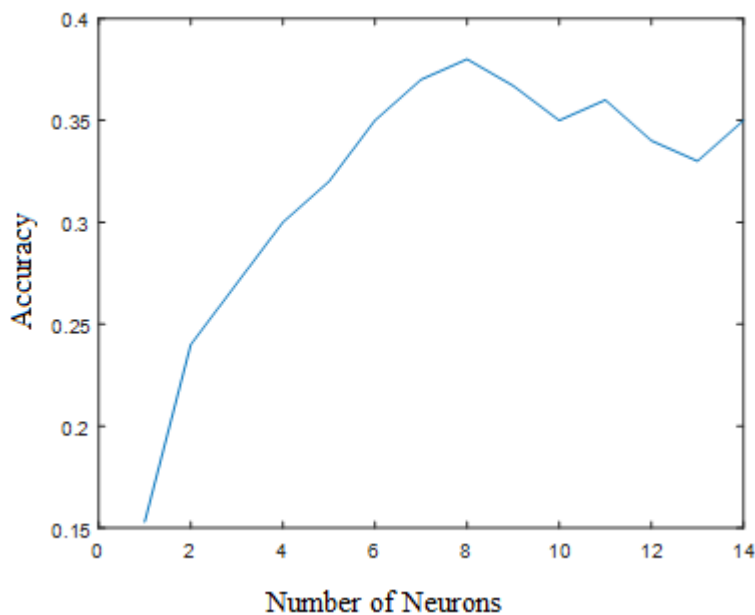
شکل ۳-۴ مقادیر ویژه استخراج بعد از اعمال PCA در تصاویر سیستم اسکوپ

مشخصات و پارامترهای مربوط به فرآیند یادگیری شبکه با استفاده از روش گرادینان نزولی در جدول ۱-۴ نشان داده شده است.

جدول ۱-۴ پارامترهای شبکه MLP برای تشخیص بیماری در تصاویر سیستوسکوپی مثانه

پارامترها	مقادیر
تعداد لایه‌های پنهان	۱
تعداد نورون در لایه ورودی	۱۵
تعداد نورون در لایه پنهان	۸
تعداد نورون در لایه خروجی	۴
حداکثر تعداد دفعات تکرار	۲۰۰
حد آستانه خطا برای توقف یادگیری	۰.۱
نحوه وزن دهی اولیه وزن‌ها	الگوریتم ژنتیک

لازم به ذکر است تعداد نورون‌های لایه پنهان با توجه به مجموعه اعتبارسنجی محاسبه شده است. در واقع در هر بار انتخاب مجموعه آموزش، ۱۰ درصد از داده‌ها برای مجموعه اعتبارسنجی استفاده شده است. در واقع شبکه MLP با تعداد نورون‌های متفاوت در لایه میانی آموزش و مورد ارزیابی قرار گرفته و بهترین عملکرد را شبکه با ۸ نورون در لایه میانی در مجموعه اعتبارسنجی از خود نشان داده است. شکل ۴-۴ میزان دقت شبکه را به ازای نورون‌های متفاوت نشان می‌دهد.



شکل ۴-۴ میزان دقت شبکه با تعداد نوروهای متفاوت در مجموعه اعتبارسنجی

پارامترهای الگوریتم ژنتیک برای محاسبه وزن‌های اولیه در جدول ۴-۲ نشان داده شده است. لازم به ذکر است تعداد تکرار بالا در آموزش توسط الگوریتم ژنتیک سبب بیش برآزش می‌شود به همین دلیل تعداد ۱۰ برای تکرار در نظر گرفته شده است.

جدول ۴-۲ پارامترهای پیکربندی الگوریتم ژنتیک برای محاسبه وزن‌های اولیه

پارامترها	مقادیر
تعداد تکرار	۱۰
تعداد جمعیت اولیه	۱۰۰ نسل
تابع انتخاب	@selectiontournament
تابع ترکیب	@crossoversinglepoint
نرخ جهش	۰.۰۱
تابع جهش	@mutationgaussian

همانطور که در قبلا اشاره شد، به دلیل کند بودن الگوریتم ژنتیک تعداد تکرار ۱۰ در نظر گرفته شده است. در ادامه بدست آوردن بازه تغییرات نرخ یادگیری با توجه به مجموعه آموزش بسیار مهم و

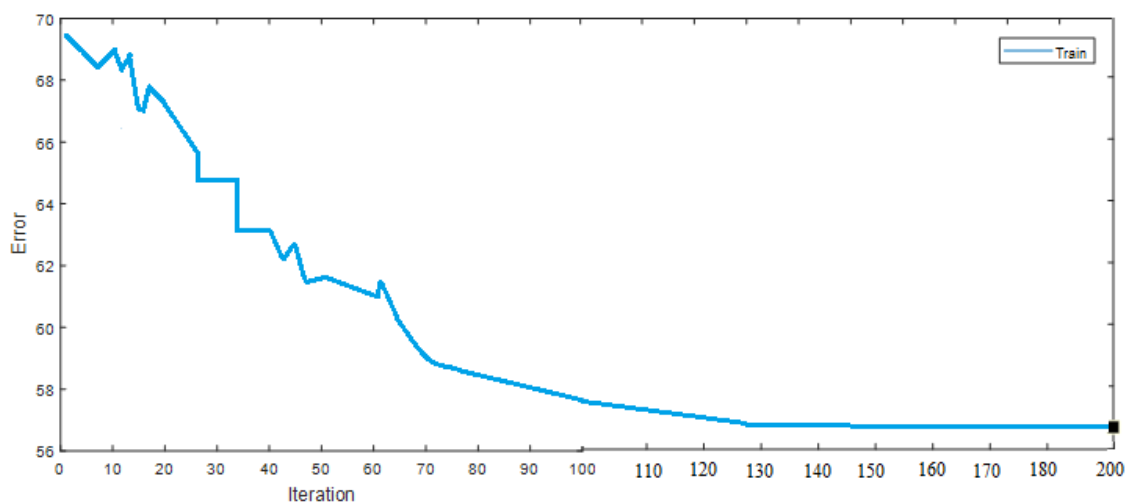
حیاتی می‌باشد. جدول ۳-۴ بازه و میزان تغییرات نرخ یادگیری با توجه به خطای مجموعه اعتبارسنجی را نشان می‌دهد.

جدول ۳-۴ میزان تغییر نرخ یادگیری با توجه به خطای اعتبارسنجی

نرخ یادگیری	بازه (خطای اعتبارسنجی)
۰.۰۱	< 0.5
۰.۰۳	$0.5 < \text{و} < 0.6$
۰.۰۶	$0.6 < \text{و} < 0.7$
۰.۱	Else

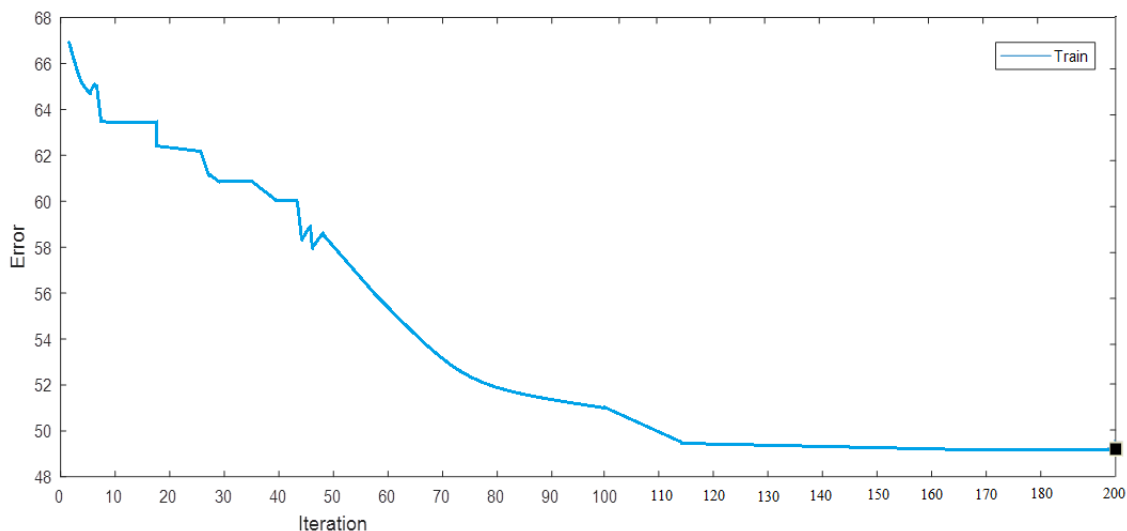
۴ - ۳ - ۱ مقایسه روش پیشنهادی بر پایه ویژگی سطح پایین با روش‌های دیگر

برای بررسی تاثیر تغییر نرخ یادگیری ابتدا نرخ یادگیر را ثابت فرض می‌کنیم. در اینجا میزان نرخ یادگیری ۰.۱ در نظر گرفته شده است. در ادامه نمودار شکل ۴-۵ خطای روش ثابت برای مجموعه آموزش نشان داده شده است. همانطور که در شکل نشان داده شده بهترین میزان خطا ۵۶.۱۶ در مرحله آموزش می‌باشد.



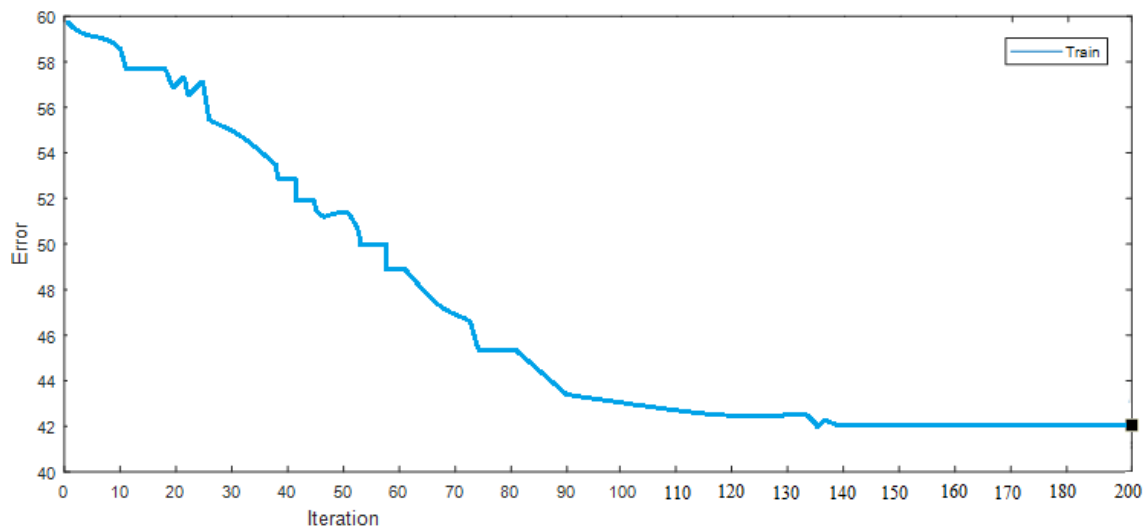
شکل ۴-۵ میزان خطای شبکه MLP با نرخ یادگیری ثابت

در ادامه تاثیر تغییر نرخ یادگیری بر روی مجموعه تصاویر سیستم اسکوپ مورد ارزیابی قرار می‌گیرد. شکل ۴-۶ میزان خطای شبکه عصبی با نرخ یادگیری متغیر و تطبیقی با میزان خطای اعتبارسنجی را نشان می‌دهد. همانطور که در شکل نشان داده شده روش پیشنهادی نسبت به روش ثابت عملکرد بهتری در مراحل آموزش داشته است. میزان خطای روش پیشنهادی در مرحله آموزش ۴۹.۲۵ می‌باشد.



شکل ۴-۶ میزان خطای شبکه MLP با نرخ یادگیری متغیر

یکی دیگر از پارامترهای تاثیرگذار در آموزش شبکه استفاده از یک روش قدرتمند در وزن‌دهی اولیه شبکه می‌باشد. همانطور که در قسمت روش پیشنهادی بیان شد از الگوریتم ژنتیک برای آموزش اولیه شبکه استفاده شده است. شکل ۴-۷ نتایج روش پیشنهادی با نرخ یادگیری متغیر و انتخاب وزن‌های اولیه با الگوریتم ژنتیک را نشان می‌دهد. همانطور که در شکل نشان داده شده این روش نسبت به روش‌های پیشین عملکرد بهتری در مجموعه آموزش از خود نشان داده است. میزان خطای روش پیشنهادی در مرحله آموزش ۴۱.۱۴ می‌باشد.



شکل ۴-۷ میزان خطای شبکه MLP با نرخ یادگیری متغیر و وزن‌دهی اولیه توسط الگوریتم ژنتیک

در ادامه روش پیشنهادی بر روی مجموعه آزمون مورد ارزیابی قرار می‌گیرد. جدول ۴-۴ نتایج حاصل از مقایسه روش پیشنهادی و روش‌های دیگر را نشان می‌دهد.

جدول ۴-۴ مقایسه روش‌ها برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه (میزان دقت)

روش MLP		
وزن‌دهی تصادفی و نرخ یادگیر ثابت	وزن‌دهی تصادفی نرخ یادگیری متغیر	وزن‌دهی با ژنتیک و نرخ یادگیری متغیر
۳۸.۵	۴۲.۱۱	۴۹.۱

همانطور که از نتایج جدول ۴-۴ مشخص می‌باشد روش پیشنهادی عملکرد مناسب‌تری نسبت به روش‌های دیگر در رده بندی تصاویر سیستم‌سکوپی مثانه از خود نشان داده است.

۴ - ۴ آزمایش‌ها و نتایج روش‌های پیشنهادی بر پایه ویژگی‌های سطح بالا و

یادگیری انتقال

در این بخش آزمایشات انجام شده بر اساس روش پیشنهادی بر اساس یادگیری انتقال و کاهش ویژگی و طبقه‌بندی بر اساس طبقه‌بند ترکیبی ارائه می‌شود. در ابتدا پارامترهای مورد نیاز برای هر بخش تشریح می‌شود. در گام بعدی نتایج بدست آمده به همراه تحلیل نتایج بیان می‌شود.

۴ - ۴ - ۱ تنظیمات مربوط به پارامترهای استفاده در روش پیشنهادی بر اساس ویژگی

سطح بالا

در این بخش مجموعه داده و سخت افزار مشابه که در بخش ۴ - ۲ - ۱ معرفی شده استفاده شده است. ویژگی‌های استفاده شده قبل و بعد از کاهش ویژگی در جدول ۴-۵ نشان داده شده است.

جدول ۴-۵ پارامترها و تعداد ویژگی‌های استخراج شده قبل و بعد از روش کاهش ویژگی

روش	پارامترهای		تعداد ویژگی	
			قبل	بعد
PCA	تعداد مولفه‌های اصلی	۱۰۰	۱۳۱۰۷۲	۱۰۰
	ThreasoldPCA	۰.۹۸		
LDA	تعداد مولفه‌ها	Min(# class-1, # of features)= Min(3,100)=3	۱۰۰	۳
	Solver	Singular Value decomposition (SVD)		

لازم به ذکر است که از کتابخانه Scikit-learn-0.22 برای طبقه‌بندی استفاده شده است و به طور پیش فرض از کلیه پارامترهای موجود در این کتابخانه استفاده شده است. به عنوان مثال، SVC (طبقه بند بردار پشتیبانی^۱) رویکرد "یک در مقابل یک"^۲ را برای طبقه بندی چند طبقه استفاده می‌شود. در مجموع شش طبقه‌بند ساخته شده است، که هر کدام داده‌های دو کلاس را آموزش داده‌اند. جدول ۴-۶ اصلی ترین پارامترهای طبقه‌بندهای پایه را نشان می‌دهد.

جدول ۴-۶ ابرپارامترهای طبقه‌بندهای پایه

مقادیر	پارامترها	روش
1.0	C	SVM
RBF	Kernel	
1.0	C	Logistic Regression
L2	Penalty	
Lbfgs	Solver	
100	Max_iter	
2	Min_samples_split	Random Forest
1	Min_samples_leaf	
None	Priors	Gussian Bayes
1e-09	Var-smoothing	

همانطور که در قبلا اشاره شد، تصاویر اصلی سیستم‌سکوپی با اندازه‌های متفاوت بوده است. با این حال، هنگامی که آنها به CNNهای عمیق از قبل آموزش دیده تغذیه می‌شدند، به طور خودکار با توجه به اندازه ورودی شبکه، اندازه آنها تغییر می‌کند. به عنوان مثال، برای معماری VGG، اندازه ورودی $224 \times 224 \times 3$ تغییر کرده است.

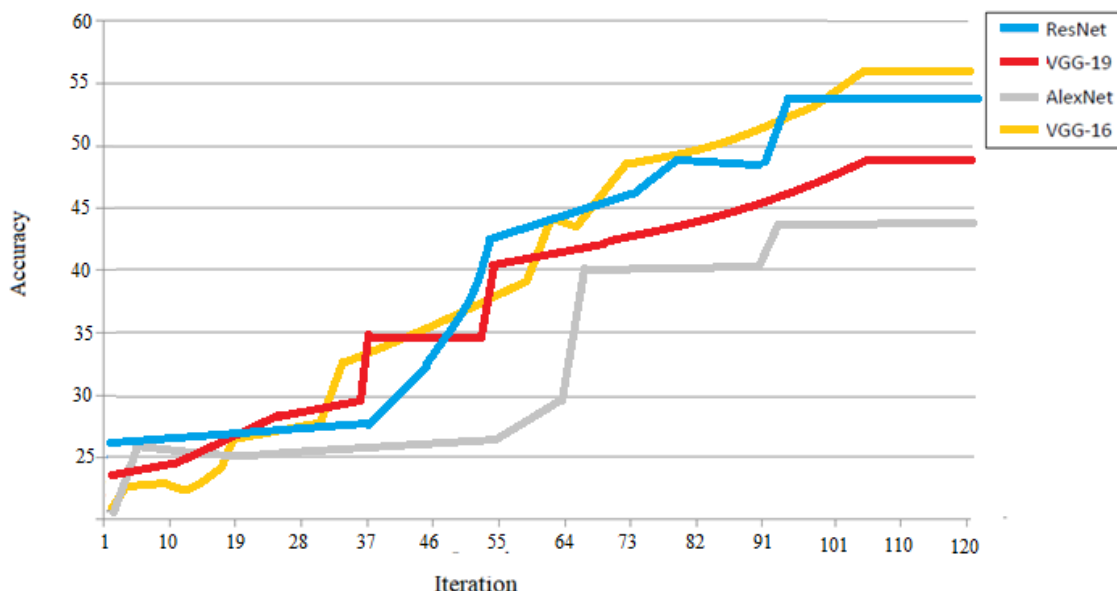
۴ - ۴ - ۲ انتخاب معماری

همانطور که در بخش روش پیشنهادی ذکر شد، انتخاب معماری عمیق در شبکه پیچشی بسیار بر عملکرد روش پیشنهادی تأثیر می‌گذارد. چهار معماری شامل VGG-16، VGG-19، AlexNet و ResNet مورد بررسی قرار گرفته اند. لازم به ذکر است که لایه‌های پیچشی آموزش داده نمی‌شوند و

¹ Support Vector Classifier (SVC)

² one-versus-one

فقط لایه‌های کاملاً متصل در این مرحله آموزش می‌بینند. در این مرحله، برای انتخاب بهترین معماری، لایه‌های کاملاً متصل به طور جداگانه برای هر نوع معماری آموزش داده می‌شوند. لایه خروجی در همه معماری‌ها شامل چهار نود است که برابر با تعداد کلاس‌ها است. در حقیقت از لایه‌های پیچشی برای استخراج ویژگی‌ها (بدون آموزش) و از لایه‌های کاملاً متصل (با آموزش) مشابه MLP جداگانه برای طبقه‌بندی تصاویر در این مرحله استفاده می‌شود. شکل ۴-۸ عملکرد هر نوع معماری را بر اعتبار سنجی تنظیم شده از نظر اندازه‌گیری دقت نشان می‌دهد. تعداد تکرارهای آموزش لایه‌های کاملاً متصل ۱۰۰ در نظر گرفته شده است. در حالیکه پارامترهای باقیمانده روی مقادیر پارامتر پیش فرض تنظیم شده‌اند. شکل ۴-۸ برتری عملکرد معماری VGG-16 را در مورد اعتبار سنجی در مقایسه با سایر معماری‌ها نشان می‌دهد. بنابراین، VGG-16 به عنوان آخرین معماری برای استخراج ویژگی‌ها استفاده می‌شود. بر اساس نتایج، یک شبکه عصبی کاملاً متصل (به عنوان مثال، یک MLP) تصاویر سیستم‌سکوپی را با دقت ۵۵.۱۳٪ طبقه‌بندی می‌کند. روش پیشنهادی می‌تواند عملکرد و میزان تشخیص را بهبود بخشد.



شکل ۴-۸ نتایج اولیه بر اساس معماری‌های مختلف در طبقه‌بندی تصاویر سیستم‌سکوپی مثانه

۴ - ۴ - ۳ مقایسه روش پیشنهادی بر پایه ویژگی‌های استخراج شده از یادگیری انتقال

در این بخش، روش پیشنهادی با سایر روش‌ها برای طبقه‌بندی تصاویر سیستم‌سکوپی مقایسه می‌شود. ویژگی‌ها پس از انتخاب معماری استخراج شدند. در روش پیشنهادی، یادگیری انتقال بر روی معماری VGG-16 اعمال شد. در مرحله بعدی، تکنیک‌های PCA و LDA به ترتیب برای کاهش ویژگی استفاده شد. در اینجا برای انتخاب ویژگی از لایه کاملاً متصل استفاده شده است که تعداد آن در جدول ۴-۵ نشان داده شده است. در هر تکرار، تعداد ویژگی‌های انتخاب شده با توجه به داده‌های آموزش و رابطه ۳-۱ متفاوت بود. جدول ۴-۷ عملکرد تشخیص چهار کلاس سرطان مثانه را از نظر میانگین و واریانس توسط طبقه‌بندهای مختلف پایه را نشان می‌دهد. Gaussian Bayes بهترین عملکرد را در مقایسه با سایر روش‌ها ارائه می‌دهد.

جدول ۴-۷ مقایسه روش‌های پایه در طبقه‌بندی تصاویر سیستم‌سکوپی

روش	Mean Accuracy	Variance
Logistic Regression	۶۷.۶	۰.۱
Support Vector Machines	۶۷.۵	۰.۱۵
Decision Trees	۶۶.۲۵	۰.۲۴
Random Forest	۶۶.۲۵	۰.۲۴
Gaussian Bayes	۶۸.۷۵	۰.۱۱

طبقه‌بند ترکیبی پیشنهادی با طبقه‌بندهای ترکیبی Adaboost و Gradient boosting مقایسه شد (جدول ۴-۸). در روش پیشنهادی، وزن طبقه‌بند های پایه شامل رگرسیون لجستیک، جنگل تصادفی، Gaussian Bayes و ماشین بردار پشتیبانی به ترتیب ۳، ۲، ۱ و ۳ در نظر گرفته شده است. این وزن‌ها با توجه به عملکرد روش‌ها در جدول ۴-۷ انتخاب شده است.

جدول ۴-۸ مقایسه روش پیشنهادی با روش‌های ترکیبی

روش	Mean Accuracy	Variance
Adaboosting	۶۵.۴۱	۰.۲۸
GradientBoosting	۶۸.۷۵	۰.۱۱
روش پیشنهادی	۶۹.۰۲	۰.۱۹

طبقه‌بند ترکیبی پیشنهادی در مقایسه با روش‌های Adaboost و Gradient Boosting نتایج بهتری را ارائه داد. لازم به ذکر است که یادگیری انتقال و کاهش ویژگی دقت طبقه‌بندی را بهبود بخشید. جدول ۵ نتایج روش پیشنهادی در معماری VGG16 را بر اساس چهار معیار ارزیابی صحت (PR)، یادآوری (حساسیت) (RE)، خاص‌بودن (Spe) و F1 در هر کلاس به صورت جداگانه نشان می‌دهد.

جدول ۴-۹ مقایسه روش پیشنهادی بر اساس چهار معیار ارزیابی صحت (PR)، یادآوری (حساسیت) (RE)، خاص‌بودن (Spe) و F1 در هر کلاس به صورت جداگانه

Methods	Label	Precision	F1	Specificity	Recall(sensitivity)
Logistic Regression	Benign	43	50	73	59
	Blood pee	65	74	84	87
	Cancer	79	56	96	44
	Normal	93	93	97	93
SVM	Benign	50	51	82	53
	Blood pee	64	62	88	60
	Cancer	67	65	89	64
	Normal	84	86	94	89
Decision Trees	Benign	57	59	84	60
	Blood pee	56	62	81	69
	Cancer	62	57	89	53
	Normal	89	87	96	85
Random Forest	Benign	58	55	87	52
	Blood pee	59	65	83	72
	Cancer	63	60	88	57
	Normal	100	100	1	100
GB	Benign	62	54	90	48
	Blood pee	78	80	92	82
	Cancer	39	44	73	50
	Normal	90	90	96	90
Adaboosting	Benign	36	36	78	36
	Blood pee	58	52	88	47
	Cancer	54	63	78	75

	Normal	100	89	1	81
GradientBoosting	Benign	71	67	91	63
	Blood pee	68	70	88	71
	Cancer	60	67	83	75
	Normal	88	70	97	58
روش پیشنهادی	Benign	58	72	77	95
	Blood pee	81	70	95	62
	Cancer	71	59	93	50
	Normal	91	87	97	83

تمام آزمایشات بر روی معماری ResNet انجام شده است. نتایج ارزیابی معماری ResNet در جدول ۴-۱۰ نشان داده شده است. یادگیری انتقال تأثیر مثبتی بر عملکرد استخراج ویژگی دارد. روش پیشنهادی مبتنی بر معماری ResNet در مقایسه با سایر روش‌ها نتایج بهتری داشت. بر خلاف جدول ۴-۸ و جدول ۴-۹، دقت روش پیشنهادی مبتنی بر معماری ResNet کمتر از دقت روش پیشنهادی مبتنی بر معماری VGG-16 بود، که این می‌تواند به دلیل تعداد زیاد لایه‌ها و تعداد کم نمونه‌های آموزشی باشد.

جدول ۴-۱۰ مقایسه روش‌های مختلف بر اساس معماری ResNet برای استخراج ویژگی

روش	Mean Accuracy	Variance
Logistic Regression	۶۳.۰۵	۰.۳۵
Support Vector Machines	۶۲.۶۳	۰.۲۸
Decision Trees	۶۰.۲۷	۰.۴۲
Random Forest	۶۰.۲۷	۰.۴۲
Gaussian Bayes	۶۳.۰۵	۰.۶۳
Adaboosting	۵۷.۷۷	۰.۴۶
GradientBoosting	۶۳.۰۵	۰.۶۳
روش پیشنهادی	۶۵.۶۳	۰.۴۲

همانطور که در جدول ۴-۵ نشان داده شده است، اگر از مرحله کاهش ویژگی استفاده نشود، تعداد ویژگی‌های استخراج شده ۱۳۱۰۷۲ است و اگر برای کاهش ویژگی از ترکیب PCA و LDA استفاده شود، تعداد ویژگی‌ها ۳ است (تعداد کلاس -۱). بنابراین، زمان اجرای الگوریتم کاهش می‌یابد.

مقایسه‌ای از نظر زمان اجرا برای ارزیابی دقیق‌تر در جدول ۴-۱۱ انجام شده است. نتایج نشان می‌دهد که میانگین زمان اجرا در روش پیشنهادی با مرحله کاهش ویژگی بسیار کاهش می‌یابد.

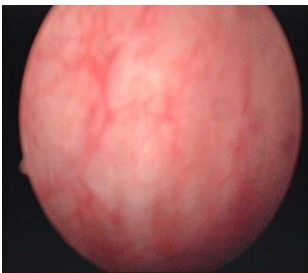
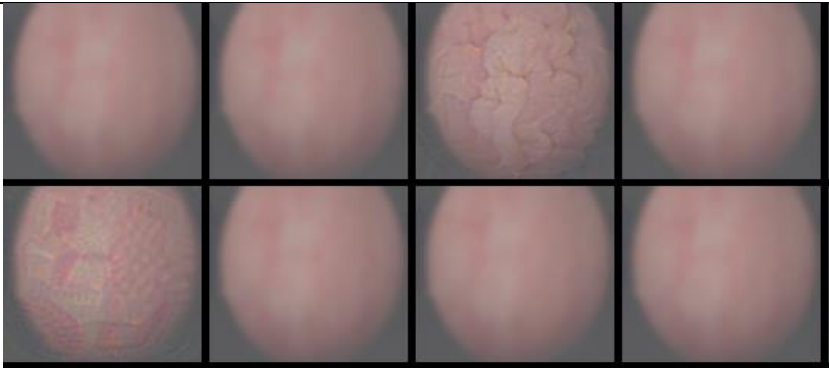
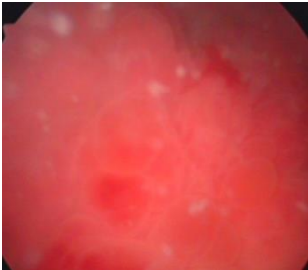
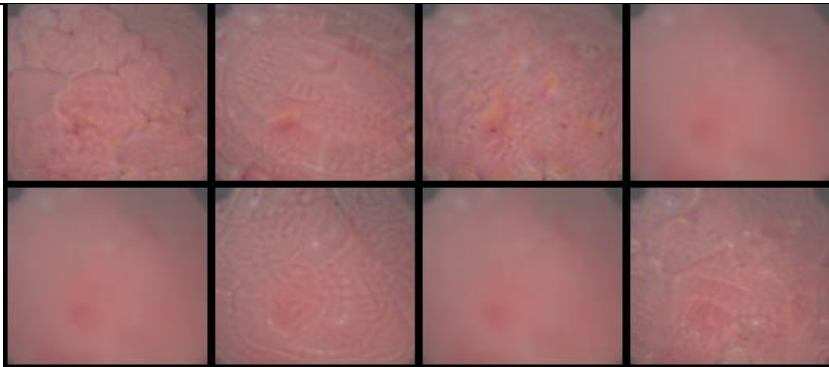
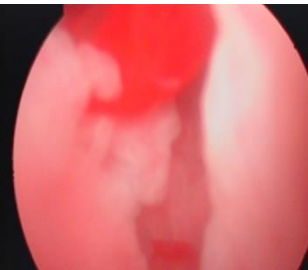
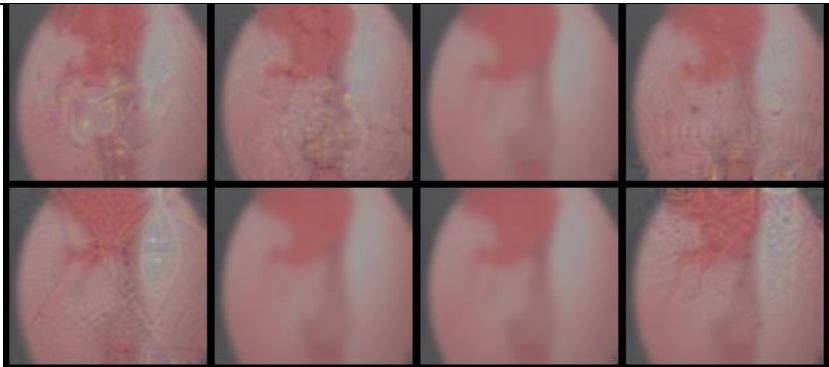
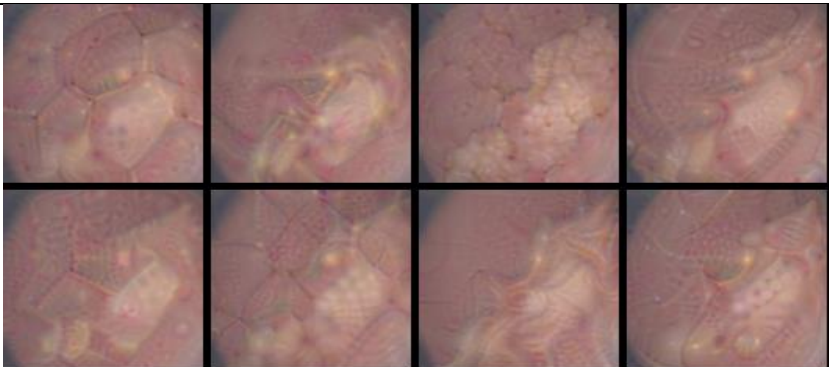
جدول ۴-۱۱ بررسی تأثیر کاهش ویژگی در روش پیشنهادی برای طبقه‌بندی تصاویر سیستم اسکوپ

روش	Mean Time (Second)	Mean Variance	Mean accuracy
روش پیشنهادی با کاهش ویژگی	۲.۳۲	۰.۱۹	۶۹.۰۲
روش پیشنهادی بدون کاهش ویژگی	۱۰.۴	۰.۲۷	۶۷.۳۶

۴ - ۵ آزمایش‌ها و نتایج روش‌های پیشنهادی بر پایه ویژگی‌های سطح بالا و

بخش‌بندی نواحی نگاشت ویژگی

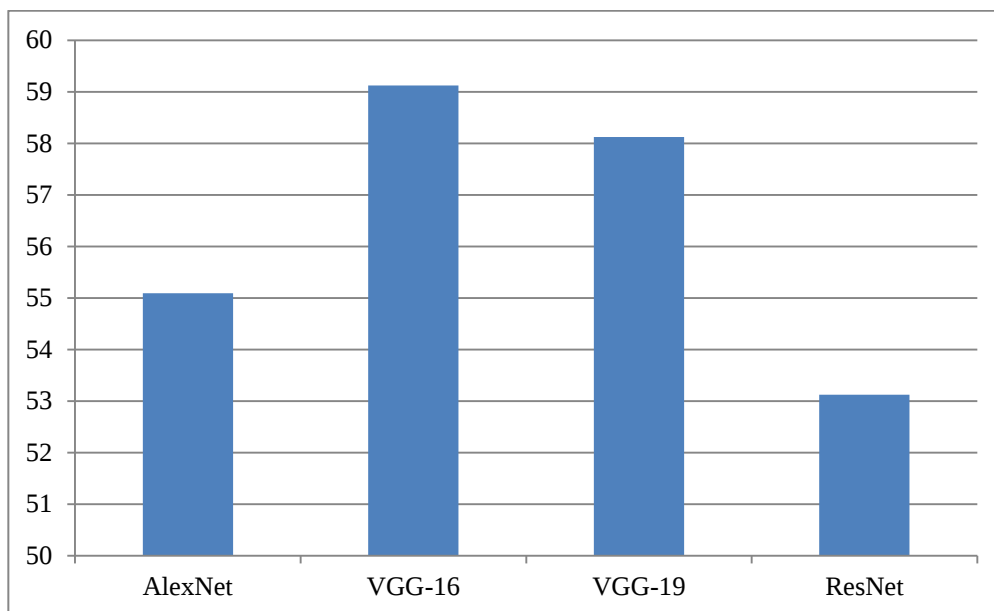
در این بخش آزمایشات انجام شده بر اساس روش پیشنهادی بر اساس یادگیری انتقال، بخش‌بندی نواحی مربوط به نگاشت ویژگی و طبقه‌بندی تصاویر بر اساس این نواحی ارائه می‌شود. در ابتدا باید اشاره نمود که تمامی پارامترهای استفاده شده برای طبقه‌بندی و شرایط پایگاه داده مشابه بخش ۴ - ۴ - ۱ می‌باشد. در ادامه مراحل انجام و تحلیل نتایج بیان می‌شود

Class	Input Image	A sample output of the last convolutional layer
Normal		
Benign		
Blood pee		
Cancer		

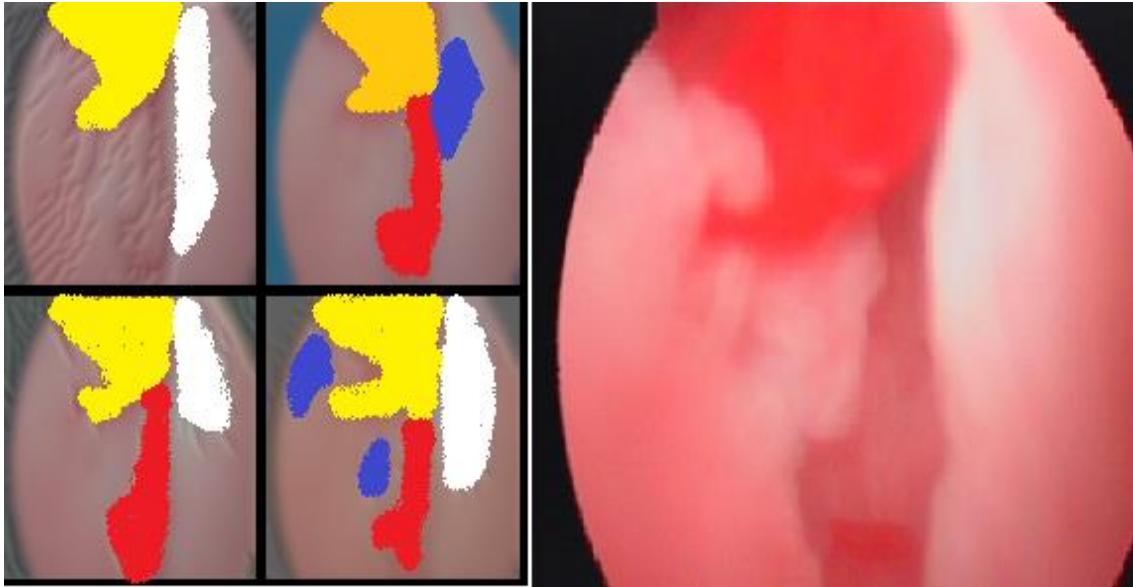
شکل ۴-۹ یک نمونه خروجی از لایه‌های پیچشی در تصاویر سیستوسکوپی برای هر کلاس نرمال، توده خوش خیم، خونی و توده بدخیم

۴ - ۵ - ۱ نتایج بدست آمده از روش بخش‌بندی نواحی نگاشت ویژگی

در این بخش نتایج آزمایشات مربوط به روش پیشنهادی بر پایه بخش‌بندی نواحی نگاشت ویژگی بیان شده است. در مرحله اول، از یادگیری انتقال برای استخراج نگاشت ویژگی استفاده می‌شود. شکل ۴-۹ نمونه‌ای از خروجی لایه‌های پیچشی را در تصاویر سیستم‌سکوپی نشان می‌دهد. شکل ۴-۱۰ مقایسه معماری‌های مختلف در طبقه‌بندی تصویر را نشان می‌دهد. در این بخش، یک لایه کاملاً متصل برای طبقه‌بندی جهت ارزیابی معماری‌های مختلف استفاده شده است. همانطور که مشخص است، شبکه VGG-16 بهترین عملکرد را نشان داده است. در مرحله بعدی از روش‌های IPSO و اتسو برای استخراج نواحی استفاده می‌شود. شکل ۴-۱۱ نمونه‌ای از نواحی انتخاب شده را بر اساس نگاشت ویژگی نشان می‌دهد. همانطور که در روش پیشنهادی بیان شده است ویژگی‌های مربوط به نگاشت ویژگی توسط روش IPSO انتخاب می‌شوند در گام از ویژگی‌های انتخاب شده برای بخش‌بندی توسط اتسو استفاده شده است. این نواحی توسط این روش انتخاب شده است.



شکل ۴-۱۰ مقایسه طبقه‌بندی با معماری‌های مختلف



شکل ۴-۱۱ نمونه ای از نواحی انتخاب شده بر اساس نگاشت ویژگی

پس از کاهش ویژگی‌ها با روش PCA، از روش‌های مختلفی برای طبقه‌بندی استفاده شده است. جدول I مقایسه نتایج بدست آمده را نشان می‌دهد. همانطور که نتایج نشان می‌دهد، روش پیشنهادی بهترین عملکرد را ارائه می‌دهد. در تمام این نتایج، از معماری Vgg-16 برای استخراج نگاشت ویژگی استفاده می‌شود.

جدول ۴-۱۲ مقایسه روش‌های طبقه‌بندی براساس معیارهای صحت (PR)، یادآوری (حساسیت) (RE)، خاص بودن (Spe) و F1 در هر کلاس به صورت جداگانه با روش پیشنهادی بر پایه بخش‌بندی

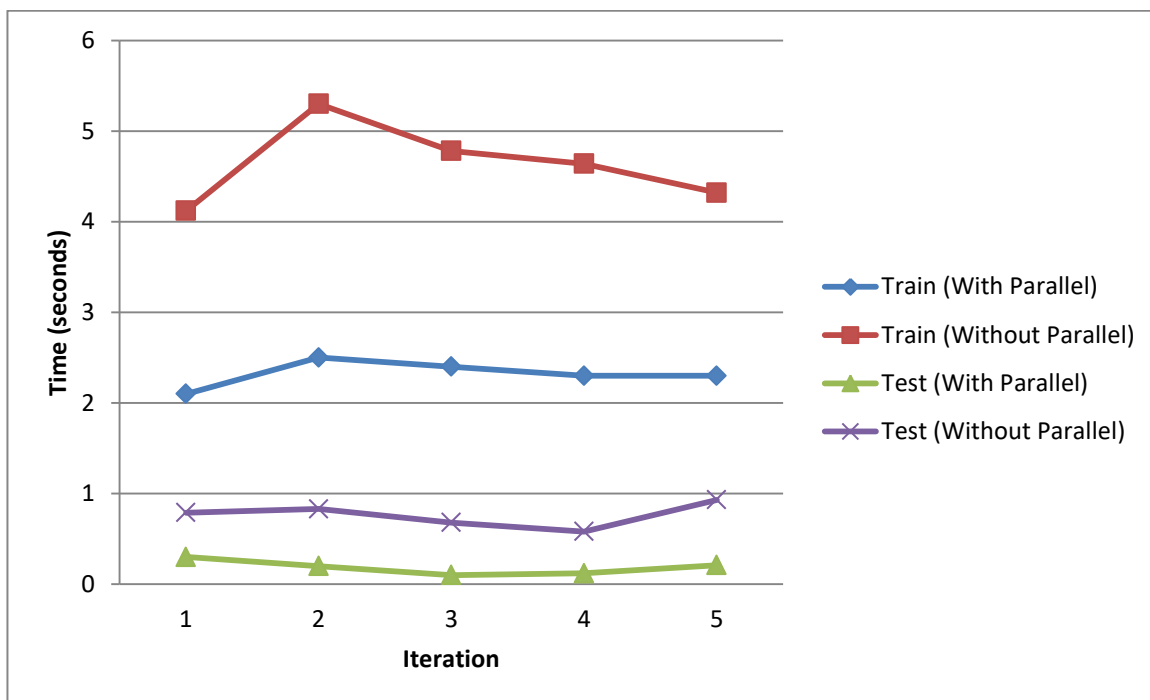
Methods	Label	Precision	F1	Specificity	Recall(sensitivity)
Logistic Regression	Benign	0.58	0.63	0.84	0.60
	Blood pee	0.65	0.55	0.90	0.59
	Cancer	0.79	0.44	0.96	0.56
	Normal	0.74	0.80	0.90	0.76
SVM	Benign	0.57	0.6	0.84	0.58
	Blood pee	0.56	0.69	0.81	0.61
	Cancer	0.62	0.53	0.89	0.57
	Normal	0.89	0.85	0.96	0.86
Random Forest	Benign	0.5	0.53	0.82	0.51
	Blood pee	0.64	0.6	0.88	0.61
	Cancer	0.67	0.64	0.89	0.65
	Normal	0.5	0.62	0.79	0.55
GB	Benign	0.68	0.52	0.91	0.58
	Blood pee	0.59	0.72	0.83	0.64
	Cancer	0.63	0.57	0.88	0.59
	Normal	0.90	0.8	0.97	0.84

Majority Voting (BG, SVM)	Benign	0.7	0.77	0.89	0.73
	Blood pee	0.72	0.73	0.90	0.72
	Cancer	0.69	0.68	0.89	0.68
Proposed Method	Normal	0.71	0.72	0.90	0.71

۴ - ۵ - ۲ تحلیل پیچیدگی زمانی

در این بخش روش پیشنهادی از نظر پیچیدگی زمانی بررسی می‌شود. همانطور که گفته شد، در روش پیشنهادی از مجموعه از قبل آموزش دیده استفاده شده است. در مرحله بعد، پس از استخراج نواحی توسط الگوریتم‌های IPSO و اتسو، از PCA برای کاهش ویژگی استفاده می‌شود. در انتها، طبقه‌بندی مجزا برای هر ناحیه به صورت موازی اجرا می‌شود.

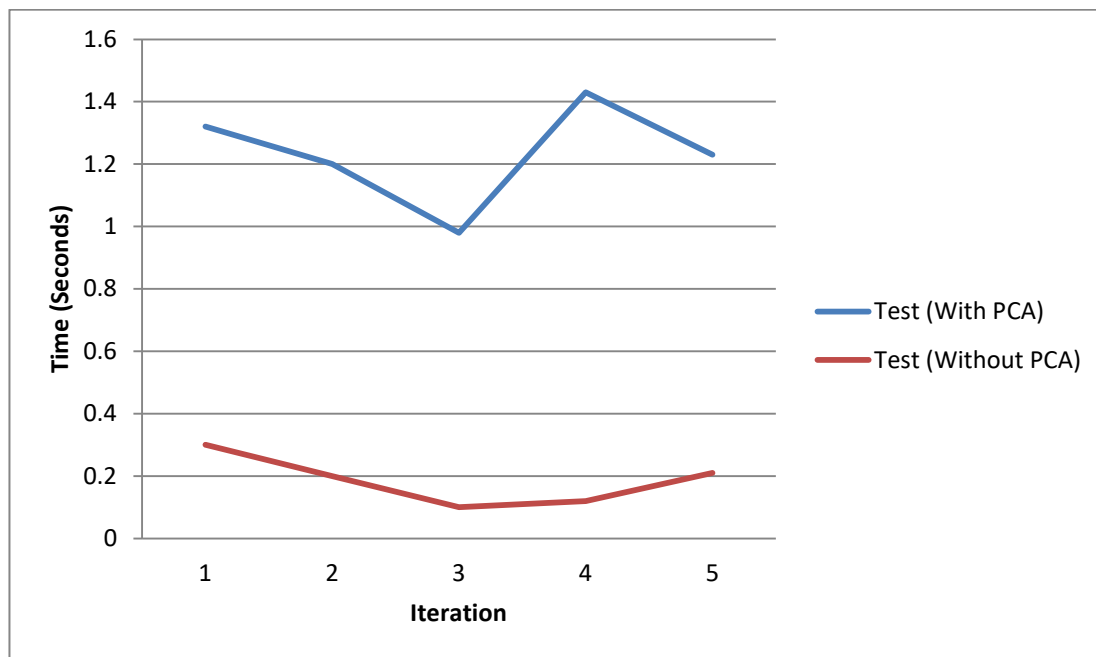
شکل ۴-۱۲ زمان اجرای الگوریتم را با پردازش موازی و بدون آن نشان می‌دهد. در زمان که از موازی‌سازی استفاده می‌شود عمل طبقه‌بندی بر اساس نواحی هر بخش به صورت همزمان انجام می‌گیرد. همانطور که در این شکل نشان داده شده است، زمان الگوریتم هم در مجموعه آموزش و هم در آزمون کاهش یافته است.



شکل ۴-۱۲ مقایسه زمان اجرای الگوریتم پیشنهادی برای مجموعه آموزش و آزمون، در دو حالت با/بدون Parallel

شکل ۴-۱۳ زمان اجرا در مجموعه آزمون را نشان می‌دهد که وقتی PCA برای کاهش ویژگی

استفاده می‌شود. همانطور که در این شکل نشان داده شده است، زمان اجرا به دلیل کاهش ویژگی‌ها کاهش می‌یابد. لازم به ذکر است که زمان اجرا به صورت متوسط برای ۵ بار اجرا برای تمام تصاویر مجموعه آموزش و آزمون گزارش شده است.



شکل ۴-۱۳ زمان اجرای الگوریتم در مجموعه آزمون، با/بدون PCA

۴ - ۶ مقایسه روش‌های مختلف و ارزیابی نتایج

در این بخش روش‌های پیشنهادی در کاربرد طبقه‌بندی تصاویر سیستم‌سکوپی مثانه مقایسه می‌شود. جدول ۴-۱۳ مقایسه نتایج روش‌های پیشنهاد شده را نشان می‌دهد. روش‌های معرفی شامل سه دسته کلی بر اساس ویژگی‌های سطح پایین و سطح بالا به همراه بخش بندی نواحی نگاشت ویژگی می‌باشد. در روش اول از شبکه MLP برای طبقه‌بندی استفاده شده است. لازم به ذکر است که در این روش برای انتخاب وزن‌های اولیه MLP، از یک روش مبتنی بر جستجوی تکاملی (الگوریتم ژنتیک) استفاده شده است. برای استخراج ویژگی‌ها از تصاویر سیستم‌سکوپی، از روش LBP استفاده شد. همانطور که گفته شد، از روش PCA به دلیل تعداد زیاد ویژگی‌های استخراج شده توسط LBP برای کاهش ویژگی‌ها استفاده شده است.

در اینجا برای بررسی عملکرد CNN از دو معماری VGG16 و ResNet برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه استفاده شده است. مقایسه نتایج نشان می‌دهد که لایه‌های پیچشی در CNN ویژگی‌های بهتر و موثرتری را در مقایسه با روش LBP سنتی استخراج کرده‌اند. از این لایه‌ها می‌توان برای استخراج ویژگی‌ها استفاده کرد.

در ادامه نتایج روش‌های ارائه شده بر اساس ویژگی‌های سطح بالا نشان داده شده است. برای این منظور نتایج مربوط به طبقه‌بندی بر اساس ویژگی یادگیری انتقال و طبقه‌بند ترکیبی ارائه شده است. همانطور که از نتایج مشخص می‌باشد این روش نسبت به روش‌هایی که از ویژگی سنتی استفاده می‌کردند عملکرد بهتری داشته است. سپس نتایج مربوط به طبقه‌بندی بر اساس استخراج ویژگی سطح بالا بر پایه نواحی بخش‌بندی شده ارائه شده است که با توجه به نتایج بدست آمده بهترین عملکرد را از خود نشان داده است.

همچنین باید اشاره نمود که، تنها مطالعه انجام شده بر روی این تصاویر توسط مرجع [۷۵] ارائه شده است. در این روش ترکیبی از شبکه عصبی پیچشی عمیق برای تشخیص ناهنجاری‌ها در تصاویر سیستم‌سکوپی معرفی شد. طبقه‌بند پایه برای ترکیب VGGNets بوده است. که با استفاده از روش جمع‌آوری^۱ bootstrap ترکیب شدند.

جدول ۴-۱۳ مقایسه روش‌های موجود در طبقه‌بندی تصاویر سیستم‌سکوپی مثانه

Methods	#class	#images	Accuracy	Variance
Randomized weighing and variable learning rate	3	540	41.57	-
Randomized weighing and constant learning rate	3	540	38.18	-
روش پیشنهادی بر پایه MLP+GA	3	540	49.36	
Vgg16	4	720	58.68	-
Resnet	4	720	54.37	-
Ensemble VGG net [75]	4	720	63	
روش پیشنهادی بر پایه یادگیر انتقال و طبقه‌بند ترکیبی	4	720	69.02	0.19

¹ aggregating

روش پیشنهادی بر پایه بخش - بندی نواحی نگاشت ویژگی	4	720	72.5	0.00023
--	---	-----	------	---------

در این رساله، یک روش پویا و قدرتمند مبتنی بر یادگیری عمیق برای طبقه بندی تصاویر سیستم‌سکوپی مثانه ارائه شده است. به دلیل تعداد محدود نمونه‌ها در تصاویر سیستم‌سکوپی مثانه، از روش یادگیری انتقال برای استخراج ویژگی‌های سطح بالا استفاده شده است. در واقع، در این مرحله، از خروجی لایه پیچشی یا همان نگاشت ویژگی استفاده شد. از روش‌های IPSO و اتسو برای بخش‌بندی نواحی ویژگی‌ها بر اساس مقدار آنها استفاده شد. در واقع با استفاده از این روش، نواحی با سطوح ویژگی ارزشمندتر از نگاشت ویژگی استخراج می‌شوند. در مرحله بعدی، از روش PCA برای کاهش ویژگی استفاده می‌شود. سرانجام، از یک روش ترکیبی مبتنی بر SVM و GB برای طبقه بندی تصاویر استفاده شده است. روش پیشنهادی مبتنی بر معماری شبکه عصبی عمیق VGG-16 بر روی تصاویر سیستم‌سکوپی مثانه با چهار دسته مختلف بود. نتایج نشان دهنده عملکرد بالای روش پیشنهادی در مقایسه با سایر روش‌ها است.

تشخیص به موقع سرطان مثانه به طور قابل توجهی احتمال موفقیت در درمان این نوع از سرطان‌های را افزایش می‌دهد. در مراحل اولیه، پزشکان بیماران را برای تصویربرداری سیستم‌سکوپی از مثانه جهت بررسی بصری داخل مثانه معرفی می‌کنند. تصویربرداری سیستم‌سکوپی، به خصوص در مراحل اولیه در مقایسه با آزمایشات آسیب شناسی، بسیار مقرون به صرفه است. بینایی ماشین و یادگیری ماشین بیش از سه دهه است که در زمینه مهندسی پزشکی برای طراحی سیستم‌های مختلف تشخیص به کمک رایانه مورد توجه محققان قرار گرفته‌اند. با این حال، عدم وجود یک مجموعه داده استاندارد از تصاویر سیستم‌سکوپی مثانه باعث عدم رشد تعداد مطالعات در این حوزه شد. در این رساله، یک مجموعه داده از تصاویر سیستم‌سکوپی جمع آوری شده از یک مرکز پزشکی در کشور هلند برای آزمایشات استفاده شده است. تصاویر سیستم‌سکوپی به چهار دسته نرمال، ادرار خونی، توده خوش خیم و توده بدخیم طبقه بندی شدند. تعداد محدود نمونه‌های آموزشی یکی از اصلی‌ترین چالش‌های

معمول در طبقه‌بندی تصویر پزشکی است که می‌تواند به شدت بر دقت طبقه‌بندی تأثیر بگذارد. برای غلبه بر این مشکل، از یادگیری انتقال استفاده شد تا از ویژگی‌های افتراقی استخراج شده از لایه‌های پیچش قبلی استفاده شود. در این راستا، CNN‌های از قبل آموزش دیده (به عنوان مثال، CNN‌هایی که روی مجموعه داده‌های ImageNet آموزش دیده‌اند) برای استخراج ویژگی بدون هیچ گونه آموزش بیشتر استفاده شد. در حقیقت، CNN‌ها برای استخراج ویژگی‌های و نه طبقه‌بندی استفاده شدند. روش پیشنهادی نتایج امیدوار کننده‌ای را در طبقه‌بندی تصاویر سیستم‌سکوپی نشان می‌دهد. یافته‌های از این روش به شرح زیر است:

(۱) بر اساس نتایج، معماری VGG16 از سایر معماری‌های رقیب برای استخراج و یادگیری انتقال بدون نظارت بهتر عمل کرد.

(۲) کاهش تعداد ویژگی‌های استخراج شده از لایه سطح با استفاده از PCA و به دنبال آن LDA به طور قابل توجهی دقت و پیچیدگی زمان کار طبقه‌بندی را بهبود می‌بخشد.

(۳) استفاده از یک طبقه‌بند واحد برای کاربردهای یادگیری ماشین معمول است. با این حال، روش‌های مختلف به دلیل پیچیدگی و تعداد کم تصاویر سیستم‌سکوپی، نتایج متفاوتی را ارائه می‌دهند. برای دستیابی به بهترین نتایج، چندین روش و همچنین طبقه‌بند ترکیبی از روش‌های پایه مورد تجزیه و تحلیل و ارزیابی قرار گرفت. افزایش تنوع در داده‌ها و سطح ویژگی‌ها به دلیل کمبود نمونه‌های آموزشی و همچنین تعداد ویژگی‌ها به دلیل کاهش شدید ویژگی با استفاده از PCA و به دنبال آن LDA بسیار چالش برانگیز است. از این رو، در این مطالعه از تنوع در سطح طبقه‌بندی استفاده شد. از روش‌های مختلف طبقه‌بندی برای بدست آوردن بهترین مجموعه از طبقه‌بندهای پایه استفاده شد. در روش پیشنهادی، وزن طبقه‌بندهای پایه با توجه به عملکرد آنها بهینه‌سازی شد. آزمایشات نتایج امیدوار کننده‌ای را به نفع روش ترکیبی پیشنهادی نشان داد.

۴ - ۷ جمع‌بندی

در این فصل سه روش برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه بر اساس پیچیدگی این نوع از تصاویر معرفی شده است. در روش اول، یک روش جدید بر اساس ویژگی سطح پایین LBP و شبکه عصبی پرسپترون معرفی شد. برای بهبود عملکرد این روش از الگوریتم تکاملی ژنتیک به همراه نرخ یادگیری تطبیقی استفاده شده است. نتایج نشان دهنده بهبود عملکرد روش پیشنهاد می‌باشد. در روش پیشنهادی بعدی از یادگیری انتقال بر اساس شبکه CNN برای استخراج ویژگی استفاده شد. برای جلوگیری از مشکل ابعاد بالا و همچنین کاهش زمان و پیچیدگی فضا، از روش‌های کاهش بعد استفاده شده است. سرانجام، یک طبقه‌بند ترکیبی برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه پیشنهاد شد. در روش پیشنهادی برای ترکیب از طبقه‌بندهای شناخته شده با استفاده از روش آرای اکثریت وزنی استفاده شده است. روش پیشنهادی نتایج امیدوار کننده مربوط به طبقه‌بندی این نوع از تصاویر سیستم‌سکوپی را نشان می‌دهد. در ادامه بهبود روش ارائه شده، از یادگیری انتقال برای استخراج نواحی مربوط به نگاشت‌های ویژگی استفاده شده است. آخرین لایه کانولوشن برای انتخاب نگاشت ویژگی استفاده شد. در مرحله بعدی، ویژگی‌های سطح بالا و نواحی مربوط به این ویژگی‌ها بر اساس روش بهینه‌سازی ازدحام ذرات بهبود یافته (IPSO) و اتسو استخراج شد. سپس، با استفاده از روش PCA، ویژگی استخراج شده در بخش‌ها کاهش می‌یابد. در این روش، از یک روش ترکیبی مبتنی بر رای‌گیری اکثریت با BG و SVM استفاده شده است. نتایج بیانگر بهبود عملکرد روش پیشنهادی نسبت به روش‌های موجود در حیطه طبقه‌بندی تصاویر سیستم‌سکوپی مثانه است.

۵ جمع بندی و کارهای آتی

۵ - ۱ مقدمه

در این فصل جمع‌بندی کارهای انجام شده و تحلیل نتایج بدست آمده در رساله ارائه می‌شود. در ادامه کارهای آتی در بهبود سیستم‌های طبقه‌بندی تصاویر سیستم‌سکوپی مثانه پیشنهاد می‌شود.

۵ - ۲ جمع‌بندی پژوهش انجام شده در رساله

در طول سه دهه اخیر، استفاده از روش‌های هوشمند در سامانه‌های تشخیص پزشکی مورد توجه بسیاری از محققان قرار گرفته است. ولی برای تشخیص سرطان مثانه از طریق تصاویر سیستم‌سکوپی با وجود شیوع بسیار زیاد در سطح جهان فعالیت هوشمندی در حوزه پردازش تصویر پزشکی ارائه نشده است. در این راستا در روش پیشنهادی اول این رساله، از شبکه عصبی چندلایه برای رده‌بندی تصاویر سیستم‌سکوپی مثانه استفاده شده است. یکی از مهمترین مشکلات در الگوریتم‌های شبکه عصبی تعیین میزان نرخ یادگیری می‌باشد اگر نرخ یادگیری کوچک انتخاب شود باعث کاهش سرعت همگرایی و اگر خیلی بزرگ انتخاب شود موجب نوسان و واگرایی می‌شود لذا نیاز به الگوریتمی داریم که بتواند بصورت پویا نرخ همگرایی را تغییر دهد. در این راستا برای بهبود شبکه عصبی چند لایه روشی تطبیقی برای تعیین نرخ یادگیری ارائه شده است. در این روش نرخ یادگیری با استفاده از یک ضربی بر اساس اختلاف دقت آموزش و اعتبارسنجی با توجه به خروجی خطا مشخص می‌شود. نرخ تغییرات با توجه به میزان تغییرات وزن و خطای خروجی بروز رسانی می‌شود. روش پیشنهادی بر روی پایگاه داده جمع‌آوری شده از تصاویر سیستم‌سکوپی مثانه مورد ارزیابی قرار گرفته است. نتایج شبیه‌سازی نشان دهنده کاهش خطا و افزایش سرعت همگرایی روش پیشنهادی نسبت به روش‌های دیگر در رده‌بندی تصاویر سیستم‌سکوپی می‌باشد.

یک روش مرسوم برای طبقه‌بندی تصاویر پزشکی استفاده از ویژگی‌های سطح پایین می‌باشد. در همین راستا در ادامه روش‌های ارائه شده، در گام نخست روشی بر پایه الگوهای باینری محلی با

شبکه‌های عصبی پرسپترون ارائه شده است. مراحل اصلی روش پیشنهادی شامل استخراج و کاهش ویژگی، بهبود شبکه عصبی با نرخ یادگیری تطبیقی و به دست آوردن وزن‌های اولیه توسط الگوریتم ژنتیک است. در این روش، یک رویکرد نوین از MLP برای پیش‌بینی و طبقه‌بندی تصاویر سیستم‌سکوپی با دقت و صحت بالا ارائه شده است.

انتخاب وزن‌های اولیه در بهبود عملکرد شبکه MLP بسیار تاثیرگذار هست. معمولاً از وزن‌هایی به صورت تصادفی برای وزن‌های اولیه استفاده می‌شود. استفاده از وزن‌های اولیه تصادفی در مجموعه داده‌های کوچک می‌تواند اثرات مخربی داشته باشد و آموزش شبکه به خوبی انجام نمی‌شود. معمولاً تصاویر پزشکی شامل مجموعه داده‌های بزرگ نمی‌باشند و تعداد نمونه‌های استخراجی از بیماران کم است. در همین راستا، نباید از وزن‌های اولیه تصادفی برای آموزش شبکه استفاده نمود. در واقع در مجموعه داده‌های کوچک باید وزن‌های اولیه با دقت مناسب‌تری انتخاب شوند. به همین دلیل، یک روش نوین بر پایه الگوریتم ژنتیک برای انتخاب وزن‌های اولیه شبکه MLP ارائه شده است.

نتایج بدست آمده از روش پیشنهادی بر پایه LBP، عملکرد بهتری نسبت به روش‌های دیگر بر پایه ویژگی‌های سنتی یا سطح پایین از خود نشان داده است. ولی همچنان این نتایج برای کاربردهای واقعی مناسب نمی‌باشند در همین راستا در ادامه روش‌های ارائه شده، یک روش نوین برپایه یادگیری عمیق پیش‌بینی برای طبقه‌بندی تصاویر سیستم‌سکوپی ارائه شده است. نتایج نشان دهنده بهبود عملکرد روش پیشنهادی نسبت به روش‌های قبل می‌باشد.

در ادامه، یادگیری انتقال بر اساس شبکه CNN برای استخراج ویژگی استفاده شد. روش پیشنهادی به گونه‌ای طراحی شده است که نسبت به تغییرات شرایط تصویربرداری مستقل باشد. این روش با استخراج یک ویژگی سطح بالا از تصاویر سیستم‌سکوپی آغاز می‌شود. از آنجا که اثبات شده است که یادگیری عمیق در استخراج ویژگی با سطوح معنایی مختلف موثر است، این نوع شبکه برای استخراج ویژگی استفاده شده است. به منظور حل مشکل تعداد محدود تصاویر سیستم‌سکوپی، از انتقال

یادگیری استفاده شد تا با استفاده از مجموعه آموزش کوچک موجود، یک مدل یادگیری مناسب ساخته شود. در رویکرد یادگیری انتقال، آخرین لایه‌های یک شبکه از پیش آموزش دیده، که قبلاً روی یک مجموعه داده بزرگ مانند ImageNet آموزش دیده‌اند، استفاده شد. برای جلوگیری از مشکل ابعاد بالا و همچنین کاهش زمان و پیچیدگی فضا، از تعداد ویژگی‌ها با استفاده از روش‌های کاهش بعد کاسته می‌شود. سرانجام، یک طبقه‌بند ترکیبی برای طبقه‌بندی تصاویر سیستم‌سکوپی مثانه پیشنهاد می‌شود. در روش پیشنهادی برای ترکیب از طبقه‌بندهای شناخته شده با استفاده از روش آرای اکثریت وزنی استفاده شده است. روش پیشنهادی نتایج امیدوار کننده مربوط به طبقه‌بندی این نوع از تصاویر سیستم‌سکوپی را نشان می‌دهد.

در ادامه بهبود روش‌های ارائه شده، از یادگیری انتقال برای استخراج نواحی مربوط به نگاشت‌های ویژگی استفاده شده است. آخرین لایه کانولوشن برای انتخاب نگاشت ویژگی استفاده شد. در مرحله بعدی، ویژگی‌های سطح بالا و نواحی مربوط به این ویژگی‌ها بر اساس روش‌های بهینه سازی ازدحام ذرات بهبود یافته (IPSO) و اتسو استخراج می‌شوند. سپس، با استفاده از روش PCA، ویژگی استخراج شده در بخش‌ها کاهش می‌یابد. در آخر، طبقه بندی انجام می‌شود. در این روش، از یک روش ترکیبی مبتنی بر رای گیری اکثریت با BG و SVM استفاده شده است. نتایج بدست آمده نشان دهنده بهترین عملکرد روش‌های موجود در طبقه‌بندی تصاویر سیستم‌سکوپی می‌باشد.

در مجموع می‌توان نوآوری‌های ارائه شده در این رساله را به صورت زیر بیان نمود:

۱. روش‌های سنتی یادگیری ماشین قادر به ارائه نتایج قابل قبولی برای استخراج ویژگی با توجه به تعداد محدود تصاویر سیستم‌سکوپی نیستند. انتقال یادگیری، که مبتنی بر یادگیری عمیق است، یکی از موثرترین روش‌ها برای استخراج ویژگی‌ها از مجموعه کوچکی از داده‌های آموزش است. در این روش، آموزش اولیه بر اساس مجموعه ای گسترده از ImageNet انجام شده است و ویژگی‌های مرتبط استخراج می‌شوند. ویژگی‌های استخراج شده از نظر بصری با

تصاویر سیستم‌سکوپی متفاوت هستند. از این رو، این ویژگی‌ها برای تصاویر سیستم‌سکوپی قابل استفاده هستند.

۲. استفاده از یادگیری عمیق در استخراج ویژگی دارای مزایای زیر می‌باشد:

آ. استقلال ورودی: با توجه به ماهیت عمیق یادگیرنده، اندازه تصاویر ورودی در استخراج ویژگی در لایه‌های بالاتر تاثیری ندارد.

ب. مقاومت به چرخش و زاویه: حتی اگر تصاویر در یک محیط ایزوله و مشخص گرفته نشده باشد و تصاویر موقعیت‌های مختلف مربوط به نواحی بیماری گرفته باشد، یادگیری عمیق نتایج امیدوارکننده‌ای را نشان می‌دهد.

ج. مقاومت در برابر تغییرات روشنایی و محیط: ترکیب لایه‌های مختلف برای شبکه‌های عصبی عمیق، اثر تغییر محیط و شرایط روشنایی را کاهش می‌دهد. بنابراین، ویژگی‌های استخراج شده نسبت به محیط در حال تغییر مقاوم ارزیابی می‌شوند.

۳. انتخاب مناسب معماری در یادگیری انتقال بسیار موثر است. در این روش، معماری‌های مختلف عمیق مانند VGG-16، VGG-19، AlexNet و ResNet بر روی تصاویر سیستم‌سکوپی ارزیابی می‌شوند. نتایج نشان می‌دهد که معماری VGG-16 نسبت به مجموعه اعتبارسنجی و مجموعه آزمون از سایر معماری‌های رقیب بهتر است.

۴. تعداد و ابعاد بالای ویژگی‌های استخراج شده باعث ایجاد برخی مسائل پیچیده برای مراحل بعدی تجزیه و تحلیل و تشخیص می‌شود. در این روش، PCA برای ارزیابی ویژگی‌ها و کاهش تعداد ویژگی‌ها استفاده شده است. نتایج نشان می‌دهد که کاهش ویژگی دقت طبقه‌بندی را بهبود می‌بخشد و پیچیدگی زمان را کاهش می‌دهد.

۵. انتخاب مناسب تعداد ویژگی‌های انتخاب شده بر عملکرد طبقه‌بندی در این حوزه تأثیر می‌گذارد. هنگامی که ویژگی‌ها توسط بردارهای ویژه PCA استخراج می‌شوند و

طبقه‌بندی‌کننده بر روی مجموعه اعتبار سنجی ارزیابی می‌شود، یک بردار ویژگی از ۱۰۰ ویژگی پس از مرحله نگاشت ویژگی برای طبقه‌بند استفاده می‌شود.

۶. پس از کاهش ویژگی‌ها توسط PCA، LDA با در نظر گرفتن برچسب‌های کلاس استفاده می‌شود. استفاده متوالی از LDA پس از ویژگی‌های وزن‌دهی توسط PCA، عملکرد طبقه بندی تصاویر سیستم‌سکوپی را بهبود می‌بخشد.

۷. استفاده از یک طبقه‌بند تنها برای کاربردهای بر اساس یادگیری ماشین معمول است در حالی که روش‌های مختلف به دلیل پیچیدگی و تعداد کم تصاویر سیستم‌سکوپی نتایج متفاوتی را ارائه می‌دهند. برای دستیابی به بهترین نتیجه، چندین روش تحلیل و ارزیابی می‌شود. در مرحله بعدی، ترکیبی از طبقه‌بندهای پایه استفاده می‌شوند. آزمایشات نتایج امیدوار کننده‌ای را به نفع روش پیشنهادی ترکیبی را نشان می‌دهد.

۸. کاهش تعداد ویژگی‌های استخراج شده از لایه مسطح با استفاده از PCA و به دنبال آن LDA به طور قابل توجهی دقت و پیچیدگی زمان کار طبقه بندی را بهبود می‌بخشد.

۹. استفاده از یک طبقه‌بند واحد برای کاربردهای یادگیری ماشین معمول است. با این حال، روش‌های مختلف به دلیل پیچیدگی و تعداد کم تصاویر سیستم‌سکوپی، نتایج متفاوتی را ارائه می‌دهند. برای دستیابی به بهترین نتایج، چندین روش و همچنین طبقه بند ترکیبی از روش‌های پایه مورد تجزیه و تحلیل و ارزیابی قرار گرفت. افزایش تنوع در داده‌ها و سطح ویژگی‌ها به دلیل کمبود نمونه‌های آموزشی و همچنین تعداد ویژگی‌ها به دلیل کاهش شدید ویژگی با استفاده از PCA و به دنبال آن LDA بسیار چالش برانگیز است. از این رو، در این مطالعه از تنوع در سطح طبقه بندی استفاده شد. از روش‌های مختلف طبقه‌بندی برای بدست آوردن بهترین مجموعه از طبقه‌بندهای پایه استفاده شد. در روش پیشنهادی، وزن طبقه‌بندهای پایه با توجه به عملکرد آنها بهینه سازی شد. آزمایشات نتایج امیدوار کننده ای را به نفع روش ترکیبی پیشنهادی نشان داد.

۵ - ۳ پیشنهادهایی برای کارهای آینده

در این پژوهش راهکاری نوین برای طبقه‌بندی تصاویر سیستم‌توسکویی مثانه ارائه شده است. در این بخش، پیشنهادهایی برای کارهای آینده آورده می‌شود.

(۱) یکی از چالش‌های اصلی در طبقه بندی تصاویر سیستم‌توسکویی مثانه، فقدان مجموعه داده استاندارد با نمونه کافی برای تحقیقات مبتنی بر یادگیری ماشین و بینایی ماشین است. برای این منظور، در اولین مرحله باید یک مجموعه داده استاندارد جمع آوری شود.

(۲) نوآوری این رساله استفاده از روش‌های یادگیری انتقالی برای استخراج ویژگی‌های موثر مبتنی بر یادگیری عمیق است. همانطور که در فصل‌های قبل بحث شد، دلیل استفاده از این روش فقدان مجموعه داده‌ای به اندازه کافی بزرگ برای آموزش صحیح شبکه بود. برای کارهای آینده توصیه می‌شود از معماری‌های دیگری مانند شبکه‌های کپسولی استفاده شود که در مقایسه با معماری‌های دیگر به آموزش با وزن کمتری نیاز دارند.

(۳) پیشنهاد می‌شود از روش‌های انتخاب ویژگی و کاهش ویژگی در لایه‌های مختلف استفاده شود، که می‌تواند منجر به استخراج دقیق تر ویژگی‌ها شود.

٦ منابع

- [1] M. Sullivan, S. Bhattacharjee, R. Wynn, and N. Nadiminty, "MP57-20 PARP INHIBITORS AS POTENTIAL THERAPEUTIC AGENTS AGAINST BLADDER CANCER," *The Journal of Urology*, vol. 201, pp. e820-e820, 2019.
- [2] A. Anees, A. M. Siddiqui, and F. Ahmed, "Chaotic substitution for highly autocorrelated data in encryption algorithm," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, pp. 3106-3118, 2014.
- [3] J. Ahmad and S. O. Hwang, "Chaos-based diffusion for highly autocorrelated data in encryption algorithms," *Nonlinear Dynamics*, vol. 82, pp. 1839-1850, 2015.
- [4] H. Jütte, M. Reike, R. M. Wirtz, M. Kriegmair, P. Erben, K. Tully, *et al.*, "KRT20, KRT5, ESR1 and ERBB2 Expression Can Predict Pathologic Outcome in Patients Undergoing Neoadjuvant Chemotherapy and Radical Cystectomy for Muscle-Invasive Bladder Cancer," *Journal of Personalized Medicine*, vol. 11, p. 473, 2021.
- [5] L. Cai, J. Gao, and D. Zhao, "A review of the application of deep learning in medical image classification and segmentation," *Annals of translational medicine*, vol. 8, 2020.
- [6] H. Haick and N. Tang, "Artificial intelligence in medical sensors for clinical decisions," *ACS nano*, vol. 15, pp. 3557-3567, 2021.
- [7] Z. Yu-Qian, G. Wei-Hua, C. Zhen-Cheng, T. Jing-Tian, and L. Ling-Yun, "Medical images edge detection based on mathematical morphology," in *2005 IEEE engineering in medicine and biology 27th annual conference*, 2006, pp. 6492-6495.
- [8] J. S. Duncan and N. Ayache, "Medical image analysis: Progress over two decades and the challenges ahead," *IEEE transactions on pattern analysis and machine intelligence*, vol. 22, pp. 85-106, 2000.
- [9] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, *et al.*, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60-88, 2017.
- [10] K. V. Kumari and S. S. Barpanda, "2. An ample review on brain image classification for Alzheimer's disease," in *Predictive Intelligence in Biomedical and Health Informatics*, ed: De Gruyter, 2020, pp. 27-48.
- [11] G. Haskins, U. Kruger, and P. Yan, "Deep learning in medical image registration: a survey," *Machine Vision and Applications*, vol. 31, p. 8, 2020.
- [12] A. K. Gárate-Escamila, A. H. El Hassani, and E. Andrès, "Classification models for heart disease prediction using feature selection and PCA," *Informatics in Medicine Unlocked*, vol. 19, p. 100330, 2020.
- [13] D. Wang, N. Pang, Y. Wang, and H. Zhao, "Unlabeled skin lesion classification by self-supervised topology clustering network," *Biomedical Signal Processing and Control*, vol. 66, p. 102428, 2021.
- [14] M. Ghasemi, M. Kelarestaghi, F. Eshghi, and A. Sharifi, "AFDL: a new adaptive fuzzy dictionary learning for medical image classification," *Pattern Analysis and Applications*, vol. 24, pp. 145-164, 2021.
- [15] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, pp. 1798-1828, 2013.
- [16] H. Greenspan, B. Van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE transactions on medical imaging*, vol. 35, pp. 1153-1159, 2016.
- [17] S. Dash, B. R. Acharya, M. Mittal, A. Abraham, and A. Kelemen, *Deep learning techniques for biomedical and health informatics*: Springer, 2020.
- [18] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85-117, 2015.
- [19] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354-377, 2018.
- [20] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural computation*, vol. 18, pp. 1527-1554, 2006.
- [21] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science*, vol. 313, pp. 504-507, 2006.
- [22] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in *Advances in neural information processing systems*, 2007, pp. 153-160.

- 
- [23] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. De Vries, M. J. Benders, and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE transactions on medical imaging*, vol. 35, pp. 1252-1261, 2016.
 - [24] G. E. Hinton, "A practical guide to training restricted Boltzmann machines," in *Neural networks: Tricks of the trade*, ed: Springer, 2012, pp. 599-619.
 - [25] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, pp. 1929-1958, 2014.
 - [26] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in *Neural networks: Tricks of the trade*, ed: Springer, 2012, pp. 437-478.
 - [27] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1-9.
 - [28] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, *et al.*, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical image analysis*, vol. 36, pp. 61-78, 2017.
 - [29] Y. Song, L. Zhang, S. Chen, D. Ni, B. Lei, and T. Wang, "Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning," *IEEE Transactions on Biomedical Engineering*, vol. 62, pp. 2421-2433, 2015.
 - [30] W. Yang, Y. Chen, Y. Liu, L. Zhong, G. Qin, Z. Lu, *et al.*, "Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain," *Medical image analysis*, vol. 35, pp. 421-433, 2017.
 - [31] A. Prason, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network," in *International conference on medical image computing and computer-assisted intervention*, 2013, pp. 246-253.
 - [32] A. A. A. Setio, F. Ciompi, G. Litjens, P. Gerke, C. Jacobs, S. J. Van Riel, *et al.*, "Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks," *IEEE transactions on medical imaging*, vol. 35, pp. 1160-1169, 2016.
 - [33] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431-3440.
 - [34] W. Sun and R. Wang, "Fully convolutional networks for semantic segmentation of very high resolution remotely sensed images combined with DSM," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, pp. 474-478, 2018.
 - [35] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234-241.
 - [36] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *International conference on medical image computing and computer-assisted intervention*, 2016, pp. 424-432.
 - [37] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 fourth international conference on 3D vision (3DV)*, 2016, pp. 565-571.
 - [38] J. Kawahara, C. J. Brown, S. P. Miller, B. G. Booth, V. Chau, R. E. Grunau, *et al.*, "BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038-1049, 2017.
 - [39] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," in *International conference on information processing in medical imaging*, 2015, pp. 588-599.
 - [40] J. Kawahara and G. Hamarneh, "Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers," in *International workshop on machine learning in medical imaging*, 2016, pp. 164-171.
 - [41] X. Gao, S. Lin, and T. Y. Wong, "Automatic feature learning to grade nuclear cataracts based on deep learning," *IEEE Transactions on Biomedical Engineering*, vol. 62, pp. 2693-2701, 2015.
 - [42] D. Nie, H. Zhang, E. Adeli, L. Liu, and D. Shen, "3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients," in *International conference on medical image computing and computer-assisted intervention*, 2016, pp. 212-220.
 - [43] Q. Zhang, Y. Xiao, W. Dai, J. Suo, C. Wang, J. Shi, *et al.*, "Deep learning based classification of breast tumors with shear-wave elastography," *Ultrasonics*, vol. 72, pp. 150-157, 2016.

- 
- [44] G. van Tulder and M. de Bruijne, "Combining generative and discriminative representation learning for lung CT analysis with convolutional restricted Boltzmann machines," *IEEE transactions on medical imaging*, vol. 35, pp. 1262-1272, 2016.
 - [45] J.-Z. Cheng, D. Ni, Y.-H. Chou, J. Qin, C.-M. Tiu, Y.-C. Chang, *et al.*, "Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans," *Scientific reports*, vol. 6, pp. 1-13, 2016.
 - [46] M. Kallenberg, K. Petersen, M. Nielsen, A. Y. Ng, P. Diao, C. Igel, *et al.*, "Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring," *IEEE transactions on medical imaging*, vol. 35, pp. 1322-1331, 2016.
 - [47] A. Criminisi and J. Shotton, *Decision forests for computer vision and medical image analysis*: Springer Science & Business Media, 2013.
 - [48] K. D. Toennies, *Guide to medical image analysis*: Springer, 2017.
 - [49] N. Zhou, A. Fedorov, F. Fennessy, R. Kikinis, and Y. Gao, "Large scale digital prostate pathology image analysis combining feature extraction and deep neural network," *arXiv preprint arXiv:1705.02678*, 2017.
 - [50] M. Chen, X. Shi, Y. Zhang, D. Wu, and M. Guizani, "Deep features learning for medical image analysis with convolutional autoencoder neural network," *IEEE Transactions on Big Data*, 2017.
 - [51] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, pp. 2278-2324, 1998.
 - [52] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.
 - [53] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
 - [54] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
 - [55] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, pp. 102-127, 2019.
 - [56] G. Cheng, C. Yang, X. Yao, L. Guo, and J. Han, "When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs," *IEEE transactions on geoscience and remote sensing*, vol. 56, pp. 2811-2821, 2018.
 - [57] G. Cheng, J. Han, P. Zhou, and D. Xu, "Learning rotation-invariant and fisher discriminative convolutional neural networks for object detection," *IEEE Transactions on Image Processing*, vol. 28, pp. 265-278, 2018.
 - [58] Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, *et al.*, "Content-based brain tumor retrieval for MR images using transfer learning," *IEEE Access*, vol. 7, pp. 17809-17822, 2019.
 - [59] F. Abdolali, R. A. Zoroofi, Y. Otake, and Y. Sato, "A novel image-based retrieval system for characterization of maxillofacial lesions in cone beam CT images," *International journal of computer assisted radiology and surgery*, vol. 14, pp. 785-796, 2019.
 - [60] S. Ayyachamy, V. Alex, M. Khened, and G. Krishnamurthi, "Medical image retrieval using Resnet-18," in *Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications*, 2019, p. 1095410.
 - [61] Y. Anavi, I. Kogan, E. Gelbart, O. Geva, and H. Greenspan, "Visualizing and enhancing a deep learning framework using patients age and gender for chest x-ray image retrieval," in *Medical Imaging 2016: Computer-Aided Diagnosis*, 2016, p. 978510.
 - [62] A. Shah, S. Conjeti, N. Navab, and A. Katouzian, "Deeply learnt hashing forests for content based image retrieval in prostate MR images," in *Medical Imaging 2016: Image Processing*, 2016, p. 978414.
 - [63] J. Shi, X. Zheng, Y. Li, Q. Zhang, and S. Ying, "Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease," *IEEE journal of biomedical and health informatics*, vol. 22, pp. 173-183, 2017.
 - [64] A. Payan and G. Montana, "Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks," *arXiv preprint arXiv:1502.02506*, 2015.
 - [65] E. Hosseini-Asl, G. Gimel'farb, and A. El-Baz, "Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network," *arXiv preprint arXiv:1607.00556*, 2016.
 - [66] H.-I. Suk and D. Shen, "Deep ensemble sparse regression network for Alzheimer's disease diagnosis," in *International Workshop on Machine Learning in Medical Imaging*, 2016, pp. 113-121.

- 
- [67] N. Ali, C. Bolenz, T. Todenhöfer, A. Stenzel, P. Deetmar, M. Kriegmair, *et al.*, "Deep learning-based classification of blue light cystoscopy imaging during transurethral resection of bladder tumors," *Scientific reports*, vol. 11, pp. 1-10, 2021.
- [68] A. Ikeda, H. Nosato, Y. Kochi, T. Kojima, K. Kawai, H. Sakanashi, *et al.*, "Support system of cystoscopic diagnosis for bladder cancer based on artificial intelligence," *Journal of endourology*, vol. 34, pp. 352-358, 2020.
- [69] N. R. Freitas, P. M. Vieira, E. Lima, and C. S. Lima, "Automatic T1 bladder tumor detection by using wavelet analysis in cystoscopy images," *Physics in Medicine & Biology*, vol. 63, p. 035031, 2018.
- [70] S. Parveen, S. M. S. Ahmad, N. H. Abbas, W. A. W. Adnan, M. Hanafi, and N. Naeem, "Face liveness detection using dynamic local ternary pattern (DLTP)," *Computers*, vol. 5, p. 10, 2016.
- [71] M. Ringnér, "What is principal component analysis?," *Nature biotechnology*, vol. 26, pp. 303-304, 2008.
- [72] L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, ed: IGI global, 2010, pp. 242-264.
- [73] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, pp. 1345-1359, 2009.
- [74] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big data*, vol. 3, pp. 1-40, 2016.
- [75] E. Kozegar, "Cystoscopic Image Classification by an Ensemble of VGG-Nets," *International Journal of Nonlinear Analysis and Applications*, vol. 12, pp. 693-700, 2021.

Abstract

Cystoscopy imaging is highly recommended for the early diagnosis of bladder cancer, which is the ninth most common cancer worldwide. Unfortunately, few studies have focused on image-based bladder cancer diagnosis, which could be due to the lack of a standard dataset of cystoscopy images of bladder as well as variation in rotation, scale, and translation of the prepared images. In this regard, designing an intelligent system to extract discriminative features is essential for these kind of tasks.

In this thesis, two main methods have been proposed based on low-level and semantic features, respectively. In the first method, modified multi-layer perceptron based on adaptive learning rate and weight initialization based on Genetic Algorithm have been used for cystoscopic image classification. The proposed method improved the convergence speed and classification accuracy. For feature extraction, local binary patterns followed by principle component analysis have been applied on the original images. The experimental results demonstrated the outperformance of the first method compared to other conventional approaches based on low-level features. However, the obtained results were not proper for real world applications. In this regard, a novel approach for cystoscopic image classification based on convolutional neural network has been presented.

In the presented work, transfer learning is employed for extracting features from the collected images due to the limited number of training samples. The proposed method is robust to imaging variations. In fact, previously tuned weights by ImageNet are used in the deep learning phase. To avoid the curse of dimensionality and time-space complexity, the number of features have been decreased by feature reduction methods. Finally, an ensemble classifier has been proposed for cystoscopic image classification. In the proposed method, weighted majority vote strategy is adopted to combine a set of well-known classifiers. The proposed method achieved promising results in cystoscopic image classification. Afterwards, to improve the classification performance, regional segmentation in feature maps obtained from convolution layers has been used to select relevant areas. High level features and the related areas have been extracted from the Particle Swarm Optimization (PSO) and Otsu method. Then, the number of features has been reduced using PCA. At the end, an ensemble classifier including Gaussian Bayes and Support Vector Machine have been utilized based on majority vote.

The proposed method is evaluated on 720 cystoscopy images collected in a medical center in the Netherlands. Next, the suggested method is categorized into four different classes including bloody urine, benign masses, malignant masses, and normal cases. The results of the experiments indicated that the presented work achieved a precision of 72.5%, an accuracy of 70.5%, a recall of 71%, a specificity of 89.5%, and an F1-measure of 72.5% which outperformed other competing methods.

Keywords: Transfer learning, Cystoscopic image classification, high level semantic features, ensemble classifier, deep learning.



Shahrood University of Technology
Computer Engineering and IT Department

Classification of bladder cystoscopic images based on deep learning

A Thesis Submitted in Partial Fulfillment of the Requirement for the
Degree of Doctor of Philosophy in Computer Engineering

By:

Mohammad Reza Hashemi

Supervisor:

Prof. Hamid Hassanpour

Advisor:

Dr. Tao Tan

Dr. Ehsan Kozegar

January 2021