



دانشگاه تربیت مدرس

دانشگده مهندسی برق و کامپیوتر

سمینار کارشناسی ارشد

عنوان:

طبقه بندی گلبول های سفید خون و شناسایی بیماران سرطان خونی

White blood classification and identification patients with blood cancer

پژوهشگر:

شماره دانشجویی:

استاد راهنما:

تیرماه ۱۴۰۰



الحمد لله الذي
جعلنا من آل
إبراهيم وإسماعيل
عليهم السلام
فلا يهيننا
الذين آمنوا
بالحق
فلا يهيننا
الذين آمنوا
بالحق

چکیده

مشاهده و شناسایی سلول‌های خونی راهی مستقیم برای تشخیص به موقع بیماری‌های خونی است [۱]. سرطان خون نوعی پاتولوژی است که جوانان و بزرگسالان را درگیر می‌کند، باعث مرگ زودرس و چندین علائم دیگر می‌شود. می‌توان از سیستم‌های رایانه‌ای برای کاهش روش‌های درمان‌های نامناسب و کمک به متخصصان در تشخیص این بیماری کمک گرفت [۲]. در این میان، شمارش گلبول‌های سفید خون (WBC) یا لکوسیت‌ها) در خون محیطی یکی از موارد مهم است، زیرا می‌تواند به آسیب شناسان کمک کند تا بیماری‌هایی مانند سرطان خون و سایر بیماری‌های خونی را تشخیص دهند [۳]. در این سمینار، به مطالعه الگوریتم‌های مختلف در زمینه تشخیص بیماران مبتلا به سرطان خون، با استفاده از شناسایی و طبقه‌بندی لکوسیت‌ها پرداخته شده است.

کلمات کلیدی:

سرطان خون (لوسمی)، گلبول سفید (لکوسیت)، طبقه‌بندی، تشخیص، بخش‌بندی

فهرست مطالب

فهرست اشکال	أ.....
فهرست جداول	ب.....
فصل ۱ مقدمه	۱.....
۱-۱ مقدمه	۲.....
۲-۱ روش‌های تشخیص سرطان خون	۲.....
۳-۱ داده‌های مورد استفاده برای تشخیص	۳.....
فصل ۲ شناسایی و بخش‌بندی لکوسیت‌ها	۶.....
۱-۲ مقدمه	۷.....
۲-۲ معرفی روش‌های شناسایی و بخش‌بندی لکوسیت‌ها	۷.....
۱-۲-۲ بخش‌بندی معنایی	۸.....
۱-۱-۲-۲ ساختار DeepLab	۹.....
۲-۱-۲-۲ استراتژی‌های آموزش و الگوریتم پیشنهادی	۱۰.....
۳-۱-۲-۲ طبقه‌بندی	۱۲.....
۴-۱-۲-۲ نتایج	۱۷.....
۲-۲-۲ بخش‌بندی مبتنی بر الگوریتم k-means	۱۸.....
۳-۲-۲ شناسایی و طبقه‌بندی لکوسیت‌ها با استفاده از شبکه عصبی کانولوشنی	۱۹.....
۱-۳-۲-۲ تشخیص WBCها	۲۰.....
۲-۳-۲-۲ طبقه‌بندی WBCها	۲۳.....
۴-۲-۲ بخش‌بندی هسته و سیتوپلاسم سلول با استفاده از تصاویر میکروسکوپی	۲۴.....
۱-۴-۲-۲ استخراج هسته	۲۴.....
۵-۲-۲ بخش‌بندی هسته و سیتوپلاسم سلول با استفاده از تصاویر ابر طیفی	۲۵.....
۱-۵-۲-۲ بخش‌بندی سلول	۲۶.....
۲-۵-۲-۲ بخش‌بندی هسته‌ها و سیتوپلاسم	۲۶.....
۳-۲ نتیجه‌گیری	۲۷.....
فصل ۳ تشخیص سرطان خون (لوسمی) از تصاویر گلبول‌های سفید	۲۹.....
۱-۳ مقدمه	۳۰.....
۲-۳ روش‌های تشخیص با استفاده از بخش‌بندی	۳۰.....
۱-۲-۳ طبقه‌بندی سرطان خون با استفاده از مورفولوژی مبتنی بر دانش پزشکی و نشانگر CD	۳۰.....

۳۲.....	۱-۱-۲-۳ استخراج ویژگی
۳۲.....	۲-۱-۲-۳ طبقه‌بندی
۳-۲-۲	سیستم پشتیبانی تصمیم برای طبقه‌بندی سرطان خون حاد بر اساس تصاویر میکروسکوپی
۳۳.....	دیجیتال
۳۳.....	۱-۲-۲-۳ استخراج ویژگی
۳۴.....	۲-۲-۲-۳ درخت تصمیم
۳۴.....	۳-۲-۲-۳ شبکه عصبی مصنوعی (ANN)
۳-۳	روش‌های تشخیص بدون استفاده از بخش‌بندی
۳۵.....	۱-۳-۳ شبکه‌های عصبی کانولوشنی
۳۶.....	۲-۳-۳ استخراج ویژگی
۳۶.....	۳-۳-۳ طبقه‌بندی
۳۶.....	۴-۳ نتیجه‌گیری
۴۰.....	فصل ۴
۴۰.....	جمع‌بندی و ارائه پیشنهادات

فهرست اشکال

- شکل ۱: انواع گلبول‌های سفید خون..... ۳
- شکل ۲: الگوریتم روش طبقه‌بندی مبتنی بر روش یادگیری انتقالی..... ۹
- شکل ۳: ساختار DeepLabv3+..... ۹
- شکل ۴: انواع گلبول‌های سفید خون..... ۱۱
- شکل ۵: ساختار AlexNet..... ۱۴
- شکل ۶: نقشه تصمیم برای تنظیم شبکه از قبل آموزش دیده روش مبتنی بر یادگیری انتقالی..... ۱۴
- شکل ۷: فرآیند آموزش روش مبتنی بر یادگیری انتقالی..... ۱۵
- شکل ۸: ماتریس درهم ریختگی مسئله پنج کلاسه در روش مبتنی بر یادگیری انتقالی..... ۱۶
- شکل ۹: فلوجارت روش شناسایی و طبقه‌بندی لکوسیت‌ها با استفاده از شبکه عصبی کانولوشنی..... ۲۰
- شکل ۱۰: مقایسه تغییر رنگ روش پیشنهادی [۳] با برخی از روش‌های دیگر برای شناسایی ابتدایی WBC..... ۲۱
- شکل ۱۱: ساختار بتونی CNN در روش پیشنهادی برای استخراج ویژگی‌ها در سطح بالا با استفاده از ۵ لایه کانولوشن و ۲ لایه پولینگ..... ۲۴
- شکل ۱۲: انتخاب سلول‌های خون سالم..... ۲۵
- شکل ۱۳: هیستوگرام هسته و سیتوپلاسم در سلول خونی..... ۲۵
- شکل ۱۴: فلوجارت فرآیند روش بخش‌بندی هسته و سیتوپلاسم سلول با استفاده از تصاویر ابر طیفی..... ۲۷
- شکل ۱۵: فلوجارت روش طبقه‌بندی سرطان خون با استفاده از مورفولوژی مبتنی بر دانش پزشکی و نشانگر CD..... ۳۱
- شکل ۱۶: طبقه‌بندی انواع ALL و AML..... ۳۲
- شکل ۱۷: فلوجارت روش تشخیص سرطان خون با استفاده از CNN..... ۳۵
- شکل ۱۸: توضیح ساده از ساختار CNN..... ۳۵

فهرست جداول

جدول ۱ : انگیزه روش‌های طبقه‌بندی و تکنیک‌های بکار رفته در مطالعات [۲۶]..... ۳۷

جدول ۲ : دسته‌بندی چالش‌های طبقه‌بندی و تشخیص سرطان خون در مطالعات [۲۶]..... ۳۸

فصل ۱

مقدمه

۱-۱ مقدمه

سرطان خون به سرطان در سلول‌های خونی اشاره دارد و مغز استخوان، سلول‌های سرطان خون (لوسمی) را تولید می‌کند که در نتیجه آن سلول‌های سفید خون، غیر طبیعی می‌شوند [۴]. سرطان خون حاد باعث وخامت سریع بیمار می‌شود، در حالی که سرطان خون مزمن با پیشرفت آهسته مشخص می‌شود و ممکن است لنفوسیتی یا میلوژنی باشد [۵]. علاوه بر این زیرگروه‌ها، سرطان خون از نظر پزشکی به چهار نوع تقسیم می‌شود، لوسمی حاد میلوژن (AML)، لوسمی لنفوبلاستیک حاد (ALL)، لوسمی میلوئید مزمن (CML) و لوسمی لنفوسیتی مزمن (CLL) [۵]. در میان چهار نوع سرطان خون، AML و ALL کودکان خردسال را تحت تأثیر قرار می‌دهد، و برای درمان به موقع بیمار نیاز به تشخیص فوری، به ویژه در مراحل اولیه وجود دارد [۶].

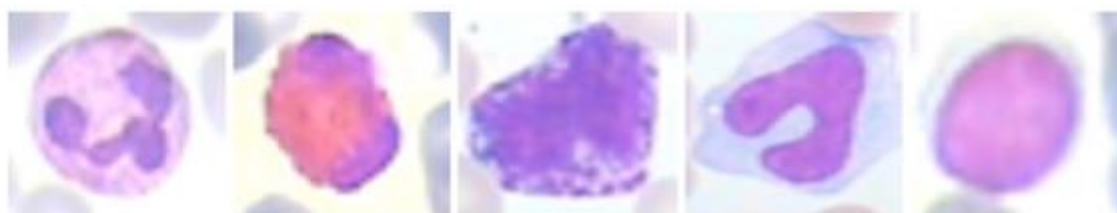
۲-۱ روش‌های تشخیص سرطان خون

در تشخیص پزشکی، تکنیک‌های پردازش تصویر برای شناسایی و تحلیل سلول‌های خون به منظور شناسایی اختلالات سلول‌های خونی مورد استفاده قرار می‌گیرند. سیستم‌های متداول تشخیص سلول‌های خونی علاوه بر اینکه وقت گیر هستند، به متخصصان پزشکی نیز نیاز دارند و صحت آنها به توانایی اپراتور بستگی دارد. همین امر، محققان را برای ایجاد یک تکنیک اتوماتیک رایانه‌ای برای تحلیل سلول‌های خونی برانگیخته است، که در آن از ریزنگارهای (میکروگراف‌های) خون شامل گلبول‌های سفید خون (WBCs)، گلبول‌های قرمز خون (RBCs) و پلاکت‌ها استفاده می‌شود. آنالیز سلول‌های خون با استفاده از ریزنگارهای خون این مزیت را دارد که ارزان هستند و به تجهیزات گران قیمت احتیاج ندارند. بسیاری از اختلالات سلول‌های خون مانند عفونت، واکنش آلرژیک، التهاب، سرطان‌های خون مانند لوسمی و لنفوم را می‌توان از طریق آنالیز گلبول‌های سفید خون تشخیص داد [۷].

لکوسیت‌ها که به گلبول‌های سفید خون نیز معروف هستند، به بدن در مبارزه با عفونت‌ها و بیماری‌ها کمک می‌کنند [۷]. گلبول‌های سفید که شامل نوتروفیل‌ها، مونوسیت‌ها، ایوزینوفیل‌ها، بازوفیل‌ها و لنفوسیت‌ها هستند [۴]، به عنوان شاخص در تشخیص سرطان خون که یکی از مهمترین بیماری‌ها است، نقش مهمی دارند [۶].

تعداد گلبول‌های سفید در مقایسه با گلبول‌های قرمز کمتر است و دارای هسته هستند و همچنین اندازه آن‌ها بزرگتر است. به طور کلی، آن‌ها در دو گروه لکوسیت‌های دانه‌دار و لکوسیت‌های بدون دانه دسته‌بندی می‌شوند [۷].

لکوسیت‌های دانه‌دار، دارای دانه در سیتوپلاسم هستند و در گروه‌های بازوفیل‌ها، ایوزینوفیل‌ها و نوتروفیل‌ها دسته‌بندی می‌شوند؛ در حالی که لکوسیت‌های بی‌دانه شامل مونوسیت‌ها و لنفوسیت‌ها هستند و درصد فراوانی آن‌ها در خون متفاوت است [۸]. شکل ۱ انواع لکوسیت‌ها را نشان می‌دهد. از این شکل می‌توان دریافت که نوتروفیل‌ها دارای یک هسته چند لوبی هستند، در حالی که ایوزینوفیل‌ها دارای یک هسته دو لوب هستند. بازوفیل‌ها دارای یک هسته دو لوب یا سه لوب هستند، مونوسیت‌ها دارای یک هسته به شکل کلیه و همچنین هسته یک لنفوسیت غیرعادی و بسیار رنگی است. انواع لکوسیت‌ها و تعداد آن‌ها برای تشخیص بیماری‌های مرتبط با خون متناسب است [۹]. در آزمایشگاه، از ماشین‌های خودکار به منظور ارزیابی خصوصیات کمی و کیفی لکوسیت‌ها استفاده می‌شود، اما برای پیدا کردن ناهنجاری‌های مورفولوژیکی از اطمینان کافی برخوردار نیستند [۱۰].



نوتروفیل

ایوزینوفیل

بازوفیل

مونوسیت

لنفوسیت

شکل ۱: انواع گلبول‌های سفید خون

۳-۱ داده‌های مورد استفاده برای تشخیص

یکی از چالش‌های ایجاد یک سیستم کمک پزشکی، توانایی تشخیص بیماری در پایگاه داده‌های با ویژگی‌های متمایز است [۲]. چندین نویسنده، روش‌های خود را فقط در پایگاه داده‌های همگن یا پایگاه داده‌های اختصاصی روش خودشان، آزمایش کرده‌اند. پایگاه داده‌هایی که در این دسته از مطالعات بکار گرفته شده‌اند، حاوی تصاویر میکروسکوپی خونی با اندازه‌های مختلف هستند. برخی مطالعات از یک

پایگاه داده و برخی نیز از ترکیب چندین پایگاه داده به منظور ایجاد پایگاه داده جدید استفاده نموده‌اند. چند نمونه از این پایگاه داده‌ها در زیر معرفی می‌شوند.

پایگاه داده در دسترس عموم LISC که شامل تصاویر با ۵۷۶*۷۲۰ پیکسل است، متشکل از تصاویر هماتولوژیکی گرفته شده از لکه خون محیطی است که توسط روش Gismo-Right رنگی شده و توسط یک دوربین دیجیتال متصل به میکروسکوپ گرفته شده است که از یک لنز آکروماتیک با بزرگنمایی ۱۰۰ استفاده می‌کند و تمام تصاویر در فرمت BMP ذخیره شده‌اند. این پایگاه داده شامل ۵۳ تصویر بازوفیل، ۵۰ تصویر نوتروفیل، ۳۹ تصویر ایوزینوفیل، ۵۲ تصویر لنفوسیت و ۴۸ تصویر مونوسیت است [۷].

پایگاه داده ALL-IDB [۱۱]، یک پایگاه داده تصاویر عمومی است که متشکل از نمونه‌های خون محیطی است که توسط برخی از متخصصان از برخی افراد عادی و بیماران سرطان خون مبتلا به سرطان خون کودکان و بیماری‌های خون جمع‌آوری شده است که شامل دو نسخه مجزا است. یکی شامل ۱۰۸ تصویر با فرمت JPG با عمق رنگ ۲۴ بیتی که بیشتر آنها با میکروسکوپ نوری در بزرگنمایی‌های مختلف از ۳۰۰ تا ۵۰۰ و دوربین Canon PowerShot G5 با وضوح ۲۵۹۲ × ۱۹۴۴ گرفته شده است. تصاویر دیگر با میکروسکوپ در یک بزرگنمایی ثابت و یک دوربین Olympus C2500L با وضوح ۱۷۱۲ × ۱۳۶۸ گرفته شده است [۳].

در یکی از مطالعات، سه پایگاه داده تصویر ناهمگن بر اساس پایگاه‌های داده سنتی ارائه شده در مطبوعات، ساخته شده است. دو تا از این پایگاه داده‌ها با توجه به تعداد لکوسیت‌های موجود در تصاویر تقسیم شده‌اند. پایگاه داده Hybrid-Leukocyte از چهار پایگاه داده تشکیل شده است که فقط یک لکوسیت در هر تصویر ارائه می‌کند: ALL-IDB1 [۱۱]، ALL-IDB2 [۱۱]، Leukocytes [۱۲] و CellaVision [۱۳]. دومین پایگاه داده ساخته شده پایگاه داده Hybrid-Slide است. این پایگاه داده حاوی لکوسیت‌های زیادی برای هر تصویر است و همچنین از چهار پایگاه داده سنتی یعنی ALL-IDB1، Atlas و دو پایگاه داده پیشنهادی در کارهای صراف زاده و همکارانش [۱۲، ۱۴] تشکیل شده است. پایگاه داده سوم Hybrid-Complete نام دارد که مرکب از دو پایگاه داده هیبریدی/ترکیبی است که قبلاً ارائه شدند [۲].

در یکی از پایگاه داده‌ها نیز، ۱۱۵ عکس میکروسکوپ نوری دیجیتال به ابعاد ۶۳۲ * ۴۸۰ پیکسل گردآوری شده‌اند. داده‌ها با استفاده از میکروسکوپ نوری پشتیبانی شده توسط دوربین CCD جمع‌آوری

شده‌اند. ششصد و بیست و چهار تصویر با ابعاد ۴۸۰ * ۶۳۲ توسط یک متخصص در سه دسته طبقه بندی شده‌اند [۵، ۱۵].

پایگاه داده دیگری که مورد استفاده قرار گرفته است شامل ۵۰۰ تصویر از سلول‌های خونی محتوی ۲۰۰ سلول سالم، ۱۵۰ ALL و ۱۵۰ سلول AML است که در بیمارستان سرطان Ubonratchathani و بیمارستان بین المللی Sunpasit در تایلند تهیه شده است [۶].

برخی مطالعات نیز از مجموعه تصاویر حاصل از تصویربرداری ابرطیفی استفاده نموده‌اند. این روش یک فناوری نوظهور است که مزایای تصویربرداری طیفی و تکنیک‌های تصویربرداری سنتی را با هم ترکیب می‌کند. هر تصویر تک باند تولید شده از طریق سیستم تصویربرداری ابرطیفی مربوط به یک طول موج باند باریک خاص است که می‌توان همزمان از آن اطلاعات مکانی و طیفی را بدست آورد. در نتیجه تصویر ابرطیفی از طریق تجسم طیفی دارای اطلاعات غنی‌تری نسبت به تصویر خاکستری یا رنگی است [۱].

فصل ۲

شناسایی و بخش‌بندی لکوسیت‌ها

به طور کلی، یک سیستم تشخیص اتوماتیک WBC عمدتاً از سه مرحله اصلی تشکیل شده است: (۱) تشخیص WBCها از تصویر خون محیطی، (۲) استخراج ویژگی‌های موثر، (۳) طراحی طبقه‌بندی [۱۶]. به این معنی که ابتدا WBCها را از تصویر خون محیطی تشخیص داده شده و سپس ویژگی‌های موثر WBCها برای طبقه‌بندی استخراج می‌شوند. تا حدی، یک روش تشخیص خوب برای شناسایی صحیح WBC از پس زمینه آن‌ها، اولین قدم برای موفقیت است [۳]. در این فصل به بررسی انواع روش‌های شناسایی و بخش‌بندی لکوسیت‌ها یا WBCها پرداخته شده است.

۲-۲ معرفی روش‌های شناسایی و بخش‌بندی لکوسیت‌ها

برای شناسایی انواع مختلف لکوسیت‌ها از تصویر لکه (اسمیر) خون، محققان مختلف الگوریتم‌های مختلفی ساخته‌اند. بسیاری از محققان، سیستم‌های با کمک رایانه را برای آنالیز سلول‌ها به واسطه تصاویر میکروسکوپی ساخته‌اند. برخی از کارها، الگوریتم‌هایی را برای بخش‌بندی صرفاً هسته سلول توسعه داده‌اند، در حالی که برخی دیگر شامل تقسیم هسته و سیتوپلاسم هستند. تئرا آمپون و همکارانش [۱۷]، یک سیستم طبقه‌بندی شش کلاسه گلبول‌های سفید مغز استخوان را با استفاده از طبقه‌بند بیز و طبقه‌بند شبکه عصبی مصنوعی (ANN) پیشنهاد داده‌اند. آن‌ها از گلبول‌های سفیدی که به صورت دستی بخش‌بندی شده‌اند، ویژگی‌هایی را استخراج کرده و نشان داده‌اند که برای دستیابی به صحت بهتر، فقط ویژگی‌های هسته، ضروری هستند. زاده و همکارانش [۱۸]، الگوریتمی برای شناسایی ۵ نوع گلبول سفید در خون محیطی ارائه داده‌اند. آن‌ها برای بخش‌بندی هسته و سیتوپلاسم گلبول‌های سفید خون از متعامدسازی گرام-اشمیت همراه با الگوریتم مار(اسنیک) استفاده کرده‌اند. در مطالعه انجام شده توسط حبیب‌زاده و همکارانش [۱۹]، بخش‌بندی گلبول سفید با استفاده از یک رویکرد تکراری انجام شده که از ماسک دایره‌ای، تخمین اندازه RBC و فیلترینگ استفاده کرده است. آن‌ها از ۱۴۰ تصویر لکه خونی پنج دسته گلبول سفید استفاده کرده و به یک مقدار صحت متوسط دست یافته‌اند. لیایو و دنگ، اولین افرادی بودند که برای بخش‌بندی لکوسیت‌های خون روش آنالیز شکل را معرفی کرده‌اند [۲۰]. آن‌ها بعد از بخش‌بندی اولیه بر اساس آستانه گذاری ساده، محیط لکوسیت‌ها را توسط آنالیز شکل مشخص کرده‌اند. این الگوریتم علیرغم سادگی و اثر بخشی اولیه، تنها جهت بخش‌بندی لکوسیت‌های کروی شکل، نظیر لنفوسیت‌ها مناسب است. مطالعه دیگری جهت بخش‌بندی نمودن سلول‌ها از بزرگی شیب، آستانه

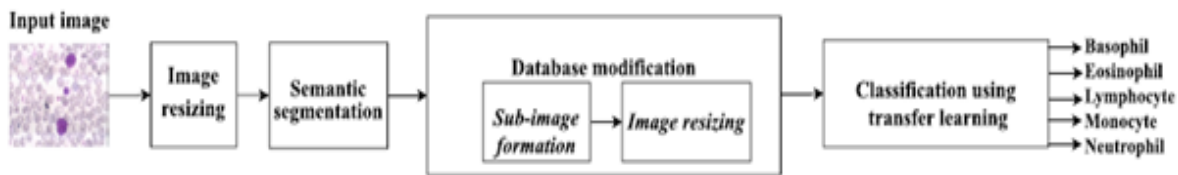
گذاری، عملیات مورفولوژیکی و تبدیل حوضه استفاده نموده است [۲۱]. برای آزمایش این روش از ۵۰ تصویر از زیرگروه‌های M2، M5 و M6 استفاده شده و نتیجه نشان داده است که این روش توانسته نتایج بخش‌بندی کیفی خوبی بدست آورد. هدف از این بخش‌بندی در این مطالعه، یافتن تصاویر فرعی بوده است که فقط حاوی یک سلول لکوسیت یا بلاست باشد و پس زمینه را حذف کند.

روش‌های دیگری نیز در راستای شناسایی و بخش‌بندی لکوسیت‌ها استفاده شده‌اند که عبارتند از: الگوریتم مخروط محدب حداکثر زاویه‌ای متوالی [۱]، الگوریتم خفاش باینری بهینه شده (OBBA) [۲۲]، بخش‌بندی معنایی [۷]، خوشه بندی [۲۳]، شبکه عصبی [۲۴]، مدل‌های تبدیل فضای رنگ [۲۵]، آستانه گذاری، عملگر مورفولوژیک، تشخیص لبه، منطقه در حال رشد، رنگ‌ها، روش مبتنی بر بهینه سازی، روش فازی و ماشین بردار پشتیبان. در ادامه برخی از این روش‌ها به تفصیل توضیح داده شده‌اند.

۲-۱-۲ بخش‌بندی معنایی

بخش‌بندی معنایی فرآیندی است که هر پیکسل تصویر را در دسته متفاوت، گروه‌بندی می‌کند. به آن "طبقه‌بندی پیکسل - سطح" نیز گفته می‌شود. در یکی از کارهای انجام شده، بخش‌بندی معنایی به عنوان یک مسئله طبقه‌بندی باینری تنظیم شده که هر پیکسل یا به عنوان WBC و یا پس زمینه، دسته‌بندی می‌شود. این، نوعی پیش‌بینی متراکم است که می‌توان آن را با استفاده از یک شبکه عصبی با چندین لایه کانولوشن با لایه یکسان، اجرا کرد. اما این نوع پیاده‌سازی از پیچیدگی محاسباتی بالایی برای حفظ وضوح کامل در تمام فرایندها برخوردار است. بنابراین، رویکرد رایجی که مورد استفاده قرار می‌گیرد، ساختار رمزگذار-رمزگشا (انکودر-دیکودر) است که عملیات نمونه‌برداری پایین در بخش رمزگذار و عملیات نمونه‌برداری بالا در بخش رمزگشا است. در این کار از ساختار DeepLab برای اجرای بخش‌بندی معنایی جهت بخش‌بندی WBC ها از تصاویر لکه خون استفاده شده است [۷].

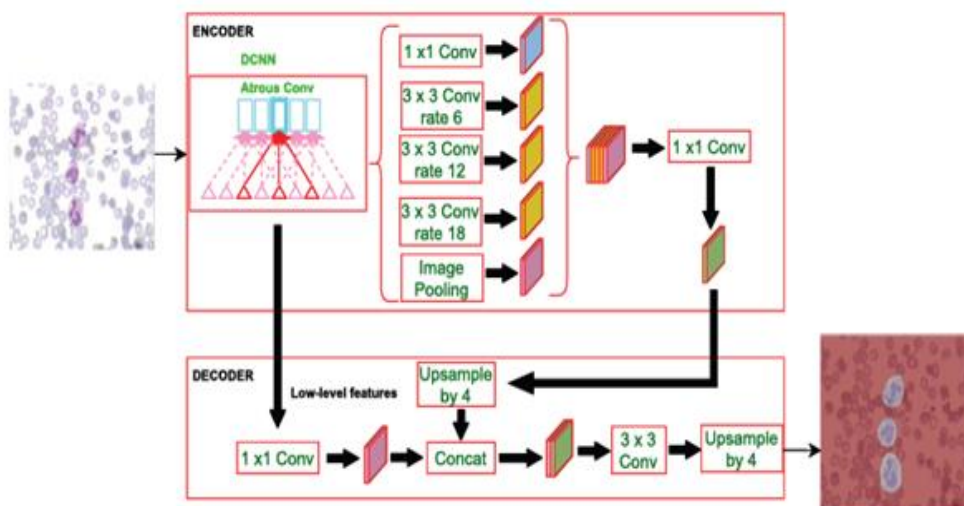
سهم اصلی این کار این است که با استفاده از بخش‌بندی معنایی، WBC ها را از تصویر لکه خون متمایز کند تا WBC ها را با مقطع مشترک بالاتر از اجتماع و دقت میانگین متوسط مکان‌یابی کند که این نیز منجر به طبقه‌بندی بهتر می‌شود. در ادامه به طور مفصل در مورد این موارد بحث شده است. الگوریتم روش پیشنهادی در شکل ۲ نشان داده شده است [۷].



شکل ۲: الگوریتم روش طبقه بندی مبتنی بر روش یادگیری انتقالی

۲-۱-۲ ساختار DeepLab

DeepLab، یک مدل یادگیری عمیق برای بخش بندی معنایی متشکل از دو بخش رمزگذاری و رمزگشایی است. در بخش رمزگذاری از CNN از قبل آموزش دیده جهت استخراج اطلاعات از تصاویر استفاده شده است. با استفاده از این اطلاعات، رمزگشا خروجی را با همان اندازه ورودی، بازسازی می کند. بیشتر شبکه های رمزگشا از نمونه برداری افزایشی (فرا نمونه برداری) دو خطی استفاده می کنند، اما این عیب را دارد که قادر به بازسازی دقیق تر جزئیات نیستند. در این روش پیشنهادی، از DeepLabv3+ که یک نسخه از DeepLab است، مبتنی بر ResNet 50، به عنوان مدل بخش بندی استفاده شده است. ساختار DeepLabv3+ در شکل ۳ نشان داده شده است. در بخش رمزگذار، اطلاعات دقیق، با استفاده از کانولوشن (پیچیدگی) های تفکیک پذیر آتروس بدست می آید. در بخش رمزگشای، یک کانولوشن 1×1 به ویژگی های سطح پایین استخراج شده در مرحله رمزگذاری اعمال می شود و نتیجه با نسخه ۴ عاملی فرا نمونه برداری شده ویژگی های رمزگذاری شده، الحاق می شود. سپس چند کانولوشن 3×3 اعمال شده و عملیات فرا نمونه برداری ۴ عاملی انجام می شود تا اندازه تصویر خروجی همانند ورودی شود.



شکل ۳: ساختار DeepLabv3+

برای به دست آوردن ویژگی‌های متراکم و کاهش پیچیدگی محاسباتی از کانولوشن تفکیک‌پذیر آتروس استفاده می‌شود چون کانولوشن آتروس در کانولوشن عمقی اتخاذ می‌شود. کانولوشن آتروس که به عنوان "کانولوشن متسع شده/ منبسط شده" نیز شناخته می‌شود، از معادله زیر بدست می‌آید:

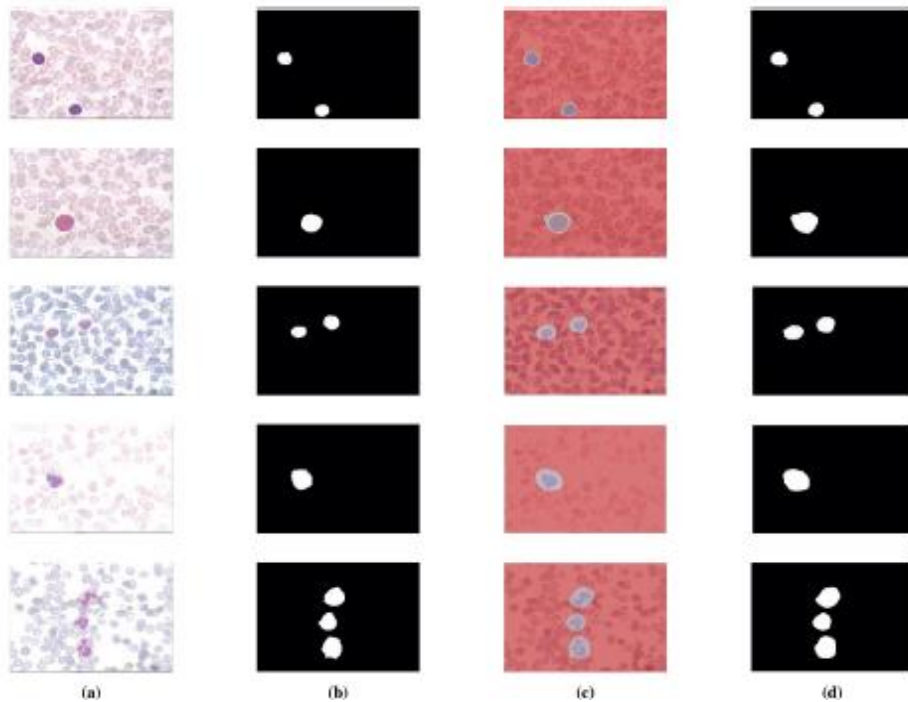
$$z(k) = \sum_j x(k + r_t * j) c(j)$$

در عبارت داده شده، خروجی z از طریق مجموع ورودی x با هسته با قرار دادن صفر $1 - r_t$ بین دو ضریب مجاور بدست می‌آید. کانولوشن آتروس از پارامتری به نام "نرخ منبسط" استفاده می‌کند که فاصله بین مقادیر در هسته را توصیف می‌کند. این بدان معناست که می‌توان میدان پذیرنده فیلتر را با تغییر نرخ r_t ، تغییر داد. کانولوشن آتروس با $r_t = 1$ همان کانولوشن استاندارد است.

این، کانولوشن فضایی را به طور مستقل برای هر کانال ورودی با نرخ متفاوتی انجام می‌دهد و برای ترکیب خروجی از کانولوشن نقطه‌ای (نقطه به نقطه) استفاده می‌کند.

۲-۱-۲-۲ استراتژی‌های آموزش و الگوریتم پیشنهادی

شبکه طراحی شده، با استفاده از نزول گرادیان تصادفی با مقدار حرکت، الگوریتم بهینه سازی معروف با همگرایی سریع آموزش داده شده است. نرخ یادگیری اولیه ۱۰ به توان منفی ۳ است و سیستم به تدریج با نرخ یادگیری ۰/۳ کاهش داده شده است. حداکثر تعداد دوره ۳۰ قرار داده شده و آن با فاکتور ۱۰ کاهش یافته است. ضریب حرکت ۰/۹ تنظیم شده است، که زمان آموزش پس‌انتشار را کنترل می‌کند. برای جلوگیری از بیش برآزش، آموزش با استفاده از تنظیم L_2 با وزنهایی که ۰/۰۰۰۵ آغاز شده، انجام گردیده است. برای بخش‌بندی، تصاویر ورودی در جهت چپ و راست منعکس شده و سپس در دو جهت افقی و عمودی تا ۱۰ پیکسل منتقل گردیده و تا ۲۰ درجه برای ساخت تصاویر تکمیل شده چرخانده شده‌اند. مراحل ویژه مکان‌یابی WBC در استراتژی‌های آموزش الگوریتم و الگوریتم پیشنهادی نشان داده شده است و در قسمت‌های a تا d شکل ۴ نتایج بخش‌بندی WBC دیده می‌شود.



شکل ۴: انواع گلبول‌های سفید خون

الگوریتم ۱. الگوریتم بخش‌بندی معنایی WBC

ورودی: تصویر لکه خون در قالب RGB و تصاویر واقعی زمینه در قالب سیاه و سفید.
خروجی: تصویر ماسک.

۱. سایزبندی دوباره تصاویر در سایز 224×224
۲. ایجاد یک مخزن داده تصاویر برای گزینش تصاویر RGB از پایگاه داده
۳. تنظیم تعداد کلاس‌ها به ۲
۴. ایجاد مخزن داده برچسب (لیبل) با برچسب‌گذاری هر پیکسل در تصویر ورودی به عنوان "WBC" و "پس زمینه" بر اساس تصاویر واقعی زمینه
۵. ایجاد لایه‌هایی برای شبکه $DeepLabv3+$ با استفاده از شبکه از پیش آموزش دیده ResNet50
۶. جایگزین کردن لایه طبقه‌بندی در شبکه $DeepLabv3+$ با یک لایه خروجی طبقه‌بندی پیکسلی که وزن آن با وزن کلاس فرکانس متوسط شروع می‌شود.
۷. تقسیم تصاویر مخزن داده به ۱۰ گروه
۸. انجام مراحل ۹ تا ۱۴ برای ۱۰ تکرار به وسیله تغییر داده تست و آموزش در هر تکرار
۹. انتخاب یک گروه برای اداف آزمایش/تست و مابقی گروه‌ها برای اهداف آموزش

۱۰. تقویت نمودن تصاویر داده آموزش با استفاده از چرخش، انتقال و مقیاس گذاری

۱۱. تنظیم عملیات آموزش و آموزش شبکه

۱۲. بدست آوردن تصویر بخش بخش شده با استفاده از تصاویر آزمایش تغذیه شده

۱۳. باینری کردن نتایج بخش بخش شده

۱۴. پر کردن سوراخ های تصویر باینری و انجام عمل گشایش/افتتاح ناحیه برای حذف اشیای

کوچک، در صورت وجود

الگوریتم ۲. الگوریتم برداشتن/بریدن WBC

ورودی: تصویر ماسک.

خروجی: تصاویر فرعی لوکوسیت.

۱. تعیین موقعیت اولیه کادر محدودکننده: موقعیت شروع (x, y) ، عرض و ارتفاع کادر. برای هر

شبی، $x = \text{Minimumvalueofrow} - 0.5$ ، $y = \text{Minimumvalueofcolumn} - 0.5$ ، عرض

برابر $1 + \text{min.valueofrow} - \text{max.valueofrow}$ و ارتفاع برابر $\text{max.valueofcolumn} -$

$\text{min.valueofcolumn} + 1$.

۲. برش WBCها از تصویر لکه خون با استفاده از کادرهای محدودکننده و سایزبندی مجدد تصاویر

فرعی با سایز $3 * 227 * 227$

۳-۱-۲-۲ طبقه بندی

تمام الگوریتم های طبقه بندی براساس کاربردها و در دسترس بودن مجموعه داده ها انتخاب می-

شوند. بیشتر طبقه بندی های مرسوم استفاده شده مانند SVM، جنگل تصادفی، LDA و غیره از بردارهای

ویژگی استخراج شده از یک مجموعه داده آموزشی برای پیش بینی کلاس نمونه های آزمون استفاده می-

کنند. با استفاده از شبکه های عصبی عمیق، می توان ویژگی های سطح بالا را مستقیماً از تصاویر

استخراج کرد که منجر به صحت بالاتر در هزینه پیچیدگی محاسباتی می شود. در کار پیشنهادی، طبقه

بندی با استفاده از یادگیری انتقالی انجام شده که از دانش یک شبکه از قبل آموزش دیده برای حل

مسئله فعلی استفاده می کند.

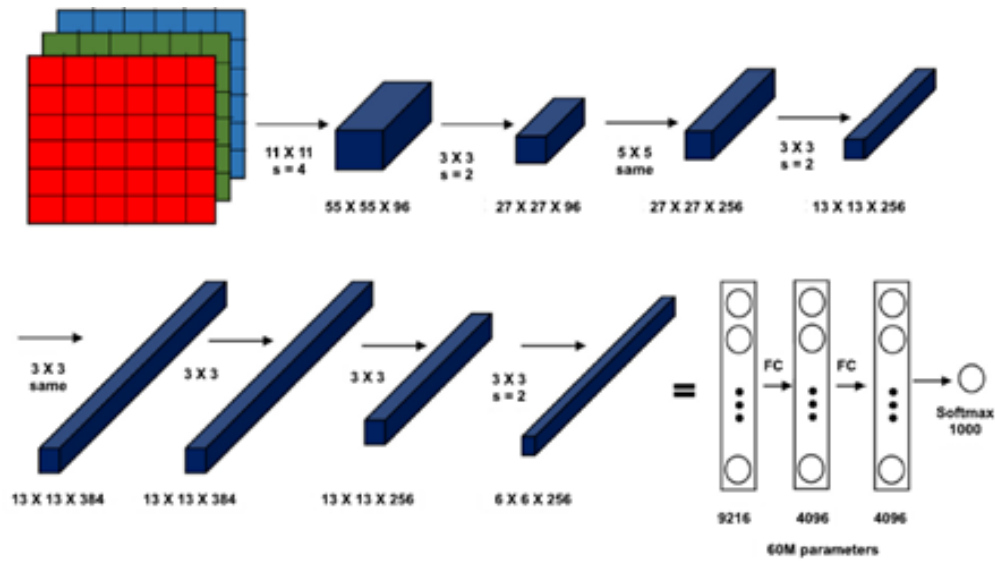
۳-۱-۲-۱-۲ یادگیری انتقالی با استفاده از AlexNet

فرآیند یادگیری انتقالی شامل انتخاب یک مدل از قبل آموزش دیده و انطباق مدل با مسئله مورد

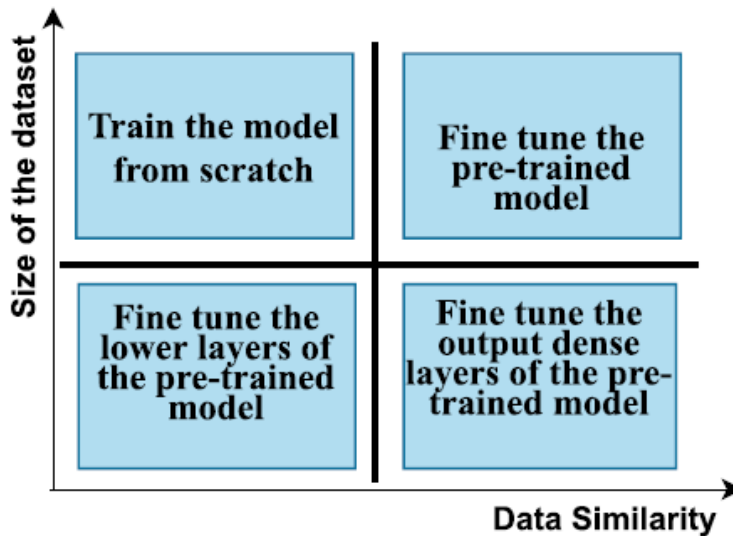
نظر است. در اینجا، AlexNet به عنوان شبکه عصبی کانولوشنی از قبل آموزش دیده برای طبقه بندی

انتخاب شده، زیرا از ReLU (واحد خطی اصلاح شده) به عنوان تابع فعال سازی استفاده می کند، بنابراین

می‌توان آن را بسیار سریع‌تر آموزش داد؛ شبکه دارای ۸ لایه عمیق است، سبب کاهش پیچیدگی زمان اجرا می‌شود. شبکه دارای پنج لایه کانولوشنی، دو لایه کاملاً متصل و یک لایه softmax است. ساختار آن در شکل ۵ نشان داده شده است. از این ۵ لایه کانولوشنی، ۳ لایه (لایه ۱، لایه ۲ و لایه ۵) با حداکثر لایه ادغام دنبال می‌شوند. AlexNet تصاویر ورودی با فرمت RGB را می‌پذیرد و سائز آن باید $۲۲۷*۲۲۷*۳$ باشد. لایه های کانولوشن از فیلترهای ۱۱×۱۱ با گام ۴ استفاده می‌کنند و حداکثر ادغام از فیلترهای ۳×۳ با گام ۲ استفاده می‌کند. در شبکه از قبل آموزش دیده AlexNet سه لایه آخر برای طبقه بندی تصاویر در ۱۰۰۰ کلاس پیکربندی شده‌اند. شکل ۶ نحوه پیکربندی لایه‌ها را در شبکه براساس اندازه مجموعه داده موجود نشان می‌دهد. وقتی هم، اندازه مجموعه داده و هم، شباهت داده‌ها کم باشد، آموزش مجدد و سفارشی سازی لایه های بالاتر با توجه به مجموعه داده جدید، مهم است. اندازه محدود مجموعه داده از طریق حفظ لایه های اولیه به صورت از قبل آموزش دیده متعادل شده و وزن آن لایه‌ها ثابت می‌شود. اگر اندازه مجموعه داده کوچک باشد و شباهت داده زیاد است، نیازی به آموزش مجدد مدل نیست، فقط لایه‌های متراکم و لایه سافت مکس نهایی جهت تولید n دسته خروجی، تغییر داده می‌شوند. اگر اندازه مجموعه داده بزرگ و شباهت داده‌ها کم باشد، پیش بینی با استفاده از مدل های از قبل آموزش دیده کارا نخواهد بود زیرا داده ها با داده هایی که مدل روی آنها از قبل آموزش دیده است، بسیار متفاوت است. بر اساس داده های جدید، آموزش شبکه عصبی از ابتدا امکان پذیر است، زیرا داده های با اندازه بزرگ در شبکه عصبی عملکرد بهتری دارند. حالت ایده آل این است که هم مجموعه داده‌ها و هم شباهت داده‌ها بسیار زیاد باشد و سپس با استفاده از وزنه‌هایی که در مدل از قبل آموزش دیده شده، مجدداً مدل آموزش ببیند.



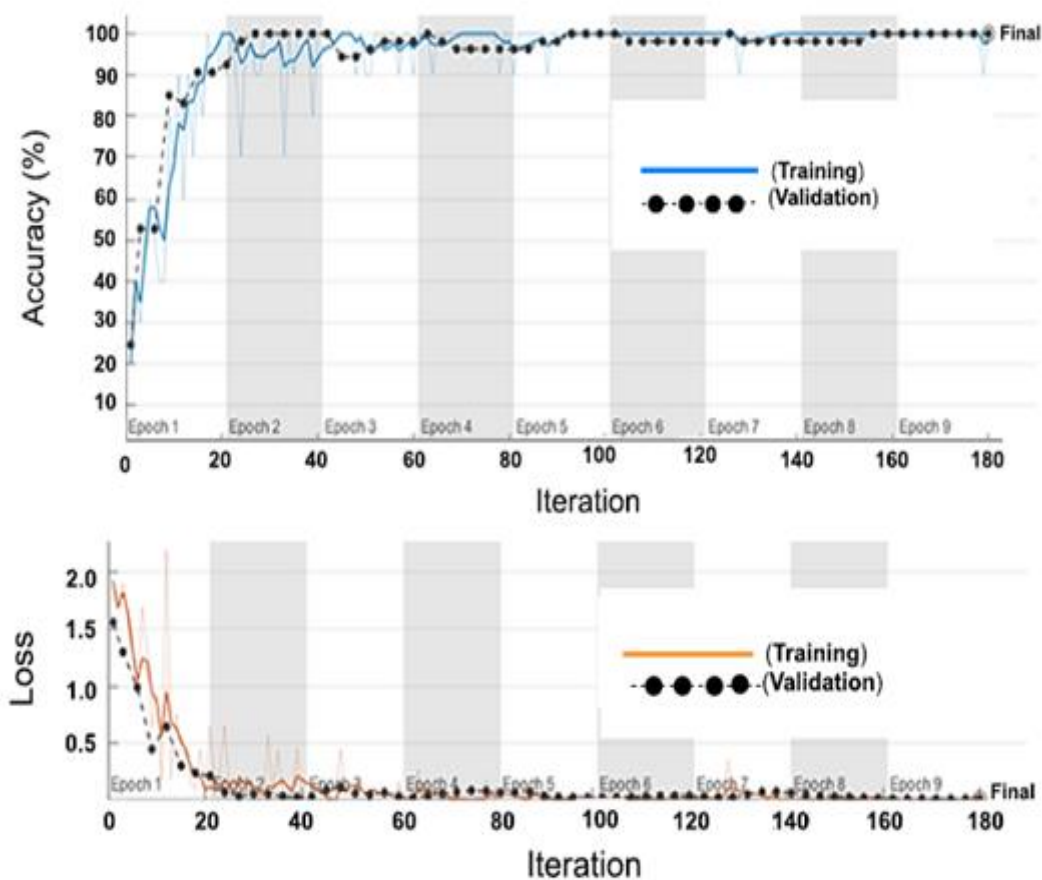
شکل ۵: ساختار AlexNet



شکل ۶: نقشه تصمیم برای تنظیم شبکه از قبل آموزش دیده روش مبتنی بر یادگیری انتقالی

برای مسئله طبقه بندی مدنظر، سه لایه آخر با لایه های جدید جایگزین شده اند: یک لایه کاملاً متصل، یک لایه سافت مکس و یک لایه خروجی طبقه بندی متناسب با مجموعه داده فعلی. عدد ۵ برای اندازه لایه های کاملاً متصل و تعداد کلاس ها، تنظیم گردیده است. مقادیر نرخ های یادگیری برای وزن ها و بایاس ها نیز روی ۲۰ تنظیم شده اند. در اینجا، برای آموزش شبکه از روش نزول گرادیان/شیب تصادفی با

حرکت (SGDM) استفاده می شود. در نزول گرادپان، وزن بر اساس بیشترین نرخ که تابع خطا با آن نزول می کند، آپدیت می شود. هشتاد درصد نمونه ها برای آموزش و ۲۰ درصد آن ها برای اعتبار سنجی استفاده می شود و این نمونه ها به طور تصادفی در هر تکرار انتخاب می شوند. آموزش با تنظیم 0.0001 برای نرخ یادگیری اولیه، ۹ به عنوان حداکثر تعداد دوره ها برای آموزش، با استفاده از یک دسته کوچک با ۱۰ مشاهده در هر تکرار و تعیین فرکانس اعتبارسنجی در ۳، انجام گرفته است. نمونه ای از نمودار پیشرفت آموزشی به دست آمده، در شکل ۷ نشان داده شده است. از آن می توان مشاهده کرد که در نهایت هر دو افت آموزش و افت اعتبارسنجی، تقریباً یکسان هستند، بنابراین کاملاً مناسب است.



شکل ۷: فرآیند آموزش روش مبتنی بر یادگیری انتقالی

۲-۲-۱-۲-۳ معیارهای عملکرد

معیار مورد استفاده برای ارزیابی عملکرد بخش بندی WBC شامل صحت متوسط، IoU متوسط (فصل مشترک بالاتر از اجتماع)، نمره BF متوسط و دقت میانگین متوسط، است. صحت متوسط میانگین درصد پیکسل های پیش بینی شده صحیح را برای هر گروه اندازه گیری می کند و IoU اشتراک یا

سازگاری بین مرز مورد انتظار و مرز واقعی را که توسط معادله ۳-۱ شرح داده شده اندازه گیری می کند. نمره BF ثبات مرزهای تقسیم شده را اندازه گیری می کند. مقدار آن در محدوده [۰ ۱] قرار دارد. اگر حدفاصل شی بخش بندی شده و زمینه واقعی دقیقاً مطابقت هم باشند، نمره BF برابر یک است. در تمام موارد تشخیص شی، ما یک آستانه IoU تعریف می کنیم، اگر IoU محاسبه شده بیش از آستانه باشد، گفته می شود که شی شناسایی شده و پیش بینی مربوطه مثبت درست تلقی می شود. دقت متوسط نیز محاسبه گردیده شده است که سطح زیر منحنی بازیابی-دقت است و گرفتن میانگین آن برای همه کلاس، دقت میانگین متوسط را نشان می دهد.

$$IoU = \frac{prediction \cap groundtruth}{prediction \cup groundtruth}$$

معیارهای مورد استفاده برای تعیین عملکرد مدل طبقه بندی چند کلاسه عبارتند از: صحت، دقت، بازیابی، مشخصه و نمره F1. این پارامترها را می توان از ماتریس درهم ریختگی محاسبه کرد که خلاصه-ای از پیش بینی های درست و نادرست را ارائه می دهد. شکل ۸ ماتریس درهم ریختگی یک مسئله پنج کلاسه را نشان می دهد. B، E، L، M و N نمایانگر کلاس ها هستند. ورودی های قطری تعداد همه نمونه های طبقه بندی شده صحیح را نشان می دهد و عنصر غیر قطری MIS_{ij} نشان دهنده طبقه بندی نادرست است، به عنوان مثال، کلاس i به عنوان کلاس j پیش بینی شده است.

		Predicted				
		B	E	L	M	N
Actual	B	TP _B	MIS _{BE}	MIS _{BL}	MIS _{BM}	MIS _{BN}
	E	MIS _{EB}	TPE	MIS _{EL}	MIS _{EM}	MIS _{EN}
	L	MIS _{LB}	MIS _{LE}	TPL	MIS _{LM}	MIS _{LN}
	M	MIS _{MB}	MIS _{ME}	MIS _{ML}	TPM	MIS _{MN}
	N	MIS _{NB}	MIS _{NE}	MIS _{NL}	MIS _{NM}	TPN

شکل ۸: ماتریس درهم ریختگی مسئله پنج کلاسه در روش مبتنی بر یادگیری انتقالی

صحت به صورت مجموع طبقه بندی های صحیح تقسیم بر تعداد کل طبقه بندی ها محاسبه می شود. دقت، بازیابی، مشخصه و نمره F1 با استفاده از معادلات زیر محاسبه می شوند.

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$specificity = \frac{TN}{TN + FP}$$

$$F1score = \frac{2 \times precision \times recall}{precision + recall}$$

اینجا، TP و TN نشان دهنده تعداد مثبت واقعی و منفی واقعی هستند، FP و FN نیز به ترتیب تعداد مثبت کاذب و منفی کاذب را نشان می‌دهد. برای یک کلاس معین، مجموع مقادیر در ردیف مربوطه به استثنای TP تعداد کل FN ها را بدست می‌دهد، در حالی که نتیجه‌ای که با افزودن مقادیر در ستون مربوطه غیر از TP بدست می‌آید، بیانگر تعداد کل FP ها است. TN با افزودن مقادیر در همه سطرها و ستون‌ها به جز سطر و ستون مرتبط با آن کلاس بدست می‌آید.

۴-۱-۲-۲ نتایج

آزمایشات با استفاده از MATLAB R2019b در رایانه شخصی با پردازنده Intel CORE i5 با سرعت ۱/۸۵ گیگاهرتز و ۱۶ گیگابایت رم و پردازنده گرافیکی Intel UHD Graphics 620 انجام شده است. مکان‌های WBC با استفاده از بخش‌بندی معنایی بر اساس مدل DeepLabv3+ مشخص شده و برای برش سلول مربوطه از تصاویر میکروسکوپی از کادر محدودکننده استفاده شده است. به منظور نشان دادن کارآمدی مدل بخش‌بندی، اعتبارسنجی ۱۰ گانه بکار برده شده است. در طول مرحله آزمایش یا همان تست مرحله بخش‌بندی، صحت میانگین برابر ۹۸/۲۲٪، IoU میانگین برابر ۸۴/۲٪ و دقت میانگین متوسط برابر ۹۸/۴۲٪ بدست آمدند.

در طول طبقه‌بندی، ۸۰٪ از داده‌ها برای اهداف آموزشی و ۲۰٪ باقیمانده برای اعتبارسنجی استفاده شده است. پنج تکرار با انتخاب تصادفی نمونه‌های آموزشی و آزمایشی از پایگاه داده اصلاح شده انجام گردیده است. صحت اعتبارسنجی در هر تکرار و میانگین و انحراف معیار آن محاسبه شده است. صحت میانگین کلی در هنگام اعتبارسنجی $98,87 \pm 1$ ٪ بدست آمده است.

صحت میانگین بخش‌بندی برای همه سلول‌ها به جز لنفوسیت‌ها و مونوسیت‌ها بیش از ۹۹٪ است. همچنین، صحت طبقه‌بندی برای بازوفیل‌ها و ایوزینوفیل‌ها ۱۰۰٪ است، در حالی که این مقدار برای مابقی سلول‌ها صحت بیش از ۹۸٪ است.

۲-۲-۲ بخش‌بندی مبتنی بر الگوریتم k-means

در مطالعه‌ای که از این روش استفاده نموده است [۵]، تصاویر توسط سه مولفه رنگی RGB مشخص شده‌اند. سه مولفه رنگی تصویر، قرمز، سبز و آبی بصری آزمایش شده‌اند و معلوم گردیده است که مولفه سبز حاوی بیشترین اطلاعات کنتراست است. این صفحه برای مرحله بخش‌بندی انتخاب شده است.

روش پیشنهادی مبتنی بر الگوریتم k-means با بخش‌بندی گسترده بردارها به گروه‌هایی که تعداد آن‌ها یکسان است، آغاز می‌شود. نقطه مرکز نشان دهنده گروه است. با بخش‌بندی یک تصویر به خوشه‌های K، طبقه‌بندی و گروه‌بندی موارد در k گروه (k تعداد گروه‌های از پیش انتخاب شده است)، به حداقل رساندن مجموع فاصله‌های مربع بین موارد و مرکز متناظر مورد استفاده در گروه‌بندی بخش‌بندی انجام شده است. این روش شامل مراحل زیر است:

(۱) مراکز خوشه‌ای اولیه انتخاب شده‌اند.

(۲) توزیع نمونه‌ها $\{x\}$ در میان خوشه‌های K با استفاده از رابطه با k امین مرحله تکرار.

$$x \in C_j(K) \text{ if } \|x - Z_j(K)\| < \|x - Z_i(K)\|$$

$$i=1, \dots, k$$

$C_j(k)$ مجموعه نمونه‌ها را نشان می‌دهد که مرکز خوشه‌ای آن‌ها $Z_j(k)$ است.

(۳) $Z_j(k+1), j = 1, 2, \dots, K$ مرکز خوشه‌های جدید است. به حداقل رساندن خوشه

جدید با توجه به مجموع فواصل مربع از تمام نقاط $C_j(k)$. حداقل سازی با میانگین $C_j(k)$ اندازه‌گیری می‌شود.

$$Z_j(K+1) = 1/N_j [x \in C_j(K) \sum x, j = 1 \dots K]$$

N_j تعداد نمونه‌ها در $C_j(k)$ است.

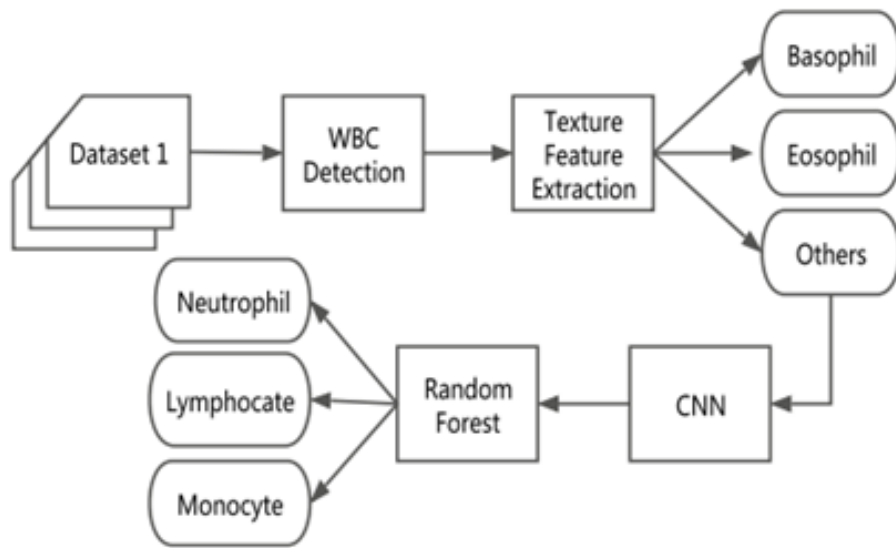
۴) اگر $Z_j(k+1) = Z_j(k)$ for $j = 1, 2, \dots, K$, الگوریتم خاتمه می یابد. در غیر این صورت به مرحله ۲ می رود.

مقدار مراکز خوشه اولیه K مورد استفاده در خوشه بندی به عنوان ورودی الگوریتم مشخص شده است. برای بخش بندی سلول های بلاست یک الگوریتم خوشه بندی K -mean با $k = 3$ و به دنبال آن $K = 2$ اعمال شده است. در هر مرحله، پیکسل های خوشه ای برگشتی حداکثر مقدار سبز هستند.

مناطق ناخواسته در تصویر بخش بندی شده وجود دارد. پس پردازش، فقط برای بلاست لازم است. مشکلات تصویری، مانند همپوشانی سلول و اعوجاج سلول، در مرحله تقویت حل می شود. در این روش، یک الگوریتم جداسازی سلول ارائه شده که شکل اصلی سلول خونی را حفظ کرده و از اطلاعات شکل آن برای تقسیم مناطق همپوشانی شده، استفاده می کند. پیکسل ها بر اساس واریانس با استفاده از حداقل کمی سازی واریانس به گروه ها پیوند می یابند. با بخش بندی فاصله ای واترشد این مشکل، رفع شده است. این روش به طور خودکار موارد بیش برآزش شده را از هم جدا میکند.

۲-۲-۳ شناسایی و طبقه بندی لکوسیت ها با استفاده از شبکه عصبی کانولوشنی

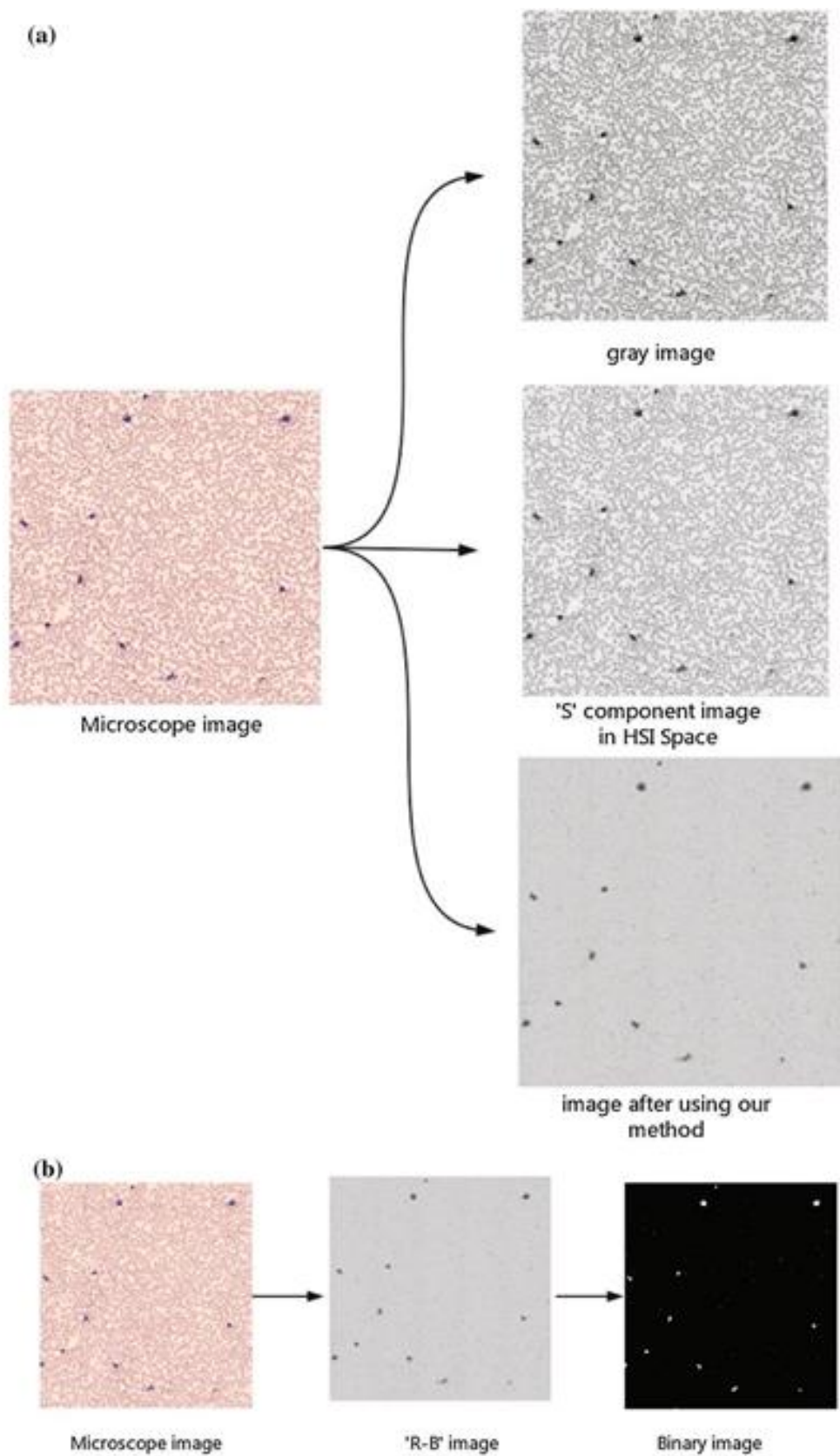
سیستم پیشنهادی [۳] ابتدا WBC ها را از روی تصاویر میکروسکوپی بر اساس ارتباط ساده رنگ ها و عملکردهای ریخت شناسی تشخیص داده است. سپس برای هر WBC، یک ویژگی گرانولاریتی (PRICoLBP) استخراج شده و SVM بر روی این ویژگی های گرانولاریتی اعمال شده است تا ایوزینوفیل و بازوفیل را از سایر انواع WBC تشخیص دهد. برای سه نوع WBC باقیمانده، CNN بکار برده شده است تا ویژگی های آن ها را به صورت خودکار در سطح بالایی استخراج کند. سپس طبقه بند جنگل تصادفی، برای شناسایی این که به کدام نوع WBC تعلق دارد، اعمال می شود. فلوجارت روش پیشنهادی در شکل ۹ نشان داده شده است.



شکل ۹: فلوچارت روش شناسایی و طبقه بندی لکوسیت‌ها با استفاده از شبکه عصبی کانولوشنی

۱-۳-۲-۲ تشخیص WBCها

تصویر خون محیطی I_0 نه تنها شامل انواع مختلف WBC، بلکه تعداد زیادی گلبول قرمز نیز می‌باشد. بنابراین، اولین قدم در این روش پیشنهادی، شناسایی WBCها از تصویر I_0 است. به طور کلی، پس از رنگ آمیزی سلول‌های خونی با برخی از روش‌ها، به عنوان مثال، روش رنگ آمیزی رایت، هسته‌های WBC به رنگ ژرفتری ظاهر می‌شوند، که باعث می‌شود WBCها به راحتی از سلول‌های دیگر متمایز شوند. روش‌های کلاسیک برای شناسایی هسته، اغلب تصویر میکروسکوپ را از فضای RGB به برخی از فضای رنگی دیگر مانند فضاهای HSI و HSV تبدیل می‌کنند و سپس اجزای قابل توجه با مقدار آستانه مناسب را انتخاب می‌کنند. در اینجا، در این روش، ویژگی خاص تصویر سلول‌های خونی برای تشخیص WBCها به کار گرفته شده است. یعنی ابتدا از مقدار اختلاف $R - B$ از رنگ R و B استفاده نموده تا یک تصویر $R - B I_1$ بدست آید، سپس روی تصویر $R - B I_1$ با یک مقدار آستانه باینریز شده تا تصویر I_2 بدست آید که هسته‌ها را ایجاد می‌کند. در شکل ۱۰ مشاهده می‌شود که تصویر I_2 به دست آمده شامل برخی از اشیا کوچک است که هسته WBC نیستند و ممکن است برخی هسته‌ها نیز کامل نباشند.



شکل ۱۰: مقایسه تغییر رنگ روش پیشنهادی [۳] با برخی از روش‌های دیگر برای شناسایی ابتدایی WBC

به منظور حذف آن اشیا کوچک از تصویر I_2 و تکمیل هسته ها، برخی از عملگرهای عملیات مورفولوژیکی مانند عملیات فرسایش و اتساع بر روی تصویر I_2 اعمال شده است. عملیات فرسایش می تواند اجسام کوچک و ناچیز را از بین ببرد. فرسایش B روی X را می توان چنین توصیف کرد

$$X \ominus B = \{z \in \mathbb{Z}^2 \mid (B)_z \subset X\},$$

که X تصویر باینری، B عنصر ساختاری و $(B)_z$ انتقال B نسبت به نقطه z است. در تضاد عملگر فرسایش، عملگر اتساع می تواند نواحی خالی درون اشیا را پر کند. عملیات اتساع B بر روی X را به این صورت می توان توصیف نمود:

$$X \oplus B = \{z \in \mathbb{Z}^2 \mid (\hat{B})_z \cap X \neq \emptyset\},$$

که B^{\wedge} بازتاب B است. در روش پیشنهادی، ابتدا برای خلاص شدن از اشیا کوچک و ناچیز، یک بار از عملیات فرسایش و برای رشد هسته WBC ها از دو بار اتساع استفاده شده است. فرمول بدست آوردن یک تصویر واضح I_3 از تصویر I_2 را می توان به شرح زیر توصیف کرد:

$$I_3 = (I_2 \ominus B \oplus B \oplus B) \cap I_2,$$

که در آن B برخی از عناصر ساختاری انتخاب شده است. با اطلاعات هسته ها در تصویر I_3 ، مکان یابی و برش WBC ها شروع می شود. در نظر بگیرید (x_i, y_i) مختصات مرکزی حداقل مستطیل محدود کننده A_i باشد که شامل WBC i ام است، $i = 1, 2, \dots, N$ ، گاهی اوقات، مستطیل انتخاب شده ممکن است یک سلول کامل را پوشش ندهد، بلکه فقط یک لوب هسته باشد. بنابراین باید مستطیلی پیدا کنیم که شامل کل WBC باشد. در اینجا، فاصله اقلیدسی مراکز دو مستطیل مورد محاسبه قرار گرفته است. اگر فاصله از قطر هسته کمتر باشد و تعداد پیکسل ها در دو منطقه از حداکثر مقدار کمتر باشد، در نظر می گیریم که این دو لوب در یک هسته قرار دارند و مستطیل و مرکز آن را با فرمول زیر به روز می کنیم:

$$A_i \leftarrow A_i \cup A_j \quad \text{and} \quad (x_i, y_i) \leftarrow \left(\frac{x_i + x_j}{2}, \frac{y_i + y_j}{2} \right)$$

تا زمانی که شرط فوق از بین برود. به طور کلی، قطر هسته و حداکثر مقدار، مربوط به بزرگنمایی تصویر است.

تا به این اینجا، WBC ها به صورت خودکار از تصویر محیطی شناسایی می شوند.

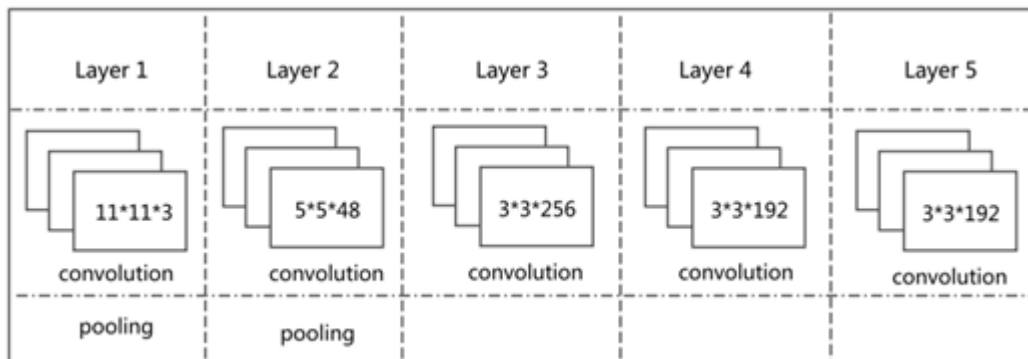
۲-۳-۲-۲ طبقه‌بندی WBCها

روش‌های سنتی برای طبقه‌بندی WBCs معمولاً برخی از ویژگی‌های هسته یا سیتوپلاسم مانند ویژگی‌های هندسی، ویژگی‌های بافت یا ویژگی‌های رنگ برای طبقه‌بندی استفاده می‌کنند. بنابراین، نتایج طبقه‌بندی بستگی زیادی به عدم شباهت ویژگی‌های طراحی شده دارد. به عبارت دیگر، نحوه طراحی یک ویژگی موثر برای طبقه‌بندی WBC بسیار مهم است. در روش طبقه‌بندی پیشنهادی، به دنبال استخراج ویژگی WBCها بر اساس CNN بوده‌اند. CNN هنگام برخورد با مجموعه تصاویر بزرگتر عملکرد بهتری از خود نشان می‌دهد، در حالی که نسبت ائوزینوفیل‌ها و بازوفیل‌ها در تصویر خون محیطی به ترتیب ۱ ~ ۵ درصد و کمتر از ۱ درصد است. بنابراین در ابتدا قبل از استفاده از CNN در WBCها ائوزینوفیل و بازوفیل از سایر WBCها متمایز شده‌اند.

از آنجا که ائوزینوفیل و بازوفیل مملو از مقداری کنسانتره و گرانول‌های بزرگ هستند، از ویژگی PRICoLBP برای تجسم گرانولیتی ائوزینوفیل و بازوفیل استفاده شده است که قدرت تمایز ائوزینوفیل و بازوفیل را از این طریق افزایش می‌دهد. انواع دیگر WBCها، PRICoLBP گونه‌ای از الگوی دودویی محلی (LBP) است که توصیف‌کننده‌ی بافت ثابت در مقیاس خاکستری است.

ویژگی PRICoLBP نه تنها می‌تواند به طور موثر اطلاعات مکانی مشترک در زمینه فضایی را حفظ کند، بلکه از عدم تغییر چرخش نیز برخوردار است. با استفاده از ویژگی‌های PRICoLBP به دست آمده برای پنج نوع WBC، از SVM برای طبقه‌بندی آنها در سه کلاس ائوزینوفیل، بازوفیل و سایر موارد استفاده شده است. اکنون کار طبقه‌بندی سه نوع WBC باقی مانده است: نوتروفیل، مونوسیت و لنفوسیت. در اینجا، CNN برای استخراج ویژگی‌های WBC در سطح بالا اعمال شده و جنگل‌های تصادفی به عنوان طبقه‌بند برای تشخیص این ویژگی‌ها در نظر گرفته شده است. ایده CNN کشف چندین سطح نمایش است، با این فرض که ویژگی‌های سطح بالاتر می‌توانند معنایی انتزاعی‌تر از داده‌ها را نشان دهند. انتظار می‌رود که ویژگی‌های استخراج شده از یک شبکه ژرف، تأثیر بیشتری در تنوع درون کلاس داشته باشند. CNN یک شبکه عصبی پیش‌خور ویژه است که از چندین لایه کانولوشن و لایه‌های پولینگ تشکیل شده است. لایه کانولوشن تصویر ورودی را با چند ماتریس کوچک از وزن فیلتر می‌کند و برخی از عملکردهای غیر خطی را به عنوان یک تابع فعال اعمال می‌کند. به عنوان مثال، برای یک تصویر ورودی I ، اگر $W \in R^{k1 \times k2}$ فیلتر وزنه باشد، سپس عملیات در لایه کانولوشن گرفتن کانولوشن دو بعدی $I \times W$ است، و تابع فعال می‌تواند $f(x) = \max(x, 0)$ باشد. لایه پولینگ حاوی وزن

نیست و به سادگی با عملکرد حداکثر-پولینگ اندازه خروجی قبلی را کاهش می‌دهد. در این مقاله، CNN اعمال شده، شامل ۵ لایه کانولوشن و ۲ لایه پولینگ است. ساختار آن در شکل ۱۱ نشان داده شده است.



شکل ۱۱: ساختار بتونی CNN در روش پیشنهادی برای استخراج ویژگی‌ها در سطح بالا با استفاده از ۵ لایه کانولوشن و ۲ لایه پولینگ

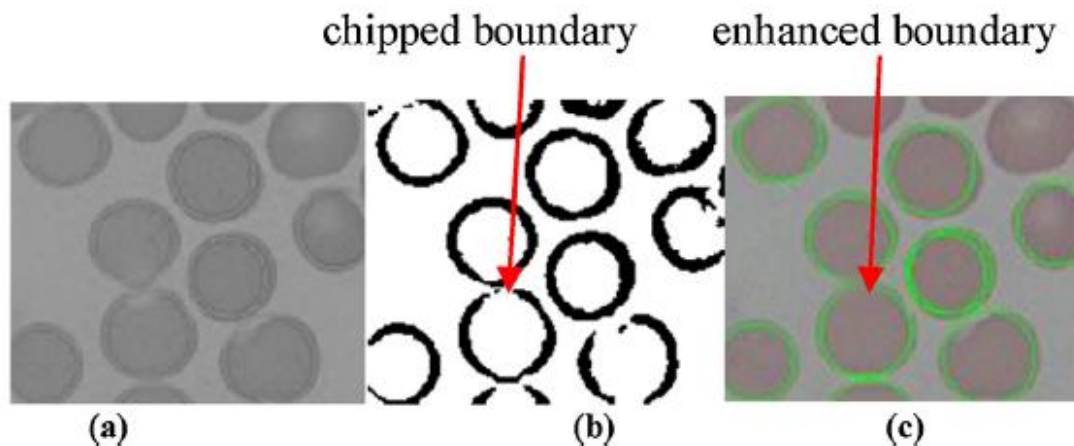
۴-۲-۲ بخش‌بندی هسته و سیتوپلاسم سلول با استفاده از تصاویر میکروسکوپی

در فرآیند بخش‌بندی هسته که در ادامه معرفی شده است، طبقه‌بندی‌های ساده مانند K-means و ... به عنوان ابزارهای مناسب در مطالعه مورد بررسی [۶]، در نظر گرفته شده‌اند تا هسته را از سلول تقسیم کنند.

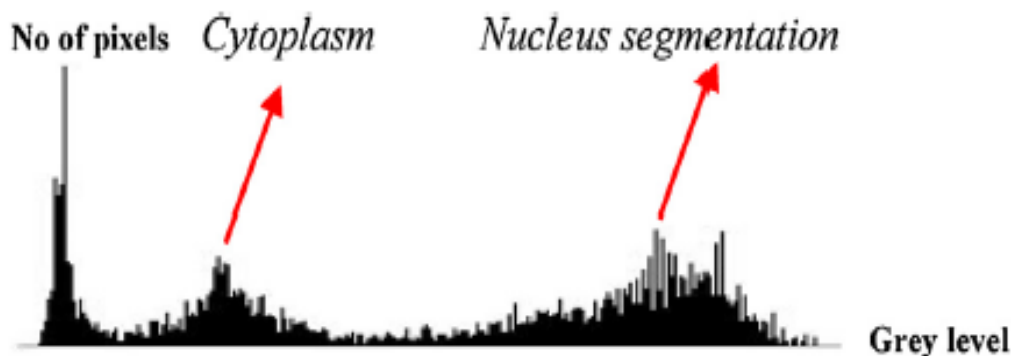
۱-۴-۲-۲ استخراج هسته

بر اساس دانش متخصصان پزشکی، سلول‌های خونی سالم، که به شکل دایره‌ای هستند، از طریق تصویر میکروسکوپ برای تشخیص، انتخاب می‌شوند. در عمل، تصویر اصلی (شکل ۱۲- (a)) باینری می‌شود و مرز برای تشخیص کاندیداهای سلول‌های خونی از پس زمینه مشخص می‌شود و سلول‌های گرد نیز بعضاً توسط فرآیند دوتایی شدن تراشه می‌شوند، همانطور که در نمونه شکل (۱۲- (b)) نشان داده شده است. سلول‌های خونی در نهایت با تغییر شکل دایره‌ای هاف همانطور که توسط دایره‌های سبز در شکل ۱۲- (c) نشان داده شده است، از بین همه کاندیداها انتخاب می‌شوند. کاربران باید طبقه‌بندی را با برخی از نمونه‌ها آموزش دهند تا مقدار آستانه مناسب گردی برای تبدیل‌ها را از قبل تعیین کنند و از آستانه آزمایش گردی استفاده کنند. در تجزیه و تحلیل تصویر سلول از نظر هیستوگرام همانطور که در شکل ۱۳ نشان داده شده است، شدت سیتوپلاسم و هسته کاملاً در سطح خاکستری قابل تفکیک هستند، بنابراین یک مقدار آستانه را می‌توان برای مراحل آموزش از قبل تعیین کرد و یک تبدیل به

سادگی استفاده می‌شود برای تقسیم هسته از تصویر سلول با این حال، می‌توان با تعیین تعداد کلاس‌ها از قبل، و بخش‌بندی تصویر سلول به دو کلاس، از ابزارهای ساده مانند K-means استفاده کرد و سپس هسته را از تصویر سلول استخراج کرد.



شکل ۱۲: انتخاب سلول‌های خون سالم



شکل ۱۳: هیستوگرام هسته و سیتوپلاسم در سلول خونی

۲-۲-۵ بخش‌بندی هسته و سیتوپلاسم سلول با استفاده از تصاویر ابر طیفی

برای مشخص کردن لکوسیت، ابتدا باید لکوسیت را از پس زمینه جدا کنیم. برخی از طبقه‌بندی‌های غلط به دلیل شباهت بین طیف‌های دو منطقه مختلف رخ می‌دهد. به دلیل تنوع WBC، به دست آوردن نتایج دقیق بخش‌بندی تصاویر RGB دشوار است. علاوه بر این، طبقه‌بندی WBC تصاویر RGB نیز وابستگی بیش از حد به عملکرد بخش‌بندی را نشان داده است. گر بخش‌بندی تصویر ناصحیح باشد، نمی‌توان استخراج ویژگی فضایی را نیز بدست آورد. بنابراین، برخی از فن‌آوری‌های جدید برای حل این

مشکل فراخوانده شده‌اند. به این ترتیب، فناوری تصویربرداری ابرطیفی میکروسکوپی می‌تواند راه حل جدیدی ارائه دهد. تصویربرداری ابرطیفی یک فناوری نوظهور است که مزایای تصویربرداری طیفی و تکنیک‌های تصویربرداری سنتی را با هم ترکیب می‌کند. برای بدست آوردن اطلاعات مفیدتر جهت استخراج ویژگی‌ها، اجزای متمایز مانند هسته و سیتوپلاسم برای طبقه‌بندی بیشتر، در مطالعه مورد بررسی [۱]، تقسیم شده‌اند. این روش در ادامه بیان می‌شود.

۲-۲-۵-۱ بخش‌بندی سلول

از الگوریتم مخروط محدب حداکثر زاویه متوالی (SMACC) برای تعیین کل لکوسیت استفاده شده است. SMACC روشی مبتنی بر مدل مخروطی محدب برای بدست آوردن اعضای نهایی از تصاویر ابر طیفی اصلی و تهیه تصاویر فراوانی به صورت خودکار است. این مدل، از قطب برای تعیین مخروط محدب استفاده می‌کند و اولین عضو آخر را تعریف می‌کند. یک طرح مورب با شرایط محدود، بر هر م موجود برای تولید عضو آخر بعدی استفاده می‌شود، روند برای تولید اعضای آخر جدید ادامه می‌یابد. از لحاظ ریاضی، SMACC از فرمول زیر برای گسترش اعضای نهایی استفاده می‌کند:

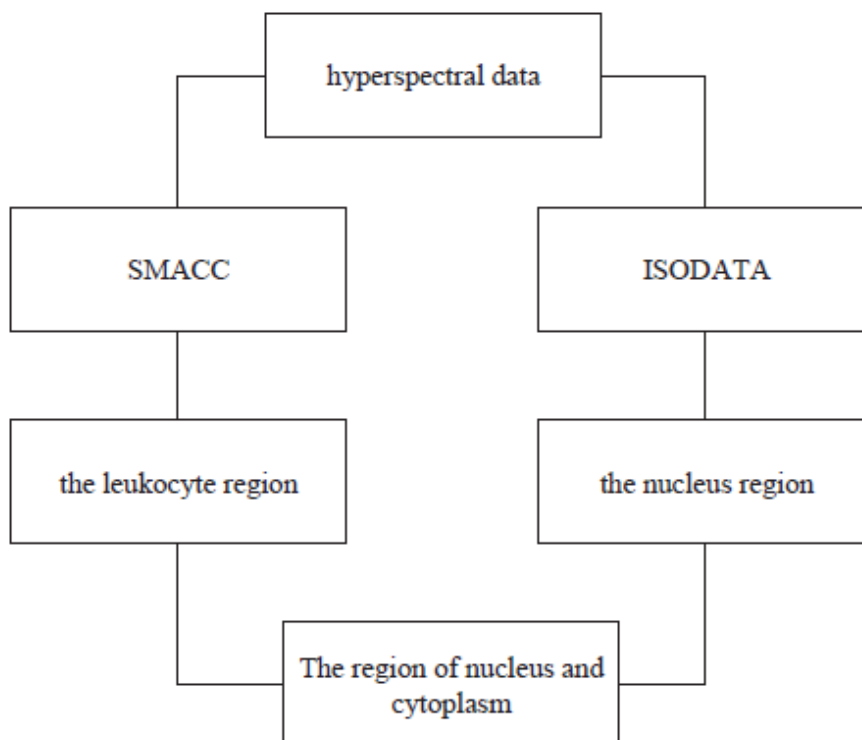
$$I(j, i) = \sum_m^X E(j, m)A(m, n)$$

که در آن $A(m, n)$ تصویر فراوانی است و $E(j, m)$ ، طیف‌های اعضای نهایی است. X نشان دهنده حداکثر شماره عضو آخر است. پس از مقداردهی تصویر فراوانی لکوسیت‌ها، بزرگترین منطقه متصل را پیدا می‌شود که از طریق آن می‌توان برخی از پیکسل‌های مربوط به منطقه سلول را حذف کرد. با توجه به اینکه ممکن است لکوسیت بخش‌بندی شده کاملاً متصل نباشد و در طول مرز شکاف وجود داشته باشد، برای حل مسئله از مورفولوژی استفاده می‌شود.

۲-۲-۵-۲ بخش‌بندی هسته‌ها و سیتوپلاسم

معمولاً مشخصه هسته‌ها هایلایت می‌شود، بنابراین لکوسیت در داخل هسته و سیتوپلاسم تفکیک-پذیرتر می‌شود. طیف هسته لکوسیت مشخصه‌ای است که قابل تفکیک پذیری بیشتری نسبت به بقیه قسمت‌های لکوسیت است. پس از بدست آوردن طیف هر پیکسل، از ISODATA تکرار شونده برای خوشه بندی پیکسل‌های با طیف مشابه استفاده می‌شود زیرا روند طبقه‌بندی را خودکارتر و سریع‌تر می‌کند. از آنجا که تعداد کلاس‌ها در تصویر گلبول‌های خون نامشخص است، روش طبقه‌بندی بدون

نظارت با مداخله کمتر انسان موثرتر است. سپس نتایج خوشه بندی با رنگ‌های مختلف مشخص می‌شود. علاوه بر این، ممکن است برخی از طبقه‌بندی‌های نادرست در منطقه هسته وجود داشته باشد، پس برای تقویت یکپارچگی منطقه هسته از عملیات مورفولوژیکی استفاده شده است. با ترکیب نتایج با منطقه لکوسیت‌های تقسیم شده می‌توان یک منطقه هسته کامل و یک منطقه سیتوپلاسم بدست آورد. نمودار روند فرآیند بخش‌بندی در شکل ۱۴ نشان داده شده است.



شکل ۱۴: فلوجارت فرآیند بخش‌بندی در روش بخش‌بندی هسته و سیتوپلاسم سلول با استفاده از تصاویر ابر طیفی

۳-۲ نتیجه‌گیری

در این فصل چندین روش بخش‌بندی لکوسیت، مورد مطالعه قرار گرفت. هر روش مزایا و معایب خود را دارد. به عنوان مثال، روش آستانه ساده است اما قادر به بخش‌بندی دقیق WBCها از پس زمینه نیست. برخی از روش‌ها (به عنوان مثال، روش SVM و روش رشد منطقه‌ای) می‌توانند به طور معقولی نتایج تشخیص دقیق داشته باشند، اما هزینه آن‌ها زمان بر است و به منابع محاسباتی بالایی نیاز دارند. در حالی که برخی از روش‌های بخش‌بندی مبتنی بر رنگ، مستقیماً بر روی فضای رنگی RGB انجام می‌شوند، در برخی از رویکردها فضای رنگی-اشباع رنگ (HSI) به ویژه در مولفه‌ی S به کار گرفته شده است. به طور کلی، روش‌های مبتنی بر مولفه‌های S از عملکردهای مبتنی بر RGB بهتر عمل می‌کنند.

از طرف دیگر، روش‌های تشخیص خودکار کمی در میان آن‌ها وجود دارد که WBCها را از تصاویر میکروسکوپی تشخیص دهند. به این معنی که بسیاری از روش‌ها، طبقه‌بندی را با تصاویر WBC برش داده شده توسط برخی از متخصصان آغاز می‌کنند، که منجر به ایجاد مشکل در برنامه‌های واقعی می‌شود [۳].

در مطالعات انجام شده، روش‌های بخش‌بندی لکوسیت معرفی شده در این فصل در کنار انواع مراحل استخراج و انتخاب ویژگی و طبقه‌بندی‌های مختلف، مورد استفاده قرار گرفته و نتایج متفاوتی حاصل شده است. در فصل بعد به بررسی روش‌های طبقه‌بندی خواهیم پرداخت.

فصل ۳

تشخیص سرطان خون (لوسمی) از تصاویر گلبول‌های سفید

۱-۳ مقدمه

چندین روش برای تشخیص لوسمی در طول سال‌ها ارائه شده است. در این میان، برخی از این روش‌ها برای طبقه‌بندی دو نوع شایع لوسمی ارائه شده است: لوسمی میلوئید حاد (AML) و لوسمی لنفوبلاستیک حاد (ALL). برخی از جدیدترین روش‌ها فقط با استفاده از تصاویر با یک لکوسیت در هر تصویر و برخی دیگر با استفاده از تصاویر با چندین لکوسیت، تشخیص را ارائه داده‌اند. برخی از مطالعات به منظور شناسایی بیماران سرطانی از تصاویر سلول خونی، با استفاده از بخش‌بندی سلول و برخی دیگر بدون استفاده از بخش‌بندی، روش‌های خود را ارائه داده‌اند. در این فصل به بررسی این روش‌ها پرداخته شده است.

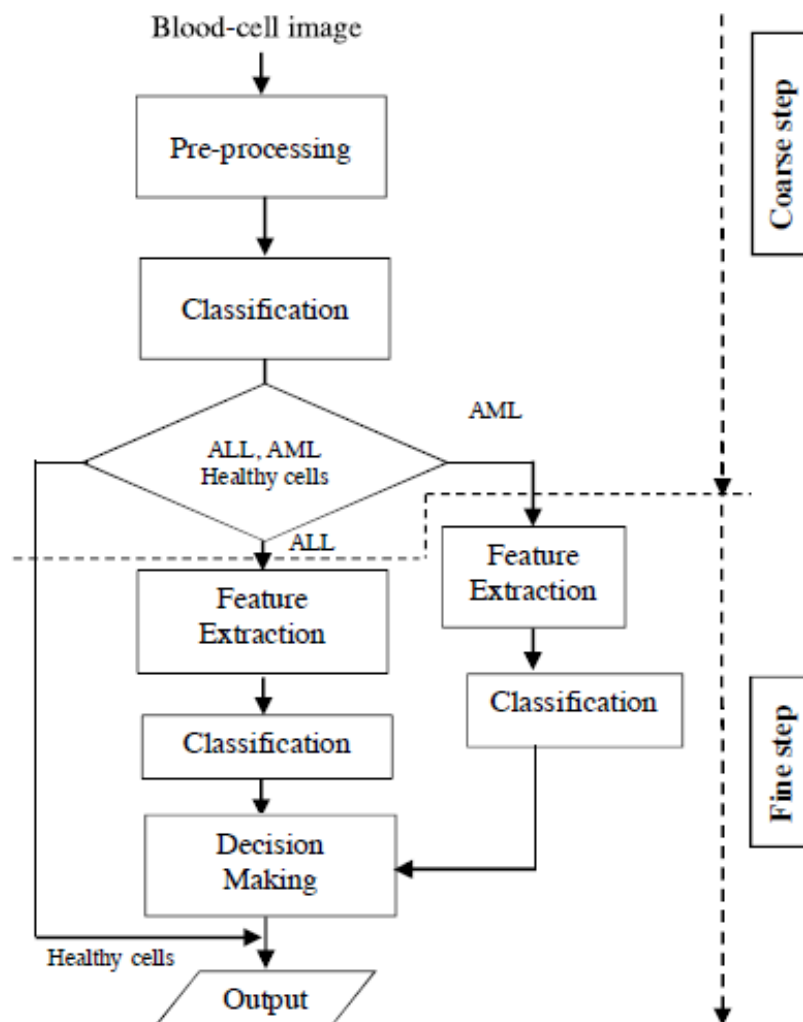
۲-۳ روش‌های تشخیص با استفاده از بخش‌بندی

در دسته‌ای از مطالعات برای شناسایی و طبقه‌بندی بیماران مبتلا به سرطان خون ابتدا از روش بخش‌بندی لکوسیت‌ها استفاده شده است و سپس روش پیشنهادی مد نظر روس تصاویر بخش‌بندی شده اعمال شده است. در این بخش به بررسی این مطالعات پرداخته شده است.

۱-۲-۳ طبقه‌بندی سرطان خون حاد با استفاده از مورفولوژی مبتنی بر دانش پزشکی و نشانگر CD

طبقه‌بندی زیرگروه‌ها در سرطان خون ALL و AML یک پیش‌فرایند حیاتی برای انتخاب یک درمان مناسب برای بیماران مبتلا به سرطان خون حاد است و به یک سیستم متخصص خودکار برای کمک به متخصصان پزشکی نیاز دارد. برای توسعه سیستم اتوماتیک، یک روش طبقه‌بندی از زیرگروه‌های لوسمی حاد مورد نیاز است. در حال حاضر، خوشه‌بندی نشانگر (CD) که توسط دانشمندان پزشکی به عنوان اطلاعات مهم ژنتیکی اثبات شده است، از نظر بالینی برای طبقه‌بندی زیرگروه‌های حاد لوسمی در مقایسه با نتایج طبقه‌بندی با استفاده از ویژگی‌های مورفولوژیکی استفاده می‌شود. در زمینه پزشکی، سلول‌های خونی ابتدا توسط ALL، AML و گروه‌های سالم توسط ویژگی‌های پروسپترون مانند تعداد هسته‌ها، نسبت سیتوپلاسمی و اندازه هسته طبقه‌بندی می‌شوند و گروه‌های ALL و AML به زیرگروه‌هایی مانند L1، L2، M1، M2 و ... با استفاده از ویژگی‌های هسته آن‌ها طبقه‌بندی می‌شوند. بنابراین در این مطالعه [۶] روشی برای طبقه‌بندی زیر مجموعه سلول مورفولوژیکی بر اساس مفهوم درشت به ریز

دانش پزشکی فعلی پیشنهاد شده است. این بدان معنی است که ابتدا ALL، AML و سلول‌های سالم در مرحله درشت از هم جدا شده و سلول‌ها به ALL و AML طبقه‌بندی می‌شوند و سپس در مرحله ریز به زیرگروه‌های خود طبقه‌بندی می‌شوند. این زیرگروه‌ها که نتایج طبقه‌بندی مورفولوژی را نشان می‌دهند، سرانجام به عنوان کاندید برای تایید زیرگروه‌های سلول با مارکرهای CD در فرایند تصمیم‌گیری استفاده می‌شوند. در ارزیابی عملکرد روش پیشنهادی، آزمایشات به ترتیب با ۲۰۰ و ۳۰۰ نمونه از سرطان خون حاد برای آموزش و آزمایش انجام شد و نتایج شامل صحت ۹۹/۶۷٪ با بهبود ۴/۹۴٪ در مقایسه با روش‌های مرسوم است. در فصل قبل نحوه استخراج هسته این روش شرح داده شد و در ادامه جزئیات ادامه روش ارائه می‌شود. فلوجارت این روش در شکل ۱۵ نشان داده شده است.



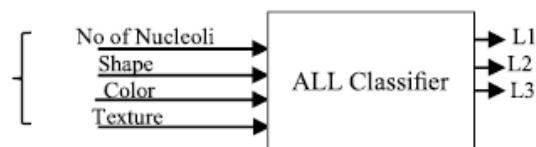
شکل ۱۵: فلوجارت روش طبقه‌بندی سرطان خون حاد با استفاده از مورفولوژی مبتنی بر دانش پزشکی و نشانگر CD

۳-۲-۱-۱ استخراج ویژگی

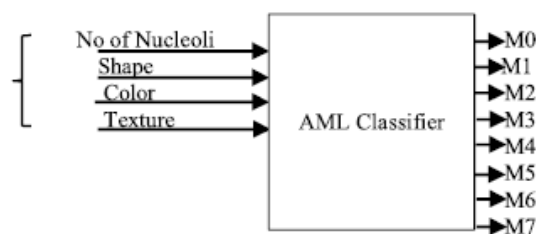
دانش پزشکی ثابت کرده و پذیرفته است که ویژگی‌هایی مانند شکل هسته، توزیع رنگ در هسته، بافت در هسته، تعداد هسته‌ها در هسته و غیره از ویژگی‌های مهم برای طبقه‌بندی ALL به L1، L2 و L3 و AML به M0، M1، M2، M3، M4، M5، M6 و M7 است. این مقاله سعی در تحلیل این ویژگی‌های دانش بنیان پزشکی دارد، تبدیل ویژگی‌ها را از نظر اطلاعات تصویر به ویژگی‌های عینی در نظر گرفته و از الگوریتم‌های موجود مناسب برای طبقه‌بندی استفاده کرده است.

۳-۲-۱-۲ طبقه‌بندی

نتایج بدست آمده در مرحله درشت، یک سلول خونی را به گروه‌های سالم، ALL و AML طبقه‌بندی می‌کند. در گروه ALL و AML، سیستم اتوماتیک متعاقباً سلول‌های خونی را به یکی از زیرگروه‌های ALL و AML برای هر گروه طبقه‌بندی می‌کند. طبقه‌بندی گروه‌های ALL و AML همان چیزی است که در شکل ۱۶ نشان داده شده است، بنابراین جزئیات بصورت زیر شرح داده شده است. اول، ویژگی‌های مبتنی بر دانش پزشکی، که در این مقاله تعداد نقاط هسته، شکل هسته، رنگ و بافت است استخراج و برای طبقه‌بندی برای آموزش و آزمایش استفاده می‌شود. در طبقه‌بندی، ویژگی‌های مهم مبتنی بر دانش پزشکی همانطور که گفته شد با استفاده از نمونه‌های شناخته شده از قبل یاد گرفته می‌شوند، و سپس نمونه‌های هدف برای انواع ALL و AML آزمایش می‌شوند.



(a) ALL Types classification



(b) AML Types classification

شکل ۱۶: طبقه‌بندی انواع ALL و AML

۲-۲-۳ سیستم پشتیبانی تصمیم برای طبقه بندی سرطان خون حاد بر اساس تصاویر میکروسکوپی دیجیتال

در عصر تصویربرداری میکروسکوپی دیجیتال، پردازش تصویر، تجزیه و تحلیل داده‌ها، طبقه بندی، سیستم‌های پشتیبانی تصمیم به عنوان یکی از مهمترین ابزارها برای تحقیقات تشخیصی هستند. پزشکان می‌توانند با تصور و تجزیه و تحلیل تصاویر، ناهنجاری‌های ساختار داخلی سلول را مشاهده کنند. شناخت سلول‌های لوسمی حاد در تصاویر میکروسکوپی رنگی یک کار چالش برانگیز است. اولین مرحله مهم در تشخیص خودکار این بیماری، بخش‌بندی تصویر، مهمترین مرحله محسوب می‌شود. در این مطالعه [۵]، یک سیستم پشتیبانی تصمیم ارائه شده است که شامل انتخاب پانل، بخش‌بندی با استفاده از خوشه بندی K-means برای شناسایی سلول‌های لوسمی و استخراج ویژگی‌ها و بازیابی تصویر است. پس از اینکه سیستم پشتیبانی تصمیم با موفقیت سلول‌ها و ساختار داخلی آن را شناسایی کرد، سلول‌ها بر اساس ویژگی‌های مورفولوژیکی آن‌ها از این تجزیه و تحلیل طبقه بندی می‌شوند سپس این الگوریتم به یک مجموعه داده دیگر که تحت نظارت یک آسیب شناس متخصص از یک بیمارستان محلی استخراج شده است، اعمال شد. کل مجموعه داده شامل ۷۵۷ تصویر جمع آوری شده از دو مجموعه داده است. تصاویر مجموعه داده‌ها با سه برچسب مختلف برچسب گذاری شده‌اند که نشان دهنده سه نوع سلول سرطان خون است: بلاست، میلویت و سلول‌های بخش‌بندی شده. روند برچسب‌گذاری این تصاویر توسط آسیب شناس متخصص بازیابی شده است. آزمایش الگوریتم با استفاده از این مجموعه داده، صحت کلی ۹۹/۵۱۷٪، حساسیت ۹۹/۳۴۸٪ و ویژگی ۹۹/۵۲۹٪ را نشان داده است. بنابراین، این الگوریتم نتایج امیدوار کننده‌ای را به همراه داشته است. جزئیات روش بخش‌بندی این مطالعه در فصل قبل مورد بررسی قرار گرفت. در ادامه به بررسی مراحل بعدی پرداخته می‌شود.

۱-۲-۲-۳ استخراج ویژگی

استخراج ویژگی‌ها، طبقه بندی و شناخت سلول‌های سرطان خون را تسهیل می‌کند. در این مرحله، ویژگی‌های استخراج شده بر اساس هندسه، آمار، بافت و نسبت اندازه از مناطق انتخاب شده در روند بخش‌بندی (هسته، سیتوپلاسم و سلول کامل) است. سپس تجزیه و تحلیل این ویژگی‌ها برای تمایز انواع سرطان خون حاد انجام شده است.

۳-۲-۲-۲ درخت تصمیم

یادگیری درخت تصمیم یکی از معمول‌ترین و اساسی‌ترین روال‌ها برای استنتاج استقرایی است. این یک استراتژی برای تقریب کارهای هدف تخمین زده-مجزا است و در آن با ظرفیت تصمیم‌گیری توسط درخت تصمیم‌گیری صحبت می‌شود.

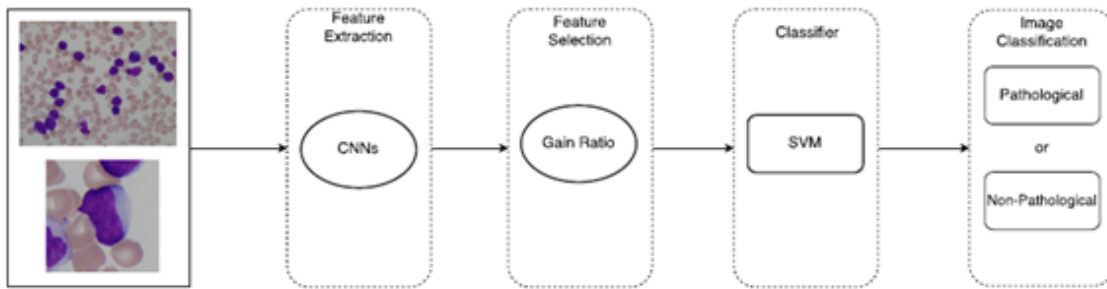
درختان تصمیم‌گیری برگزیده، ترتیب‌ها را مرتب می‌کنند تا درخت را از ریشه به یک قطب برگ مرتب کنند که ترتیب وقوع را ترتیب می‌دهد.

۳-۲-۲-۳ شبکه عصبی مصنوعی (ANN)

الگوریتم‌های چند لایه پرسپترون اغلب برای مشکلات طبقه‌بندی در پزشکی استفاده می‌شوند. تمام MLP ها از ۲ لایه پنهان با ظرفیت تبادل هذلولی انحرافی و یک لایه تک‌نورون با یک کار تبادل لجستیک تشکیل می‌شوند که عملکرد MLP را تشکیل می‌دهد. شبکه‌های عصبی با استراتژی اعتبارسنجی متقابل ۱۰ گانه آماده شده و آزمایش می‌شوند تا در صورت استفاده کامل از مجموعه اطلاعات روش پیشنهادی، امکان آزمایش نمونه مشابه را کاهش دهند. اطلاعات به طور تصادفی به ده زیر مجموعه تقسیم شد. یک زیرمجموعه برای آزمایش استفاده شده، در حالی که نه زیرمجموعه باقیمانده برای تهیه اهداف استفاده شده است. برای طبقه‌بندی شبکه عصبی از توابع طبقه‌بندی Weka Multilayer Perceptron استفاده گردیده است. ورودی‌های شبکه شامل ۶۴۸ نمونه ویژگی ۲۳ ویژگی استخراج شده از سلول‌ها است.

۳-۳ روش‌های تشخیص بدون استفاده از بخش‌بندی

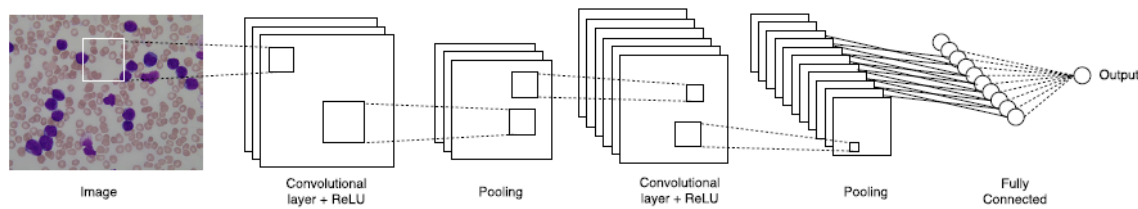
در مقاله [۲] تشخیص لوسمی با استفاده از تصاویر لکه خون انجام شده است. از فلوجارت نشان داده شده در شکل ۱۷، می‌توان مشاهده کرد که سیستم از یک تصویر بدون هیچ‌گونه پیش‌پردازش یا بخش‌بندی به عنوان ورودی استفاده می‌کند. این تفاوت اصلی بین روش پیسنهادی و روش‌های پیشرفته اخیر است. CNNها برای توصیف تصویر ورودی استفاده می‌شوند و ویژگی‌ها نیز انتخاب و کاهش می‌یابند. در مرحله طبقه‌بندی، SVM برای طبقه‌بندی تصاویر به عنوان پاتولوژیک یا غیر پاتولوژیک مورد استفاده قرار گرفته است.



شکل ۱۷: فلوجارت روش تشخیص سزطان خون با استفاده از CNN

۱-۳-۳ شبکه‌های عصبی کانولوشنی

در حال حاضر، محققانی که با هوش مصنوعی کار می‌کنند، از تکنیک‌های یادگیری عمیق برای ایجاد سیستم‌های محاسباتی قوی استفاده می‌کنند. در میان این تکنیک‌ها، CNNها جهت کمک به تشخیص، پیشنهاد شده‌اند که بهتر از روش‌های معمول استخراج ویژگی عمل کرده و نرخ صحت بهتری حاصل شده است. ساختار CNN در شکل ۱۸ نشان داده شده است.



شکل ۱۸: توضیح ساده از ساختار CNN

در این روش از CNNها در یک پایگاه داده طبیعی بزرگ محتوی تصاویر با ۱۰۰۰ دسته به نام ImageNet که از قبل آموزش دیده‌اند، استفاده شده است. در میان این دسته‌ها، پایگاه داده حاوی تصاویر لنفوسیت‌ها و لنفوبلاست‌ها نیز می‌باشد. در طول توسعه رویکرد ما، سه مدل انتخاب شده است: AlexNet، Vgg-f و CaffeNet. این مدل‌ها از ساختارهای مشابهی برخوردارند و تفاوت اصلی آنها در اندازه فیلترهای استفاده شده در لایه‌های کانولوشن و تعداد سلول‌های عصبی در لایه‌های کاملاً متصل است. تعداد ویژگی‌های گرفته شده از آخرین لایه کاملاً متصل هر ساختار ۴۰۹۶ است.

۳-۳-۲ استخراج ویژگی

پس از ایجاد بردار ویژگی، انتخاب ویژگی انجام شده است. این انتخاب با هدف از بین بردن ویژگی‌های غیرضروری و در نتیجه ساده سازی مدل پیش بینی، کاهش هزینه محاسباتی و درک بهتر از نتایج به دست آمده، است.

الگوریتم نسبت بهره یا گین جهت انتخاب ویژگی مورد استفاده قرار گرفته است. این الگوریتم بهره اطلاعات را محاسبه می کند که از آنتروپی به عنوان اندازه گیری درجه ناخالصی و درختان تصمیم برای نمایش ویژگی‌های بردار استفاده می کند. عملکرد آن بهتر از بهره اطلاعات در نظر گرفته می شود، بنابراین صحت بهتری در مورد پیچیدگی درختان تولید شده، فراهم می کند.

۳-۳-۳ طبقه بندی

برای طبقه بندی ویژگی‌های حاصل از انتخاب ویژگی، SVM انتخاب شده است. این طبقه بندی با استفاده از شباهت‌های بین ویژگی‌های هر نمونه، سعی می کند دو کلاس از یک مجموعه داده را با استفاده از یک خط جداسازی یا ابر صفحه از هم جدا کند. طبقه بندی‌های دیگری نیز در این آزمایش استفاده شده اند که عبارتند از: MLP، RF و KNN. در آزمایشات انجام شده در این مطالعه، طبقه بندی SVM بهترین نتیجه را ارائه داده است.

۳-۴ نتیجه گیری

در این فصل کارهای مربوط به تشخیص سرطان خون، مورد بررسی قرار گرفت. بیشتر رویکردهای ذکر شده، ویژگی‌های شکل، رنگ و بافت را ترکیب کرده اند. در فرآیند طبقه بندی، غالباً طبقه بندی‌های SVM، پرسپترون چند لایه (MLP)، جنگل تصادفی (RF)، K نزدیک ترین همسایگی (KNN)، شبکه تابع پایه شعاعی (RBFN) و بیزین ساده (NB) مورد استفاده قرار گرفته اند [۲].

السالم، در مقاله خود ذکر کرده است که انگیزه‌های محققان برای توسعه و استفاده از روش‌های خودکار برای تشخیص، شناسایی و طبقه بندی سرطان خون حاد بر افزایش سرعت، صحت، هزینه کم، عملکرد بدون متخصص، مجموعه داده‌ها، روش‌ها و تکنیک‌ها متمرکز است [۲۶]. انگیزه روش‌های طبقه بندی و تکنیک‌های بکار رفته در مطالعات مورد بررسی السالم در این زمینه در جدول ۱ ارائه شده است.

جدول ۱: انگیزه روش‌های طبقه‌بندی و تکنیک‌های بکار رفته در مطالعات [۲۶]

انگیزه	روش	
ساده و دارای یک صحت پیش بینی شگفت آور رقابتی، س تعمیم نتایج حتی در اندازه مجموعه داده کوچک نیز خوب است.	شبکه‌های بیزین	۱
عملکرد خوبی از خود نشان می‌دهد و به دوره آموزش کمتر نیاز دارد. پیچیده نیست، مقیاس‌پذیری خوبی را نشان می‌دهد، ساده و موثر	KNN	۲
عملکرد خوبی نشان می‌دهد و از نظر محاسباتی گران نیست. می‌واند به موارد جدید تعمیم داده شود و می‌تواند خطی یا غیرخطی باشد. ابر صفحه مطلوب و خصوصیات تعمیم خوب، حاشیه بهینه شده بین کلاس‌ها، عملکرد خوب، تعداد کمتری از نمونه‌های آموزش به طور موثرتر برای بدست آوردن بهینه‌ترین ابر صفحه استفاده می‌شود. استفاده از داده‌های میکرو آرایه در مسائل طبقه بندی مفید است. مناسب برای مسائل طبقه بندی باینری. می‌تواند بین دو کلاس تفاوت قائل شود. یکی از سریع‌ترین طبقه بندها برای آموزش و طبقه بندی؛ می‌تواند چالش‌های مجموعه داده بیان ژن، مانند اندازه کوچک و ابعاد را پردازش کند.	SVM	۳
قوانین فازی را می‌توان با توصیف و تفسیر نتایج با اصطلاحات ساده که برای انسان قابل درک است، به راحتی درک کرد. عملکرد طبقه بندی را بهبود می‌بخشد.	روش مبتنی بر نقشه فازی	۴
عملکرد خوبی را در مورد زمان آموزش نشان می‌دهد. کمتر مستعد ابتلا به over-fitting است، وابستگی زیادی به داده‌های آموزش ندارد و شفاف است.	طبقه‌بند ناهموار فازی	۵
می‌تواند داده‌های غیرخطی را مدیریت کرده و برخی از مسائل موجود در داده مانند نویز و اوت‌لایر را اداره کند.	سیستم منطق فازی فاصله نوع ۲	۶
روش فازی را با الگوریتم KNN ترکیب می‌کند و عملکرد طبقه بندی کلی را افزایش می‌دهد.	الگوریتم KNN فازی	۷
صحت طبقه بندی را بهبود می‌دهد، تعداد ویژگی‌ها را کاهش می‌دهد.	برنامه نویسی ژنتیک	۸
متناسب با طبقه بندی چند طبقه.	ماشین توزیع حاشیه بزرگ	۹
راه حل بهتری را بر اساس جستجوی انعطاف پذیر با استفاده از رفتار و الگوی زندگی مورچه‌ها ارائه می‌دهد.	الگوریتم‌های استخراج قانون وابسته به تئوری مورچه	۱۰
طبقه بندی دقیق ارائه می‌دهد و در مقایسه با سایر روش‌های یادگیری پیچیده نیست.	طبقه‌بند بیزین ساده	۱۱
مناسب برای داده‌های تک طبقه برای مدل طبقه بندی ساخته شده توسط یک ابر محیط.	شرح داده‌های بردار پشتیبان	۱۲
توانایی به حداقل رساندن تابع انرژی برای تعیین ارتباط بین بیومارکرها.	شبکه‌های تجمع بیومارکر	۱۳
نرخ دقت بالاتری را در مقایسه با الگوریتم‌های طبقه بندی اخیر در منطقه ML ارائه می‌دهد. مناسب برای مجموعه داده‌های بزرگ	جنگل تصادفی	۱۴
با انتخاب خودکار ویژگی‌های متمایز، یک مدل طبقه بندی دقیق ارائه می‌دهد.	درخت عصبی انعطاف پذیر	۱۵
برای گروه کوچکی از ژن‌ها استفاده شده است. می‌تواند مشکلات over-fitting را برطرف کند و در نتیجه عملکرد بالایی دارد.	تقویت SVM	۱۶

در سال‌های اخیر، زمینه تشخیص و طبقه‌بندی سرطان خون افزایش یافته است، اگرچه هنوز از بسیاری جهات مهم با موضوعات و چالش‌هایی روبرو است. به عنوان مثال برای چنین چالش‌هایی می‌توان به ابعاد بالا، نویز در مجموعه داده‌های آزمون، اندازه نمونه کم و غیره اشاره کرد. خلاصه‌ای از این چالش‌ها در جدول ۲-۳ ارائه شده است.

جدول ۲: دسته‌بندی چالش‌های طبقه‌بندی و تشخیص سرطان خون در مطالعات [۲۶]

چالش‌های مربوط به تشخیص	خطای انسانی، خستگی، طول مدت، تکرار شونده‌گی. برخی از دیدگاه‌های ذهنی یک متخصص. گزارش‌های غیراستاندارد، ناسازگار و ذهنی.
چالش مربوط به زمان تشخیص	معاینات لکه زمان بر است. تجزیه و تحلیل مورفولوژیک سلول‌های خونی در مقایسه با جدی بودن سرطان خون حاد کند است.
چالش مربوط به اندازه کم نمونه	چالش اصلی این است که تعداد نمونه‌ها در مقایسه با هزاران ژن کم است.
چالش مربوط به ابعاد بالا	مشکل اصلی که از طریق یادگیری ماشین مبتنی بر طبقه‌بندی پیش می‌آید، داده‌های بعد بالا است. طبقه‌بندی داده‌های ریزآرایه با "p بزرگ"، "n کوچک" مواجه است. الگوهای زائد توسط طبقه‌بند یافت می‌شود. بی‌ثباتی محاسباتی را ایجاد می‌کند.
چالش مربوط به ویژگی‌ها (ژن‌ها)	کاهش تعداد ویژگی‌ها. ویژگی‌های غیر مرتبط هزینه محاسباتی بیشتری به همراه دارد. تعیین زیرمجموعه بهینه یا افتراقی ژن دشوار است. ژن‌های بی‌ربط منجر به مشکل در شناسایی رابطه بیولوژیکی می‌شوند. ویژگی‌های نامربوط باعث افزایش اندازه مبتنی بر قانون می‌شوند. عملکرد طبقه‌بندی ممکن است با ویژگی‌های نامربوط کاهش یابد. چالش تشخیص حداقل بیماری باقی مانده (MRD). همپوشانی سلول‌ها در تصاویر مسئله جدی در طبقه‌بندی است.
چالش مربوط به چالش‌های فنی	تنوع ژنتیکی بر بیان ژن تأثیر می‌گذارد. اختلاف در بین تولیدکنندگان منجر به تفاوت مواد مورد استفاده در تجزیه و تحلیل ریزآرایه بیان ژن شده است. تفاوت در رویه‌های کاری مانند تهیه نمونه، عملکرد ابزارها و جمع‌آوری داده‌ها. تنوع رنگ آمیزی بالا، تنوع در کیفیت لکه/اسمیر نمونه‌ها.
چالش مربوط به نویز بیولوژیکی و فنی	تجزیه و تحلیل داده‌های بیان ژن به طور قابل توجهی نویزدار است. ذرات گرد و غبار روی تصویر. نویز از انعکاس لنز و پخش نور در اسکنر.
چالش مربوط به ارزیابی	دشواری ارائه مجموعه داده‌های عمومی در دسترس برای ارزیابی عملکرد الگوریتم‌ها.

صحت تنها هدف ارزیابی نیست. هزینه‌های طبقه بندی نادیده گرفته می‌شوند.	
صحت و قابلیت اطمینان دستگاه اعتماد مبدل با میزان خطای طبقه بند خاص متفق است.	قابلیت اطمینان و پیش بینی‌ها
مقدار اندازه عدم انطباق/عدم موافقت تحت تأثیر تنظیمات مختلف پارامتر قرار دارد.	تنظیمات پارامتر
پیچیدگی محاسبه زمان بر است. ویژگی‌های فراوان باعث افزایش زمان مورد نیاز می‌شود.	چالش مربوط به صرف وقت
دشواری در به دست آوردن یک نتیجه طبقه بندی مطلوب توسط ML. ژن‌های غیر مرتبط و زاید می‌توانند دچار overfitting شوند.	چالش مربوط به overfitting
همبستگی مطالعه بین سطوح بیان دشوار است.	چالش مربوط به دانش بیولوژیکی
مشکلات درخت‌های تصمیم‌گیری کلاسیک. اتوماسیون یادگیری هنوز برای ANN مسئله دشواری است. بیزین ساده با underflow و overfitting روبرو است. kNN با برابری نمونه‌های آموزش روبرو است.	چالش‌های مربوط به روش‌های طبقه بندی

فصل ۴

جمع‌بندی و ارائه پیشنهادات

در این سمینار به بررسی انواع روش‌های شناسایی و بخش‌بندی لکوسیت‌ها و در ادامه آن روش‌های شناسایی و طبقه‌بندی سرطان خون پرداخته شد. در فصل ۱ انواع روش‌های تشخیص و داده‌های مورد استفاده، ذکر گردید. در فصل ۲ روش‌های انجام شده در راستای شناسایی و بخش‌بندی لکوسیت‌ها معرفی شدند. هر کدام از این روش‌ها دارای مزایا و یا معایبی هستند که بسته به شرایط مورد نظر می‌توان از آن‌ها استفاده کرد. در فصل ۳ مطالعات انجام شده در جهت شناسایی و تشخیص بیماران مبتلا به سرطان خون با استفاده از تصاویر خونی مرور شدند. انگیزه‌ها و چالش‌های موجود در این زمینه نیز مورد بحث قرار گرفتند. به عنوان توصیه به ادامه دهندگان این مسیر می‌توان به این موارد اشاره کرد که فرآیند انتخاب ویژگی‌ها به متمایزترین ویژگی‌ها برای دستیابی به سطح بالایی از صحت طبقه‌بندی منجر می‌شود. انتخاب ویژگی بهینه همچنین برای بهبود نتایج طبقه‌بندی بسیار مهم است. در حقیقت، صحت طبقه‌بندی تحت تأثیر انتخاب ویژگی است. بنابراین، دقت در انتخاب ویژگی‌های الگوریتم برای اطمینان از مهمترین ویژگی‌های هدف، ضروری است. انتخاب ویژگی‌های حاوی اطلاعات مفید برای صحت طبقه‌بندی مفید است. ارزیابی شکل هسته بهتر از استفاده از شکل سیتوپلاسم برای تشخیص بین لنفوسیت‌های بالغ طبیعی و نابالغ بر اساس ویژگی‌های آن‌ها است. همچنین، چگونگی یافتن روش تشخیص موثرتر و چگونگی بهبود طبقه‌بندی تصاویر لنفوسیت‌ها بر اساس روش‌های انجام شده نیز از اهداف مطالعه آینده پژوهشگران است.

پیشنهادات برای کارهای آینده:

- انجام تحقیقات بیشتری تحت حضور سلول‌های همپوشان یا مجاور در تصاویر مانند افزایش مقدار نمونه‌های لکوسیت که می‌تواند قابلیت اطمینان نتایج آزمایش را بیشتر بهبود بخشد و طبقه‌بندی انواع بیشتری از سلول‌ها را به دست آورد.

- تنظیم دقیق در CNN جهت بهبود جداسازی اطلاعات لوسمی

- استفاده از پایگاه داده‌های جدیدی از تصاویر با مقدار داده بیشتر، به سیستم این اجازه را خواهد داد که در زندگی روزمره ارزیابی و استفاده شود و به پزشکان و بیماران نیز در تشخیص این بیماری کمک کند.

- بهبود شرایطی که مجموعه داده شامل سلول‌های تحت تأثیر قرار دهنده بوده و درعین حال سرعت محاسبه نیز افزایش یابد.

- ایجاد یک سیستم بخش‌بندی قوی مستقل از لکه‌های استفاده شده در تصاویر لکه خون

- گسترش اندازه مجموعه داده تا مدل طبقه بندی با تعداد بیشتری نمونه مفید در مرحله آموزش ارائه شود. این گسترش باید انواع سلول‌های لوسمی میلوئید حاد را برای تمایز کامل تمام سلول‌ها به کار گیرد.

در نهایت برای به روز ماندن، محققان باید از فناوری‌های جدید مانند برنامه‌های تلفن همراه یا تلفن‌های هوشمند پشتیبانی کنند. تحقیقات در مورد تشخیص و طبقه بندی سرطان خون حاد بر اساس دستگاه‌های موبایل (سیار) هنوز انجام نشده است، بنابراین یک تحقیق مهم ضروری است که ممکن است با چندین خط فناوری و علمی دیگر تلاقی کند.

- .١ Duan, Y., et al., *Leukocyte classification based on spatial and spectral features of microscopic hyperspectral images*. Optics & Laser Technology, 2019. **112**: p. 530-538.
- .٢ Vogado, L.H., et al., *Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification*. Engineering Applications of Artificial Intelligence, 2018. **72**: p. 415-422.
- .٣ Zhao, J., et al., *Automatic detection and classification of leukocytes using convolutional neural networks*. Medical & biological engineering & computing, 2017. **55**(8): p. 1287-1301.
- .٤ Asaad, N.Y., M.M. Abd El-Wahed, and M.M. Dawoud, *Diagnosis and prognosis of B-cell chronic lymphocytic leukemia/small lymphocytic lymphoma (B-CLL/SLL) and mantle cell lymphoma (MCL)*. J Egypt Natl Canc Inst, 2005 : (٤)١٧ .p. 279-290.
- .٥ Negm, A.S., O.A. Hassan, and A.H. Kandil, *A decision support system for Acute Leukaemia classification based on digital microscopic images*. Alexandria engineering journal, 2018. **57**(4): p. 2319-2332.
- .٦ Laosai, J. and K. Chamnongthai, *Classification of acute leukemia using medical-knowledge-based morphology and CD marker*. Biomedical Signal Processing and Control, 2018. **44**: p. 127-137.
- .٧ Reena, M.R. and P. Ameer, *Localization and recognition of leukocytes in peripheral blood: A deep learning approach*. Computers in Biology and Medicine, 2020. **126**: p. 104034.
- .٨ Lewis, S.M. and D. Roper, *Laboratory methods used in the investigation of the haemolytic anaemias*, in *Dacie and Lewis practical haematology*. 2001, Churchill Livingstone London .p. 149-166.
- .٩ Cao, H., H. Liu, and E. Song, *A novel algorithm for segmentation of leukocytes in peripheral blood*. Biomedical Signal Processing and Control, 2018. **45**: p. 10-21.
- .١٠ Clark, V.L. and J.A. Kruse, *Clinical methods: the history, physical, and laboratory examinations*. Jama, 1990. **264**(21): p. 2808-2809.
- .١١ Labati, R.D., V. Piuri, and F. Scotti. *All-IDB: The acute lymphoblastic leukemia image database for image processing*. in *2011 18th IEEE International Conference on Image Processing*. 2011. IEEE.
- .١٢ Sarrafzadeh, O. and A.M. Dehnavi, *Nucleus and cytoplasm segmentation in microscopic images using K-means clustering and region growing*. Advanced biomedical research, 2015. **4**.
- .١٣ Rollins-Raval, M.A., J.S. Raval, and L. Contis, *Experience with CellaVision DM96 for peripheral blood differentials in a large multi-center academic hospital system*. Journal of pathology informatics, 2012. **3**.
- .١٤ Sarrafzadeh, O., et al. *Detecting different sub-types of acute myelogenous leukemia using dictionary learning and sparse representation*. in *2015 IEEE International Conference on Image Processing (ICIP)*. 2015. IEEE.
- .١٥ B Kekre, H., A. B Patankar, and H. Ramesh Galiyal, *Segmentation of blast using vector quantization technique*. International Journal of Computer Applications, 2013. **72**(15): p. 20-23.
- .١٦ Su, M.-C., C.-Y. Cheng, and P.-C. Wang, *A neural-network-based approach to white blood cell classification*. The scientific world journal, 2014. **2014**.
- .١٧ Theera-Umpon, N. and S. Dhompongsa, *Morphological granulometric features of nucleus in automatic bone marrow white blood cell classification*. IEEE Transactions on Information Technology in Biomedicine, 2007. **11**(3): p. 353-359.
- .١٨ Rezatofghi, S.H. and H. Soltanian-Zadeh, *Automatic recognition of five types of white blood cells in peripheral blood*. Computerized Medical Imaging and Graphics, 2011. **35**(4): p. 333-343.

- .۱۹ Habibzadeh, M., A. Krzyżak, and T. Fevens, *Comparative study of shape, intensity and texture features and support vector machine for white blood cell classification*. Journal of Theoretical and Applied Computer Science, 2013. **7**(1): p. 20-35.
- .۲۰ Liao, Q. and Y. Deng. *An accurate segmentation method for white blood cell images*. in *Proceedings IEEE International Symposium on Biomedical Imaging*. 2002. IEEE.
- .۲۱ Nee, L.H., M.Y. Mashor, and R. Hassan, *White blood cell segmentation for acute leukemia bone marrow images*. Journal of Medical Imaging and Health Informatics, 2012. **2**(3): p. 278-284.
- .۲۲ Gupta, D., et al., *Optimized Binary Bat algorithm for classification of white blood cells*. Measurement, 2019. **143**: p. 180-190.
- .۲۳ Sheikh, H., B. Zhu, and E. Micheli-Tzanakou. *Blood cell identification using neural networks*. in *Proceedings of the IEEE 22nd annual northeast bioengineering conference*. 1996. IEEE.
- .۲۴ Yi, F., et al. *White blood cell image segmentation using on-line trained neural network*. in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. 2006. IEEE.
- .۲۵ Alreza, Z.K.K. and A. Karimian. *Design a new algorithm to count white blood cells for classification leukemic blood image using machine vision system*. in *2016 6th International Conference on Computer and Knowledge Engineering (ICCKE)*. 2016. IEEE.
- .۲۶ Alsalem, M., et al., *A review of the automated detection and classification of acute leukaemia: Coherent taxonomy, datasets, validation and performance measurements, motivation, open challenges and recommendations*. Computer methods and programs in biomedicine, 2018. **158**: p. 93-112.