



تشخیص اتوماتیک بیماریهای شبکیه چشم با استفاده از مدل‌های ریاضیاتی پردازش تصویر، مبتنی بر یادگیری دیکشنری چندلایه

آزاده منتظری¹، محبوبه شمسی²، روح‌الله دیانت³

¹ گروه مهندسی فناوری اطلاعات، دانشکده کامپیوتر، دانشگاه قم، قم، ایران.

² گروه کامپیوتر، دانشکده برق و کامپیوتر، دانشگاه صنعتی قم، قم، ایران. (نویسنده‌ی مسئول) shamsi@qut.ac.ir

³ گروه کامپیوتر، دانشکده کامپیوتر، دانشگاه قم، قم، ایران.

چکیده

هدف از این مطالعه بهبود عملکرد طبقه‌بندی روش‌های نوین، با استفاده از مدلی چندلایه به‌منظور کمک به تشخیص بیماری‌های شبکیه‌ی چشم است. این مدل از الگوریتم K-SVD پیشرفته، برای یادگیری ماتریس دیکشنری و الگوهای پایه استفاده می‌کند تا بتواند با الگوپذیری از معماری چندلایه، ویژگی‌های بهتری را در تصاویر OCT شبکیه بیاموزد. همچنین در این معماری، علاوه بر استفاده از برجسب‌های کلاس داده‌های آموزشی، اطلاعات برجسب نیز در هر ستون پایه در ماتریس دیکشنری ترکیب می‌شود تا در کدگذاری تنک در طی فرآیند یادگیری دیکشنری بیشترین تبعیض اعمال شود که این منجر به موفقیت مراحل کدگذاری تنک و جمع‌بندی، در پیدا کردن نمایش مؤثرتری از داده به‌منظور طبقه‌بندی می‌گردد. برای اعتبارسنجی الگوریتم، از مجموعه داده‌های داک استفاده شده است. این مجموعه داده شامل اسکن‌های حجمی از 45 نوع، که 15 نوع حجم از آن طبیعی، 15 نوع بیمار تخریب ماکولای وابسته به سن (AMD) و 15 نوع بیمار ورم ماکولای دیابتی (DME) است که تعداد اسکن‌های OCT در هر حجم بین 36 تا 97 تصویر متغیر می‌باشد. نتایج تجربی نشان می‌دهد که الگوریتم پیشنهادی این مقاله توانسته است با پیشی گرفتن از بسیاری از مدل‌های جدید یادگیری دیکشنری و نمایش تنک، بسیار خوب عمل نماید و با دقت خوبی منجر به طبقه‌بندی صحیح 95.85% برای تصاویر نرمال و صد درصد برای تصاویر بیمار (DME و AMD) شود.

اطلاعات مقاله

مقاله علمی - پژوهشی

دریافت: 27 اسفند 1396

پذیرش: 04 مهر 1397

واژگان کلیدی:

یادگیری دیکشنری چندلایه

نمایش تنک

الگوریتم K-SVD

مقطع‌نگاری همدوسی شبکیه

طبقه‌بندی

Automatic recognition of retinal diseases using mathematical models of image processing, based on multilayer-dictionary learning

Azadeh Montazeri¹, Mahboubeh Shamsi², Rouhollah Dianat³

¹ Department of Information Technology, Faculty of Computer, University of Qom, Iran

² Department of Computer, Faculty of Computer and Electrical, Qom University of Technology, Iran, (Corresponding author) shamsi@qut.ac.ir

³ Department of Computer, Faculty of Computer, University of Qom, Iran

ARTICLE INFORMATION

Original Research Paper

Received: 18 March 2018

Accepted: 26 September 2018

Keywords:

Multi-layer Dictionary Learning

Sparse Representation

K-SVD Algorithm

Optical Coherence Tomography

Classification

ABSTRACT

The purpose of this study is to improve the classification of new methods using a multi-layered model to address retinal diseases diagnosis. This paper presents a multi-layer dictionary learning method for classification tasks. Our multi-layer framework uses a label consistent in K-SVD algorithm to learn a discriminative dictionary for sparse coding in order to learn better features in retinal optical coherence tomography images. In addition to using class labels of training data, we also associate label information with each dictionary item (columns of the dictionary matrix) to enforce discrimination in sparse codes during dictionary learning process. In fact, it relies on a succession of sparse coding and pooling steps in order to find an effective representation of data for classification. Moreover, we apply Duke dataset for validating our algorithm: Duke spectral domain OCT (SD-OCT) dataset, consisting of volumetric scans acquired from 45 subjects 15 normal subjects, 15 AMD patients, and 15 DME patients. Our classifier leads to a correct classification rate of 95.85% and 100.00% for normal and abnormal (DME and AMD). Experimental results demonstrate that our algorithm outperforms compared to many recent proposed supervised dictionary learning and sparse representation techniques.

1. مقدمه

پردازش تصویر یکی از مسائل مهم در حوزه‌ی هوش مصنوعی می‌باشد که در موارد مختلف صنعتی، پزشکی، نظامی، سیستم‌های امنیتی و... کاربرد دارد. گرچه اغلب این مسائل باهم نامرتبط هستند، اما عموماً نیازمند روش‌هایی هستند که قادر به ارتقای اطلاعات تصویری برای تعبیر و تحلیل انسان باشد. برای نمونه در پزشکی به کمک مدل‌ها و روش‌های موجود وضوح و کیفیت تصویر را ارتقا می‌دهند یا این‌که برای تعبیر آسان‌تر تصاویر اشعه ایکس یا سایر تصاویر پزشکی، سطوح شدت روشنایی را بارنگ، رمز می‌کنند. متخصصان جغرافیایی نیز از این روش‌ها یا روش‌های مشابه برای مطالعه الگوهای آلودگی هوا که با تصویربرداری هوایی و ماهواره‌ای به دست آمده است، استفاده می‌کنند. در باستان‌شناسی نیز روش‌های پردازش تصویر برای بازیابی عکس‌های مات شده‌ای که تنها باقی‌مانده آثار هنری نادر هستند، مورد استفاده قرار می‌گیرد. در فیزیک و زمینه‌های مرتبط، فنون رایانه‌ای بارها تصاویر آزمایش‌های مربوط به موضوعاتی نظیر پلاسماهای پرانرژی و تصاویر ریزبینی الکترونی را ارتقا داده‌اند. کاربردهای موفق دیگری از پردازش تصویر را نیز می‌توان در نجوم، زیست‌شناسی، پزشکی هسته‌ای، اجرای قانون و صنعت بیان کرد [1-5].

به این ترتیب مدل‌های ریاضیاتی زیادی در شرح و توصیف داده‌ها با درجه‌های موفقیت گوناگون ارائه شده که در بین آن‌ها مدل‌های مبتنی بر یادگیری (آموزش) دیکشنری و نمایش تنک یا DLSR¹ از محبوبیت ویژه‌ای برخوردار است و در زمینه‌های مختلف مورد تحقیق و بررسی قرار گرفته است. تحقیقات اخیر نشان داده دیکشنری لرنینگ و نمایش تنک کاربرد مای گوناگونی نظیر، نمایش تصویر، طبقه‌بندی تصویر، حذف نویز، پردازش گفتار، نمایش داده و انتخاب پایه، بازشناسی صورت، طبقه‌بندی سیگنال، بخش‌بندی داده و تحلیل تصاویر پزشکی مانند تصاویر OCT² شبکه دارد [3-9].

شبکه داخلی‌ترین لایه‌ی چشم است و شامل سلول‌های گیرنده‌ی نور و نورون‌ها می‌باشد. بافت شبکه چشم یک ساختار چندلایه‌ای است که با ساختاری مرکب از چندین لایه سلولی، بخش خلفی حفره چشم را در بر گرفته، مسئول تبدیل انرژی نورانی به سیگنال-های عصبی برای تحلیل نهایی توسط مغز می‌باشد. لکه‌ی زرد یا ماکولا بخشی از شبکه است که بیشترین حساسیت به نور را دارد و موجب دید مستقیم واضح می‌شود. لکه زرد برای کارهای دقیق مثل خواندن و رانندگی لازم است. اما بعضی بیماری‌ها می‌توانند این قسمت را درگیر کنند. بیماری‌های ماکولا انواع و اقسام

مختلفی دارد؛ یک نوع آن شامل بیماری‌های تخریبی است که بعضی وابسته به سن و یا ناشی از عوامل ژنتیکی و متابولیک در ماکولا می‌باشد. یکی از بیماری‌های شناخته شده متابولیکی که شیوع بالایی هم دارد، دیابت است. دیابت یا مرض قند می‌تواند باعث ایجاد ورم در ناحیه ماکولای شبکه چشم شود. هرگونه آسیب یا بیماری شبکه، نوعی رتینوپاتی تلقی می‌شود اما زمانی که به واسطه بیماری قند خون رگ‌هایی که سلول‌های شبکه چشم را تغذیه می‌کنند دچار آسیب می‌شوند به آن رتینوپاتی دیابتی گفته می‌شود. ورم ماکولای دیابتی یا DME³ شکلی از رتینوپاتی دیابتی است که در آن از عروق خونی مایعاتی به داخل ماکولا نشت می‌کند. تورم این بخش از شبکه موجب تاری دید می‌شود. همچنین استحاله ماکولای وابسته به سن یا AMD⁴ که در افراد سالمند تظاهر می‌یابد از جمله بیماری‌های شایع مربوط به شبکه چشم است. این بیماری با کاهش دید، بیمار را دچار اختلال بینایی می‌کند [10].

از مهم‌ترین زمینه‌های کاربردی پردازش تصویر استخراج قواعد طبقه‌بندی در حیطه علم پزشکی است. با به کارگیری الگوریتم‌های قدرتمند این حوزه می‌توان سیستم‌های هوشمندی ابداع کرد که به شکل خودکار و بدون نیاز به نظارت پزشک قادر به فهم و تفسیر ویژگی‌های پزشکی افراد باشند یا اطلاعات مفیدی را اکتشاف کنند که متخصصان را در قضاوت صحیح یاری رساند. زمانی که پارامترهای لازم برای تشخیص بیماری زیاد می‌شود، تشخیص و پیش‌بینی بیماری حتی برای یک متخصص خبره پزشکی نیز به سختی ممکن می‌گردد به همین دلیل در چند دهه اخیر ابزار تشخیص کامپیوتری باهدف کمک به پزشک مورد استفاده قرار گرفته است. این مهم موجب شده است که خطاهای احتمالی ناشی از خستگی یا بی‌تجربگی فرد متخصص تا حدی کاهش پیدا کند و داده‌های پزشکی مورد نیاز، طی مدت‌زمان کمتر و با جزئیات و دقت بیشتر در اختیار پزشک قرار گیرد. ابزارهای آنالیز و مدل‌سازی، پتانسیل خوبی برای ایجاد یک محیط غنی از دانش دارند که می‌توانند کمک قابل توجهی به کیفیت تصمیمات بالینی نمایند [1-3].

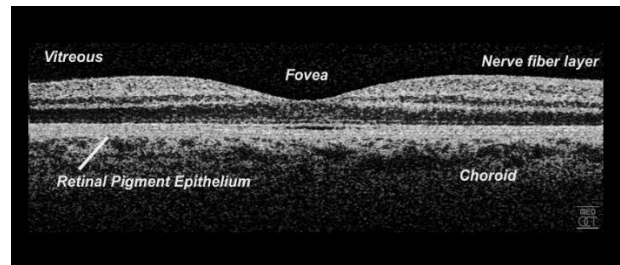
در این میان تلاش‌ها و پیشرفت مای مهمی در عرصه پردازش تصاویر شبکه نیز به منظور ارائه دستگاه‌هایی اتوماتیک جهت تشخیص بیماری‌های مختلف روی آن انجام شده است. چنین سیستم‌هایی علاوه بر اینکه امکان پردازش تصاویر شبکه را در حجم زیاد و با حداقل زمان و هزینه فراهم می‌کنند، از خستگی و ضعف‌های دیگری که شخص تشخیص‌دهنده می‌تواند دچار شود مبرا می‌باشند. تکنیک جدید تصویربرداری OCT (مقطع‌نگاری

داده شد و یک دسته‌بندی دوتایی بر روی تصاویر معمولی و بیمار ADM شکل گرفت. به دنبال آن، آلبراک و همکاران در [28] مدلی بر روی تصاویر سه‌بعدی OCT ارائه نمودند که ابتدا تصویر سه‌بعدی اولیه را به تصاویر با ابعاد کمتر تجزیه نموده سپس آن‌ها را به وسیله‌ی ساختار درختی بازسازی می‌کرد و نهایتاً با ترکیب ویژگی‌های استخراج‌شده از زیر گراف‌ها، تصاویر بیمار و نرمال را تفکیک می‌نمود.

یادگیری دیکشنری و نمایش تنک در سال‌های اخیر منجر به نتایج بسیار خوبی در مسائل کاربردی جدید و به‌ویژه در حوزه‌ی پردازش تصاویر پزشکی و طبقه‌بندی آن‌ها شده و لذا توجه زیادی را به خود جلب کرده است؛ یکی از دلایل موفقیت آن، این است که می‌تواند الگوهای اساسی در داده‌ها را به‌طور مؤثر یاد بگیرد، که منجر به عملکرد قوی آن در این حوزه‌ها شده است. کدهای تنک به‌دست‌آمده همچنین می‌توانند به‌عنوان نمایش جدیدی از سیگنال ورودی و یا به‌عنوان ویژگی‌های استخراج‌شده در فرایند طبقه‌بندی مورد استفاده قرار بگیرند [4-6]. در چنین مواردی، دیکشنری اغلب به روش بدون نظارت آموخته می‌شود. سپس، کدهای تنک به‌دست‌آمده می‌توانند یا به‌طور مستقیم برای طبقه‌بندی و یا به‌عنوان ویژگی‌های مورد استفاده در طبقه‌بندی کننده به‌کارگیری شوند. همچنین تحقیقات اخیر در این مدل‌ها به‌جای بازسازی صرف، بر مزایای مدل‌های یادگیری تنک متمایز نیز تأکید می‌کنند [4، 5، 7]. این امر معمولاً با یادگیری هم‌زمان کد مای تنک و طبقه‌بندی کننده انجام می‌شود. در عمل، هر تصویر ورودی برچسب‌گذاری شده و یادگیری دیکشنری به‌صورت ناظر انجام می‌گردد.

در این مقاله مدلی با ناظر و چندلایه مبتنی بر الگوریتم K -SVD⁵ به‌منظور آموزش دیکشنری فشرده و متمایز برای کدگذاری تنک ارائه می‌شود. در این مدل با تعیین دیکشنری مطلوب، خطای بازسازی و خطای کدگذاری تنک حداقل سازی شده و معیار عملکرد طبقه‌بندی به‌منظور دسته‌بندی تصاویر OCT شبکیه به سه دسته‌ی نرمال، AMD، DME، در تابع هدف بهینه می‌گردد. در معماری چندلایه‌ی پیشنهادی به‌منظور تحلیل دقیق‌تر ابتدا تصویر ورودی به تکه مای کوچک تجزیه‌شده، سپس به کمک دیکشنری اولیه کدگذاری می‌گردد. ضرایب تنک به‌دست‌آمده در هر لایه به‌عنوان ورودی لایه‌ی بعدی مورد استفاده قرار می‌گیرد و عملیات کدگذاری پس از یافتن دیکشنری بهینه، در چندلایه تکرار می‌گردد. این فرایند تکرار شونده منجر به تمایز بیشتر در ویژگی‌های به‌دست‌آمده و مینیمم سازی خطای بازسازی و کدگذاری تنک در خروجی لایه‌ی آخر می‌شود.

همدوسی نوری) از روی شبکیه و فراهم شدن امکان اخذ اطلاعات و تصاویر با دقت و کیفیت بالا از لایه مای مختلف شبکیه باعث شده است تا باکار بر روی این تصاویر، خطای الگوریتم مای ارائه‌شده جهت تشخیص و استخراج اتوماتیک الگوهای مرتبط با بیماری‌های شبکیه را کمتر نمود، و روش‌ها و الگوریتم مای نوینی در جهت آنالیز اتوماتیک این تصاویر معرفی کرد. تصویربرداری OCT چشم یک روش تصویربرداری از لایه مای مختلف پرده چشم و سر عصب است که با استفاده از امواج نوری، مقاطع تصویربرداری مجزایی را تهیه می‌کند. این تکنولوژی می‌تواند تصویربرداری‌های مقطعی از ساختمان میکروسکوپی بافت‌های زیستی انجام دهد. همچنین دارای رزولوشنی در مقیاس میکرومتر می‌باشد [11].



شکل 1. اسکن OCT در بافت شبکیه چشم در 800 نانومتر و رزولوشن محوری 3 میکرون

Fig 1. OCT scan in retinal tissue at 800 nm and axial resolution of 3 microns

در طول دو دهه گذشته اکثر آثار محققین در زمینه پردازش تصاویر OCT، به دو بخش اصلی اختصاص یافته که بخش اول روش‌های بخش‌بندی و تقطیع لایه مای شبکیه [12-18] است، که در این مقاله مورد بحث قرار نگرفته و بخش دوم از مقالات این حوزه، که هدف ما است به روش‌های گوناگون طبقه‌بندی تصاویر OCT پرداخته است که عموماً مبتنی بر بازسازی تصویر هستند [19-31]. در سال 2011 لیو همکارانش به‌منظور تشخیص آسیب مای ناحیه‌ی ماکولا، از جمله AMD و DME، روشی را پیشنهاد دادند که از الگوهای دودویی محلی استفاده نموده و با استفاده از یک هرم فضایی چندبعدی و تحلیل مؤلفه مای اصلی (PCA) برای کاهش ابعاد، تصاویر را بازسازی کردند [24].

در سال 2012، مدلی برای بازسازی تصاویر بر اساس گراف پیشنهاد شد [25-26] که در این مدل ابتدا، تصاویر به یک مجموعه چهار درختی تقسیم سپس به‌منظور تحلیل این چهار درخت از تکنولوژی کاهش زیرگراف استفاده و با داشتن توانایی تشخیص آن‌ها، زیر گرافی‌های مشترک برای بازسازی تصاویر انتخاب شدند. در نهایت با استفاده از بردارهای ویژگی، طبقه‌بندی کننده، آموزش

خطای بازسازی در فرمول (1)، منجر به یادگیری دیکشنری و بازسازی تنک سیگنال می‌گردد.

هدف اصلی در الگوریتم K-SVD حل معادله‌ی زیر است تا بتوان از روی آن ماتریس دیکشنری D را به دست آورد.

$$\langle D, X \rangle = \operatorname{argmin}_{D, X} \|Y - DX\|_F^2 \quad \forall i, \|x_i\|_0 < T \quad (1)$$

که در آن $D = [d_1, d_2, \dots, d_m] \in \mathbb{R}^{d \times m}$ دیکشنری اولیه، $Y = [y_1, y_2, \dots, y_n] \in \mathbb{R}^{d \times n}$ سیگنال ورودی و $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$ مجموعه ضرایب تنک می‌باشد و m تعداد عناصر دیکشنری و n هم تعداد سیگنال‌های ورودی و d بعد نمونه‌ها که در حالت خاص برای تصویر دوبعدی $d=2$ است. این الگوریتم دارای دو مرحله‌ی اساسی است. یک مرحله مربوط به نمایش تنک و اصلاح آن، مرحله‌ی بعدی اصلاح ستون‌های ماتریس دیکشنری است. فرض کنید در مرحله‌ی اول دیکشنری D ثابت و فرمول (1) نسبت به ماتریس X حل شود. بنابراین نرم فروبنیوس را به صورت زیر می‌توان بازنویسی کرد:

$$\|Y - DX\|_F^2 = \sum_{i=1}^k \|y_i - Dx_i\|_2^2 \quad (2)$$

در نتیجه معادله‌ی (1) به k مسئله به شکل زیر تقسیم می‌شود:

$$\langle D, X \rangle = \operatorname{argmin}_{D, X} \|y_i - Dx_i\|_2^2 \quad \forall i, \|x_i\|_0 < T \quad (3)$$

به اندازه‌ی کافی کوچک انتخاب می‌شود تا تنک بودن x_i ‌ها حفظ شود و تعداد کمی از ضرایب آن غیر صفر باشند. معادله‌ی (3) به روش‌های گوناگون قابل حل است. از آنجایی که یافتن جواب این معادله به صورت بهینه غیرممکن می‌باشد، اغلب روش‌های ارائه شده به صورت حل تقریبی به دست آمده که توسط الگوریتم‌های تعقیبی و یا حریصانه قابل انجام می‌باشد. این الگوریتم‌ها حریصانه اند و اتم‌های دیکشنری پی‌درپی انتخاب می‌گردند. این روش‌ها شامل محاسبه حاصل ضرب درونی سیگنال و اتم‌های دیکشنری‌اند و گاهی نیز راه‌حل‌های حداقل مربعات را به کار می‌گیرند. دسته دیگری از روش‌های تعقیبی وجود دارند که به محدب سازی مسئله تجزیه اتمی، منجر می‌گردند و یا از نرم l_p با $p \leq 1$ به عنوان جایگزین نرم l_0 استفاده می‌نمایند.

اگرچه برای $p < 1$ شباهت به اندازه‌گیری تنکی درست‌تر است، اما مسئله غیر محدب می‌گردد که ممکن است در بهینه‌های محلی گرفتار شده و به راه‌حل‌های غیرقابل اعتمادی در فضای جستجو منجر گردد. در این دسته از روش‌ها از ضرب کننده‌های لاگرانژ به منظور تبدیل قیود به تابع هدف و از روشی تکراری نتیجه شده بر اساس ایده حداقل مربعات دوباره وزن دهی شده

2. یادگیری دیکشنری

انتقال سیگنال‌ها (فضای بردارهای ویژگی) به یک فضای جدید با خواص مطلوب، یک رویکرد رایج برای بهبود کارایی، در اکثر حوزه‌ها از قبیل خوشه‌بندی و ارتقا سیگنال می‌باشد. بردارهای پایه در برخی تبدیل‌ها مثل تبدیل فوریه، ثابت هستند و به زمینه وابسته نیستند. در مقابل، برخی تبدیل‌ها مثل PCA⁶ و نمایش تنک، بردارهای پایه‌ی وابسته به زمینه‌دارند. یعنی بردارهای پایه از نمونه‌هایی که از دامنه به دست می‌آیند، استخراج می‌شوند. عمل استخراج بردارهای پایه در رویکرد نمایش تنک، اغلب یادگیری (آموزش) دیکشنری یا یادگیری کلمه‌کد نامیده می‌شود. اعمال تنظیم‌گر تنک ساز، در مسئله بهینه‌سازی یادگیری دیکشنری منجر به یادگیری اتم‌های معنایی سطح بالا در دامنه می‌گردد. در فرایند طبقه‌بندی مبتنی بر یادگیری دیکشنری و نمایش تنک، یک فرض اولیه این است که، دیکشنری (مجموعه‌ای از پایه‌ها) وجود دارد و همان داده‌های آموزشی است؛ فرض دوم آن است که با استفاده از مدل‌ها و الگوریتم‌های مناسب دیکشنری اولیه را طراحی نمود. هدف، تقریب داده‌های آزمون (تصویر جدید) با استفاده از تعداد کمی (تنک) از تصاویر آموزش است.

به عنوان یک روش ابتدایی در طبقه‌بندی می‌توان برای هر تصویر آزمون، برچسب (نام کلاس) تصویری که در بین داده‌های آموزش به آن شبیه‌تر است را به آن داد. مدل‌های مطرح شده در این حوزه سعی دارند که این الگوریتم ساده را بهبود دهند و به جای قضاوت فقط با یک داده آموزش، تصویر جدید را با مجموعه‌ای از تصاویر آموزش (تعداد کم: تنک) به بهترین وجه (حداقل کردن مربعات خطای بازسازی) تقریب بزنند و سپس با توجه به سهم هر کدام از کلاس‌هایی که تصویر آن‌ها در بازسازی تصویر آزمون نقش داشته، تشخیص بهتری انجام دهند تا نهایتاً تصویر جدید با دقت بیشتری به یکی از دسته‌ها (کلاس‌ها) طبقه‌بندی شود. برای پردازش مناسب‌تر سیگنال می‌توان آن را با استفاده از یک ساختار مشخص که در اصطلاح به آن دیکشنری گفته می‌شود به سیگنالی دیگر تبدیل نمود که تعداد عناصر غیر صفر آن نسبت به سیگنال اولیه بسیار کمتر است. به این عمل نمایش تنک آن سیگنال گفته می‌شود که عموماً گام ابتدایی در الگوریتم‌های یادگیری دیکشنری می‌باشد.

در میان روش‌های گوناگون مبتنی بر یادگیری دیکشنری و نمایش تنک، الگوریتم K-SVD [29] یک الگوریتم تکرارشونده‌ی موفق در حوزه‌ی یادگیری دیکشنری است که با به حداقل رساندن

در قسمت بعد برای یادگیری دیکشنری مناسب از روش‌هایی استفاده می‌شود که مجموعه‌ی داده‌های آموزشی را برای پیدا کردن دیکشنری به کار می‌گیرد و سعی در گزینش بهترین عناصر می‌کند. مرحله‌ی یادگیری مشابه الگوریتم K-SVD است. تفاوت آن در این است که دیکشنری D ورودی Y علاوه بر نمونه‌های آموزشی از لیبل کلاس‌ها نیز برای ارتباط دهی بهتر بین عناصر انتخابی خود و کلاس‌ها استفاده می‌کند.

2.1 آموزش دیکشنری با ناظر

همان‌طوری که گفته شد فرمول اصلی یادگیری دیکشنری و نمایش تنک DLSR، مبتنی بر کمینه کردن خطای بین سیگنال اصلی و نمایش تنک در فضای دیکشنری یافت شده است. اگرچه این فرمول برای حل مسائلی نظیر حذف نویز، فشرده‌سازی و کدگذاری بهینه است، ممکن است منجر به جواب بهینه در مورد طبقه‌بندی نشود، به طوری که آن‌ها نمی‌توانند از اطلاعات برچسب هر کلاس و داده‌های خاص آن، در یادگیری اتم‌های دیکشنری و ضرایب نمایش تنک استفاده کنند. همین امر منجر به تشکیل دسته جدیدی از فن‌ها شده که اطلاعات برچسب را در محاسبه دیکشنری، ضرایب تنک یا هر دو آن‌ها به کار می‌گیرد و یادگیری دیکشنری و نمایش تنک با ناظر نامیده می‌شود [2]، [32]، [34]. به منظور تفسیر بهتر به‌وسیله‌ی مدل‌سازی ریاضی، فرض کنید دوتایی (y_k, l_k) موجود باشد به طوری که $y_k \in Y = \mathbb{R}^{2 \times n}$ سیگنال ورودی دوبعدی و $l_k \in l = [l_1, l_2, \dots, l_N]$ برچسب نظیر آن است. اگر L تابع هزینه‌ی طبقه‌بندی و W پارامترهای آن باشد، مسئله‌ی یادگیری دیکشنری با ناظر را به صورت زیر می‌توان نوشت:

$$\hat{x}_k = \operatorname{argmin}_{x_k} \|y_k - Dx_k\|_2^2 \quad s.t. \quad \forall i, \|x_i\|_0 < T \quad (6)$$

$$\min_{D,W} \sum_{k=1}^n L(l_k, \hat{x}_k, W) \quad (7)$$

که در آن $D = [d_1, d_2, \dots, d_m] \in \mathbb{R}^{2 \times m}$ ماتریس دیکشنری و $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$ ضرایب تنک سیگنال‌های ورودی و Y و T حد آستانه و فاکتور تنظیم تنکی می‌باشند. $\|Y - DX\|_2^2$ خطای بازسازی نامیده می‌شود و بامی نیم کردن آن D و به دنبال آن شرایط تنکی حاصل می‌گردد.

در این فرایند سیگنال ورودی $y_k \in Y$ (متناظر با برچسب l_k) به‌وسیله‌ی دیکشنری D کدگذاری می‌شود و منتج به کد تنک \hat{x}_k می‌گردد که به‌عنوان ورودی تابع هزینه با پارامترهای W مورد استفاده قرار می‌گیرد. تابع هزینه $L(l_k, \hat{x}_k, W)$ محاسبه و

تکرار شونده که نرم l_p را به‌عنوان یک نرم l_2 وزن داده شده کنترل می‌نماید، استفاده می‌گردد [40, 42, 44].

در نهایت با حل معادله‌ی (3) ماتریس ضرایب X را به صورت زیر خواهد بود:

$$X = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \dots & x_{mn} \end{bmatrix} \quad (4)$$

ستون‌های X، ضرایب تنک هر سیگنال ورودی است. بنابراین ترکیبی از عناصر دیکشنری برای هر سیگنال ورودی توسط سطرها نشان داده می‌شود. برای مرحله‌ی دوم فرض می‌شود هر دو ماتریس X و D ثابت باشند و فقط یک ستون از ماتریس دیکشنری $d_i \in D$ متغیر باشد. با استفاده از معادله‌ی زیر اصلاح و به‌روزرسانی ستون‌های ماتریس D را آغاز می‌گردد:

$$E^i = Y - \sum_{\substack{j=1 \\ j \neq i}}^m d_j x_j^i \quad (5)$$

که E^i اختلاف بین سیگنال ورودی و ترکیب بازسازی شده‌ی سیگنال ورودی با استفاده از عناصر دیکشنری است. در معادله‌ی (5)، x_j^i زامین ردیف از ماتریس X است. در واقع ماتریس E^i میزان خطا را برای زمانی محاسبه می‌کند که i امین عنصر از رابطه حذف شده باشد.

در مرحله‌ی بعد هر ستون e^h که ضریب تنک متناظرش صفر است، یعنی $x_{j,h} = 0, 1 \leq h \leq n$ ، از E^i حذف می‌شود.

در آخر عملیات تجزیه مقدار منفرد (SVD) روی $E^i = U\Delta V^T$ انجام و d_i به u_1 و ضریب تنک نظیرش به v_1 تبدیل می‌شود. این روند برای عناصر بعدی دیکشنری نیز با اضافه کردن i و محاسبه مجدد معادله‌ی (5) و اجرای دوباره‌ی SVD ادامه می‌یابد؛ بنابراین تمام دیکشنری به‌روزرسانی می‌شود. با محاسبه مجدد ضرایب تنک به‌دفعات، فرایند کلی تا رسیدن به دیکشنری مطلوب تکرار می‌گردد.

طبقه‌بندی مبتنی بر نمایش تنک تا حد زیادی وابسته به انتخاب دیکشنری مناسب است؛ به طوری که اگر بر اساس عناصر کلاس‌های مختلف و به صورت بهینه انتخاب شود، بسیار قدرتمند خواهد بود. در این بخش دیدیم که الگوریتم مای غیر نظارتی مانند الگوریتم K-SVD، هیچ تمایزی بین عناصر دیکشنری قائل نیستند بنابراین می‌توانند با سایر الگوریتم‌های مای قدرتمند طبقه‌بندی مانند SVM⁷ نیز مورد مقایسه قرار بگیرند. در حالی که ما اطلاعات خوبی درباره کلاس‌ها و عناصر آن‌ها داریم که می‌توانیم از آن‌ها در بازنمایی تنک بهتری از سیگنال، بهره‌مند شویم.

روش‌های کاهش نویز در تصاویر OCT ممکن است در سیگنال OCT و قبل از تولید اندازه سیگنال صورت گیرد که به این دسته از روش‌ها، روش‌های "حوزه مختلط" یا "مبتنی بر سخت‌افزار" اطلاق می‌گردد. اما اگر کاهش نویز بر روی اندازه سیگنال OCT تداخلی صورت پذیرد، این روش‌ها را "حوزه اندازه" یا "مبتنی بر نرم‌افزار" می‌نامند. روش‌های حوزه مختلط را می‌توان به دودسته "روش‌هایی برای بهبود بخشیدن به سیستم نوری" و "سازگار نمودن شیء مورد تصویربرداری" تقسیم نمود. "روش‌های حوزه اندازه" را نیز می‌توان به دودسته "روش‌های حوزه تصویر خام" و "روش‌های مبتنی بر نمایش تُنک" تقسیم نمود. "روش‌های حوزه تصویر خام" شامل تعدادی از راه‌حل‌های معمول در پردازش تصویر مانند فیلترهای پایین‌گذر و فیلترهای متوسط‌گیری وینر، فیلترهای مدین پایین‌گذر، هموارسازی خطی می‌باشند [27]. تعدادی از روش‌های پیشرفته در پردازش تصویر نیز زیرمجموعه روش‌های حوزه تصویر خام می‌باشند، مانند: روش‌های تنظیم واگرایی آی، فیلترهای انتشار مختلط، فیلترهای های ناهمسانگردی غیرخطی، فیلترهای جهت‌دار، توابع کرنل با مقادیر برداری وقفی، روش‌های ماشین بردار پشتیبان و تخمین مای بی‌زین.

روش‌های مبتنی بر نمایش تُنک نیز در حذف نویز تصاویر OCT از جمله روش‌های نوین و بسیار پرکاربرد است که به دودسته "روش‌های غیر پارامتری" و "روش‌های پارامتری" تقسیم می‌شوند. در روش‌های غیر پارامتری می‌توان به PCA مقاوم و حذف نویز بر اساس تُنک بودن داده‌ها در فضای تبدیل اشاره نمود.

در روش‌های پارامتری نیز می‌توان به روش‌های مبتنی بر موجک، تبدیل‌های موجک مختلط دودرختی، تبدیل کرولت و پیاده‌سازی انتشار ویولت با مدل مخلوط لاپلاسی اشاره نمود. همچنین باید خاطرنشان کرد که تعداد زیادی از کارهای تحقیقاتی پیش از مرحله تقطیع تصویر از هیچ پیش‌پردازش خاصی کمک نمی‌گیرند. در بین روش‌های مبتنی بر نمایش تُنک با توجه به اینکه مبنای اصلی کار ما در طبقه‌بندی تصاویر OCT، الگوریتم KSVD است، در پیش‌پردازش نیز بنای کار را بر اساس این الگوریتم قرار می‌دهیم و به تشریح آن می‌پردازیم.

از آنجایی که عموماً در کاربردهای پردازش تصویر تعداد پیکسل‌ها بسیار زیاد می‌باشد و اعمال الگوریتم روی تمام پیکسل‌ها به روش ذکر شده بسیار زمان‌بر بوده و به دلیل از دست رفتن اطلاعات محلی، غیردقیق نیز می‌باشد، معمولاً در پردازش تصاویر از تکه‌های کوچک‌تری با اندازه از پیش تعیین شده استفاده می‌گردد که با / بدون همپوشانی از تصویر انتخاب می‌گردند.

سپس دیکشنری D و پارامترهای W برای انطباق با مسئله‌ی طبقه‌بندی به‌روز می‌شوند. هدف فرمول (7) به حداقل رساندن تابع هزینه L با توجه به پارامترهای D و W است که به‌روزرسانی شدند. به‌منظور می‌نیمم‌سازی تابع هزینه می‌توان از روشی مشابه با الگوریتم پس انتشار خطا [35] رایج در شبکه مای عصبی استفاده نمود.

محاسبه‌ی گرادیان L با توجه به پارامتر مای W معمولاً ساده است. مشکل اصلی، کمینه‌سازی تابع هزینه‌ی L با توجه به دیکشنری D است؛ زیرا D به‌طور صریح و روشن وجود ندارد و با یک مسئله‌ی بهینه‌سازی دیگر درگیر می‌شود. یک‌راه غلبه بر این مشکل محاسبه‌ی گرادیان تابع هزینه L بر اساس دیکشنری D مفروض است؛ این موضوع در [2] و [4] مورد بررسی قرار گرفته است. با توجه به این موضوع گرادیان مطلوب به‌صورت زیر می‌تواند محاسبه شود:

$$\nabla_D L(l_k, \hat{x}_k, W) = D\beta \hat{x}_k^T + (y_k - D\hat{x}_k)\beta^T \quad (8)$$

اگر فرض کنیم $\Lambda = \{i | x_i \neq 0\}$ مجموعه اندیس مای ضرایب غیر صفر متناظر با کد تنک \hat{x}_k باشد. اگر $\beta \in \mathbb{R}^K$ یک بردار باشد، در این صورت β_Λ به شکل زیر تعریف می‌گردد:

$$\beta_\Lambda = (D_\Lambda^T D_\Lambda)^{-1} \nabla_{\hat{x}_{k_\Lambda}} L(l_k, \hat{x}_k, W) \quad (9)$$

که \hat{x}_{k_Λ} ضرایب غیر صفر متناظر با کد تنک \hat{x}_k می‌باشد و $\beta_j = 0$ اگر $j \notin \Lambda$.

2.2. پیش‌پردازش تصاویر

تصاویر OCT شامل حجم عظیمی از اطلاعات هستند. در نتیجه تحلیل غیر اتوماتیک و چشمی این داده‌ها برای چشم‌پزشک تقریباً ناممکن می‌باشد. یک عامل مؤثر در مطرح‌شدن بحث‌های مربوط به پیش‌پردازش اتوماتیک این تصاویر، فراهم نمودن یک ابزار کمی مناسب برای کمک به چشم‌پزشکان در تشخیص و مونیتور نمودن بیماری‌های وابسته به چشم می‌باشد. از آنجایی که حجم زیادی از اطلاعات در چنین تصاویری نهفته است، اثربخشی چنین نمود اتوماتیکی می‌تواند در بسیاری از کاربردها به‌وضوح قابل‌مشاهده باشد. تصاویر OCT دقیقاً مشابه تصاویر ماوراء صوت، با نویز اسپکل آلوده می‌گردند؛ نویز اسپکل یک نویز ضرب شونده است و در اکثر سیستم‌های تصویربرداری که بر مبنای تصویرگری همدوس باشند، مشاهده می‌گردد. وجود این نویز باعث امکان تفسیرهای نادرست بر روی این داده‌ها می‌شود. بنابراین غربالگری چندبعدی و حذف نویز یکی از مهم‌ترین عملیات پیش‌پردازی بر روی تصاویر OCT می‌باشد [27].

برای یافتن تعداد مناسب برای تکرار الگوریتم به‌روزرسانی فوق (تعداد داده‌های مورد استفاده در ساخت دیکشنری ارتقا یافته) الگوریتم زیر پیشنهاد شده است:

(a) تعداد نمونه داده‌ها را $k=0$ فرض کن

(b) $\text{Var}_{\text{old}} = \text{var}_{\text{new}} = 100$

(c) $\text{Var}_{\text{old}} = \text{var}_{\text{new}} = k+1$

(d) برای $i=1:10$

1. تعداد k داده را به صورت تصادفی انتخاب کن

2. دیکشنری ارتقا یافته را برای k داده بساز

3. از دیکشنری ارتقا یافته برای حذف نویز یک داده مشخص (مثلاً داده شماره 1) استفاده کن.

4. نسبت کنتراست به نویز CNR را برای نتیجه‌ی حاصل محاسبه کن

5. $\text{Temp}(i) = \text{CNR}$

(e) $\text{Var}_{\text{new}} = \text{variance}(\text{Temp})$

(f) اگر $|\text{var}_{\text{new}} - \text{var}_{\text{old}}| > \varepsilon$ برو به (d)

(g) K تعداد داده‌های مورد نیاز برای ساخت دیکشنری ارتقا یافته می‌باشد.

یکی دیگر از عملیات پیش‌پردازشی مؤثر در تحلیل تصاویر OCT عملیات مسطح سازی تصاویر است. تصحیح انحنای مجموعه سلول‌های دارای بازتابندگی بالا یعنی HRC، هدف اصلی در الگوریتم‌های مسطح سازی به شمار می‌آید. این مجموعه سلول‌ها به بخش‌های داخلی و خارجی سلول‌های نور پذیر به انضمام سلول‌های رنگ‌دانه‌ای شش ضلعی اطلاق می‌گردد. با فرض اینکه انحنای موجود، ناشی از اعوجاج زمان-فرکانس دستگاه اخذ کننده داده و یا انحنای ناشی از شکل منحنی وار چشم یا حرکت بیمار باشد، حذف این انحنای باعث تنظیم امتداد مورد نیاز در تصویر خواهد شد. بررسی کارهای انجام شده در این حوزه ما را به سه دیدگاه در این حوزه رهنمون می‌سازد. در روش نخست اسرینی و همکارانش در [30] با سوار کردن لایه‌ی RPE بر روی چند جمله‌ای مرتبه‌ی دوم نواحی شبکیه را هم‌تراز و سپس شبکیه را مسطح کردند، که اگر لایه‌ی RPE شبکیه در حد زیادی آسیب دیده باشد (شکل 2 قسمت b) به علت عدم تطابق با یک منحنی مناسب دچار اشکال می‌شود.

اگر هر تکه‌ی کوچک را $x_{ij} = R_{ij}X$ با اندازه‌ی $\sqrt{n} \times \sqrt{n}$ در نظر بگیریم، ماتریس سازنده‌ی تکه‌ی کوچک یعنی R_{ij} دارای اندازه‌ی $n \times N$ بوده و با اعمال روی تصویر با اندازه‌ی $\sqrt{N} \times \sqrt{N}$ ، بلوک ij را از آن استخراج می‌کند.

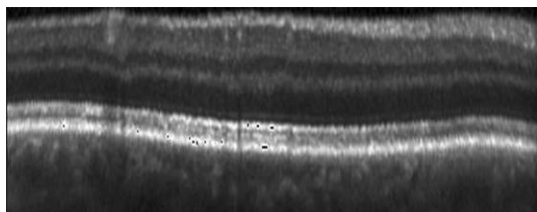
در این صورت الگوریتم K-SVD را می‌توان به صورت زیر تعمیم داد. در مرحله‌ی اول فرض می‌کنیم دیکشنری D نامعلوم است و مسئله به صورت زیر مطرح می‌شود:

$$\{\hat{D}, \hat{a}_{ij}, \hat{X}\} = \arg \min_{\hat{D}, a_{ij}, X} \gamma \|X - Y\|_2^2 + \sum_{ij} \mu_{ij} \|a_{ij}\|_0 + \sum_{ij} \|Da_{ij} - R_{ij}X\|_2^2 \quad (10)$$

که در آن عبارت اول تضمین‌کننده‌ی نزدیکی بین نسخه‌ی نویزی تصویر Y و نسخه‌ی بعد از حذف نویز (نامعلوم) X می‌باشد. عبارات دوم و سوم نمایانگر تنگ بودن ضرایب و نزدیکی پاسخ بازنمایی شده با دیکشنری به هر تکه‌ی کوچک می‌باشد. با در نظر گرفتن هر تکه‌ی کوچک به صورت $x_{ij} = R_{ij}X$ ، ضرایب a_{ij} تنگ در بازنمایی آن تکه‌ی کوچک بر روی دیکشنری D می‌باشد. در مرحله‌ی دوم فرض می‌شود D و X ثابت هستند و بازنمایی تنگ محاسبه می‌گردد. با در دست داشتن مقادیر ضرایب بازنمایی، دیکشنری با استفاده از روش K-SVD به روز می‌شود.

برای ساخت دیکشنری پایه تخصیص یافته، می‌توان از دو روش استفاده نمود. روش اول، انتخاب تصادفی یک نمونه تصویر از بین نمونه‌های موجود و اعمال آموزش دیکشنری بر روی آن نمونه خاص می‌باشد. نتایج حاصل نشان می‌دهد که نمونه‌های مختلف، منجر به نتایج حذف نویز مختلفی می‌شود، این موضوع ممکن است تعبیر صورت گرفته از عملکرد یک روش را مورد تأثیر منفی قرار دهد. بنابراین، روش دوم برای حل این مشکل پیشنهاد گردید که در آن، از بین مجموعه داده مربوطه، تعدادی داده انتخاب شده، آموزش روی آن‌ها صورت می‌گیرد و تعداد مورد انتخاب به نحوی می‌باشد که کمترین وابستگی به انتخاب داده‌های مجزا در آن وجود داشته باشد. برای این منظور، ابتدا یک دیکشنری اولیه انتخاب شده، آموزش دیکشنری روی داده اول صورت می‌پذیرد. دیکشنری حاصل، سپس به عنوان دیکشنری اولیه برای آموزش یک دیکشنری دیگر (دیکشنری به روز شده) بر روی داده بعدی به کار می‌رود. این روند تا رسیدن به یک نسخه دیکشنری ارتقا یافته برای استفاده در حذف نویز تمام داده‌های متعلق به آن کلاس به کار می‌رود [27].

و همکاران در سال (2017) [33] ارائه دادند عملیات هم‌ترازی انجام می‌شود به این ترتیب که، ابتدا تصویر شبکه‌بندی بدون برآورد مرز RPE دریافت می‌شود. در مرحله‌ی دوم، دو مجموعه داده از نقاط یک تصویر مفروض (برای انطباق خطی و منحنی) استخراج و سپس به‌طور خودکار یک مجموعه از نقاط که بیشتر نمایانگر مورفولوژی شبکه‌بندی است، انتخاب می‌گردد. نهایتاً به‌طور خودکار بین یک چندجمله‌ای مرتبه دوم و یک خط مستقیم جهت انطباق با مجموعه نقاط انتخاب‌شده، تصمیم‌گیری می‌شود.

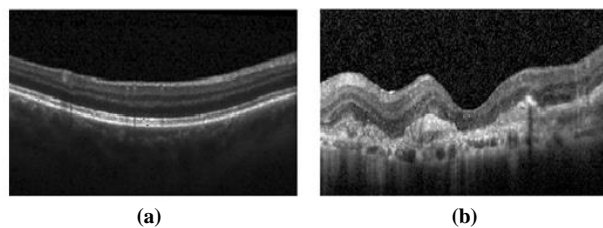


شکل 5. تصویر هم‌تراز شده‌ی OCT
Fig 5. Aligned OCT image

2.3. تجزیه پیچ و جمع‌آوری

روش‌های یادگیری دیکشنری برای بازسازی یا طبقه‌بندی تصاویر کامل یا تکه‌های کوچک تصاویر استفاده می‌شوند. به این معنی که در عمل سیگنال‌های ورودی y_k می‌توانند یک تصویر یا تکه‌های کوچکی از آن تصویر در قالب بردارهای ستونی، حاوی مقادیر پیکسل باشد. یک مدل یادگیری دیکشنری با ناظر می‌تواند به نحو موفقیت‌آمیزی الگوها را باهدف طبقه‌بندی به‌خوبی یاد بگیرد. با این حال در عمل، برخی محدودیت‌ها ممکن است مشاهده شود، بهترین اثربخشی این روش‌ها زمانی است که تصاویر ورودی نسبتاً کوچک و به‌صورت محلی همگن باشند؛ در غیر این صورت، دیکشنری مورد استفاده به‌منظور تجزیه‌ی تنگ کارآمد به تعداد اتم‌های زیادی احتیاج دارد.

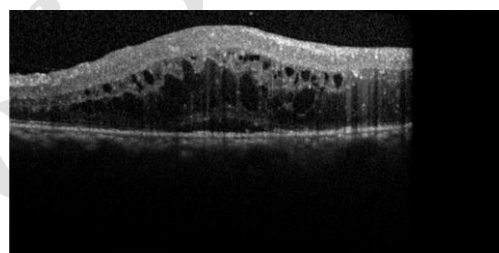
علاوه بر این، طبقه‌بندی مجموعه‌ای از تکه‌های استخراج‌شده از یک تصویر به‌جای خود تصویر، از نقطه‌نظر روش‌های یادگیری دیکشنری متفاوت است. در واقع، هنگام برخورد با تکه‌ها، روشی برای به دست آوردن اطلاعات درباره مجموعه‌ای از تکه‌ها مورد نیاز است. این مسئله خاص توسط یانگ و همکاران مورد مطالعه قرار گرفته است [17]. آن‌ها پیشنهاد دادند مرحله‌ی کدگذاری تنگ در سطح پیچ‌ها با سایز کوچک‌تر از تصویر انجام شود و به دنبال آن نیز مراحل جمع‌آوری چندگانه به‌منظور فرایند مؤثرتر کاهش ابعاد و به دست آوردن بازنمایی، در چند مقیاس ارائه شود. برای مواجهه با تصاویر بزرگ، ما تابع جمع‌آوری را معرفی می‌کنیم. عملگر جمع‌آوری به‌منظور قابلیت اطمینان و تغییرناپذیری ویژگی‌های تصویر استفاده می‌شود و رویکردی مؤثر



شکل 2. تصاویر OCT از شبکه‌بندی: (a) نمایی از لایه‌ی RPE واضح (b) لایه‌ی RPE تخریب‌شده

Fig 2. OCT images of the retina: (a) View of the clear RPE layer (b) the RPE layer damaged

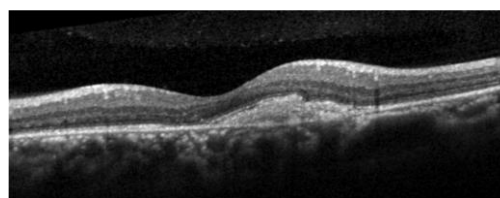
در روش دوم، لیو همکاران [24] با استفاده از انطباق کل تصویر OCT شبکه‌بندی با یک چندجمله‌ای مرتبه دوم نواحی شبکه‌بندی را هم‌تراز می‌کنند. اگرچه روش هم‌ترازی آن‌ها می‌تواند برای تصاویر شبکه‌بندی که RPE آن‌ها تخریب‌شده کارساز باشد ولی نقطه‌ضعف روش موجود این است که اثربخشی آن بر روی نوع خاصی از بیماری شبکه‌بندی که در آن بخش آسیب‌دیده متورم شده است (شکل 3)، تا حد زیادی کاهش می‌یابد.



شکل 3 تصویر OCT شبکه‌بندی که در آن بخش آسیب‌دیده متورم شده است.

Fig 3. The retinal OCT image in which the damaged part swells.

سوم، تصاویر OCT زیادی وجود دارند که در آن‌ها لایه‌ی RPE شبکه‌بندی مستقیم است اما با یک زاویه مشخص با خط افقی قرار گرفته است (شکل 4). در این حالت، انطباق شبکه‌بندی با خط مستقیم ممکن است بهتر از یک سهمی باشد.



شکل 4. تصویر OCT شبکه‌بندی که در آن لایه‌ی RPE شبکه‌بندی مستقیم و با یک زاویه مشخص با خط افقی قرار گرفته است.

Fig 4. The retinal OCT image, in which the RPE layer of the retina is straight and at a given angle with a horizontal line.

برای حل مشکلات سه‌گانه ذکر شده در بالا و هم‌ترازی شبکه‌بندی استفاده از روشی پایدارتر، در این مقاله بر مبنای روشی که یانگویی

الگوریتم مای خوشه‌بندی و پیدا کرده مراکز کلاس‌ها محقق می‌گردد. کدگذاری یک بردار با استفاده از یک دیکشنری، شبیه به تصویرسازی یک بردار در یک فضای جدید است. این تصویرسازی غیرخطی است و نتیجه‌ی آن بردار تنک می‌باشد. سپس این بردار تنک به‌عنوان ورودی جدید برای لایه بعدی استفاده می‌شود. به همین ترتیب ورودی هر لایه با استفاده از کد تنک‌لایه‌ی قبل شکل می‌گیرد. بنابراین، اضافه کردن یک‌لایه‌ی جدید برای کدگذاری تنک با استفاده از دیکشنری به‌روز شده، مشابه تصویرسازی دیگری در یک سیستم مختصات جدید است و هر دیکشنری به روش یادگیری با ناظر شکل می‌گیرد. در معماری مذکور، این فرآیند را می‌توان چندین بار تکرار کرد که در هر بار تکرار، لایه‌ی جدید با دیکشنری جدید طراحی می‌شود.

2.4.1 چارچوب کلی روش پیشنهادی

روش پیشنهادی چندلایه را می‌توان به شرح زیر توصیف کرد:

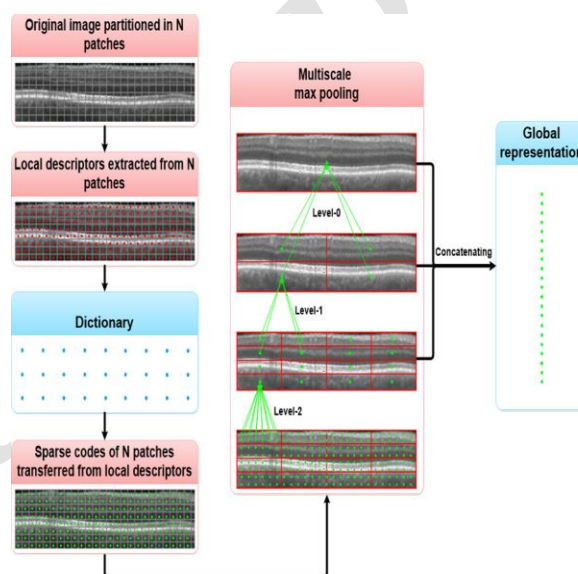
1. ابتدا هر تصویر ورودی به مجموعه‌ای از تکه مای همپوشانی کوچک γ_k با حفظ محلی سازی مکانی خود، تجزیه می‌شوند.
2. هر پچ γ_k با استفاده از اولین دیکشنری D کدگذاری می‌شود. از آنجایی که محلی سازی مکانی این تکه‌ها حفظ شده است، مجموعه‌ی کدهای تنک \hat{x}_k را می‌توان به‌عنوان بردار 3 بعدی X با عمق برابر با تعداد اتم‌ها در دیکشنری D در نظر گرفت.
3. بردار 3 بعدی حاصل به‌عنوان ورودی لایه بعدی مورد استفاده قرار می‌گیرد.

(1) و (2) را می‌توان برای هر تعدادی از لایه‌ها تکرار کرد.

برای بهینه‌سازی معماری چندلایه باهدف طبقه‌بندی، پیدا کردن دیکشنری مای بهینه در هر لایه لازم است و جهت افزایش دقت طبقه‌بندی و کارایی آن اهمیت زیادی دارد. در این معماری برای به دست آوردن بهترین دیکشنری در اولین لایه، با استفاده از الگوریتم K-maens در هر کلاس عملیات خوشه‌بندی را انجام داده و مراکز خوشه‌ها به‌عنوان اتم مای دیکشنری انتخاب می‌شوند. در آخر با جمع دیکشنری مای حاصل از کلاس‌ها دیکشنری بهینه‌ی D در لایه‌ی اول تشکیل می‌گردد. دیکشنری مای دیگر با به‌روزرسانی دیکشنری اولیه در اتمام فرایند هر لایه ایجاد می‌شود. در این فرایند از برچسب کلاس‌ها در طراحی دیکشنری استفاده می‌شود و عناصر شاخص هر کلاس به‌عنوان اتم‌های دیکشنری در فرایند طبقه‌بندی دخالت پیدا می‌کنند. چارچوب چندلایه ما، با مبنا قرار دادن الگوریتم K-SVD سازگار با برچسب، با استفاده از بهینه از برچسب کلاس‌ها و انتخاب دیکشنری مناسب، فرایند

برای کاهش ابعاد و روشی برای جمع‌کردن ویژگی‌های محلی فضایی در یک تصویر است.

طبقه‌بندی تصویر I را در نظر بگیرید. ابتدا I را به تکه مای همپوشانی محلی γ_k تجزیه می‌کنیم که به‌وسیله دیکشنری D و محاسبه‌ی ضرایب \hat{x}_k تنک کدگذاری می‌شوند. این تکه‌ها به‌طور مساوی بر روی یک توری مشبک محلی سازی می‌شوند، به‌طوری که موقعیت نسبی از تمام تکه‌ها ایجاد گردد. سپس کدهای \hat{x}_k در یک یا چند مرحله، روی گروه کوچکی از تکه‌ها به‌وسیله‌ی تابعی ترکیبی می‌شوند (مثلاً تابع ادغام بیشینه) به‌طوری که در هر مرحله منجر به کاهش تعداد تکه مای کوچک می‌گردد (به شکل 6 نگاه کنید).



شکل 6. بازنمایی تصویر با استفاده از عملیات تجزیه به تکه مای کوچک

و جمع‌آوری با تابع ماکزیمم [10]

Fig 6. Image representation using patch decomposition and pooling with maximum function [10]

2.4 روش پیشنهادی: یادگیری دیکشنری چندلایه مبتنی

بر الگوریتم K-SVD

ایده اصلی روش ما این است که لایه کدگذاری تنک را تکرار کنیم و از K-SVD سازگار با برچسب استفاده کنیم تا تبعیض پذیری ویژگی‌ها را افزایش دهد. هدف دیگر روش پیشنهادی این است که اندازه ابعاد داده ورودی را با کاهش دادن کدگذاری تنک تصویر اصلی کنترل کند. به این ترتیب که به جای کدگذاری تصویر اصلی از کدگذاری تنک تکه مای کوچک از تصویر اصلی که می‌تواند عملیات پردازش و تحلیل را بسیار مؤثرتر نماید، استفاده می‌کند. سومین ایده‌ی این روش پیدا کردن دیکشنری بهینه جهت ارتقا تأثیرگذاری فرایند طبقه‌بندی می‌باشد که این مهم با استفاده از

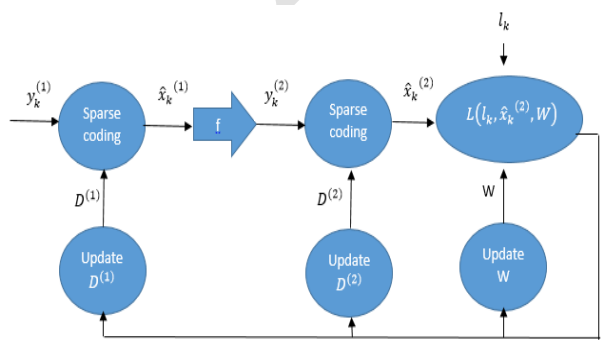
در این فرمول A نیز یک ماتریس تبدیل خطی می‌باشد؛ ما تبدیل خطی $g(X, A) = AX$ را در نظر می‌گیریم که کد مای تنک معمولی X را به کد مای تنک متمایز در فضای \mathbb{R}^m منتقل می‌کند. عبارت $\|Z^{(q)} - Ax^{(q)}\|_2^2$ در فرمول (11) میزان خطای کدگذاری تنک متمایز را مشخص می‌کند که باعث می‌شود کد مای تنک X در هر لایه، هر چه بیشتر به کدهای تنک متمایز Z نزدیک شود و نیز منجر به شبیه‌تر شدن بازنمایی تنک سیگنال مای مربوط به هر کلاس می‌شود و این یعنی اثربخشی "سازگاری برچسب" و کار آیی برچسب کلاس‌ها در کدهای تنک و در ادامه‌ی آن بهبود عملکرد طبقه‌بندی با استفاده از طبقه‌بندی کننده‌ی خطی ساده. برای بهینه‌سازی تابع هزینه‌ی طبقه‌بندی با توجه به دیکشنری‌های هر لایه، ما از الگوریتم پس انتشار خطا استفاده نموده‌ایم [35]؛ بنابراین لازم است گرادبان نسبت به دیکشنری مای آخر یعنی $\hat{x}_k^{(q)}$ به‌عنوان معیارها و شاخص مای طبقه‌بندی استفاده کرده‌ایم و تابع هزینه‌ی طبقه‌بندی روی مجموعه‌ی داده‌های آموزشی ورودی از n تصویر کمینه می‌شود (شکل 8 را ببینید).

همچنین برای میزان تلفات طبقه‌بندی با توجه به نتایج خوبی که در طبقه‌بندی مای چند کلاسه از تابع کراس آن‌تروپی دیده شده است، این تابع را انتخاب نموده‌ایم و طبقه‌بندی کننده‌ی منتخب برای این خروجی را تابع سافت مکس در نظر می‌گیریم. اگر ما مسئله‌ی طبقه‌بندی را با C کلاس فرض کنیم، تابع کراس آن‌تروپی به‌صورت زیر محاسبه می‌گردد:

$$L(l_k, \hat{x}_k^{(q)}, W) = -\sum_{i=1}^C l_{ik} \log(p_{ik}) \quad (13)$$

که در آن p_{ik} به‌صورت زیر تعریف می‌شود:

$$p_{ik} = \frac{\exp(\hat{x}_k^{(q)T} w_i)}{\sum_{j=1}^C \exp(\hat{x}_k^{(q)T} w_j)} \quad (14)$$



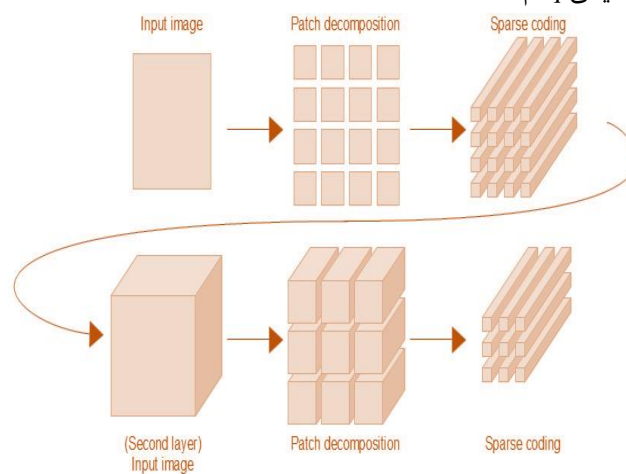
شکل 8. معماری پیشنهادی با دو لایه

Fig 8. Proposed structure with 2 layers.

طبقه‌بندی را تا حد زیادی تأثیرگذار می‌نماید؛ شکل 7 نمای کلی از این فرایند است.

2.4.2 مدل‌سازی ریاضی روش پیشنهادی

فرض کنید $Y = \{y_1^{(1)}, y_2^{(1)}, \dots, y_p^{(1)}\}$ مجموعه‌ی تکه مای کوچک ورودی باشد که از تجزیه تصویر ورودی به p تکه به‌دست آمده است. اندیس بالا در مجموعه‌ی فوق معرف اولین لایه است و در حالت کلی $y_k^{(q)}$ نمایانگر k امین تکه از سیگنال ورودی لایه‌ی q ام است.



شکل 7. معماری دو لایه‌ی روش پیشنهادی

Fig 7. Two-layered architecture of the proposed method

الگوریتم پیشنهادی با Q لایه را می‌توان به‌صورت زیر نشان داد:

$$(\hat{x}_k^{(q)}, A, D) = \operatorname{argmin}_{x, A, D} \|y_k^{(q)} - Dx^{(q)}\|_2^2 + \alpha \|Z^{(q)} - Ax^{(q)}\|_2^2 \quad s.t. \quad \forall i. \|x_i\|_0 < T \quad (11)$$

$$\min_{D^{(1)}, D^{(2)}, \dots, D^{(Q)}, W} \sum_{k=1}^n L(l_k, \hat{x}_k^{(Q)}, W) \quad (12)$$

که $y_k^{(q)}$ به‌عنوان ورودی لایه‌ی q ام، همان خروجی لایه‌ی قبل خود یعنی $\hat{x}_k^{(q-1)}$ است که در واقع نمایش تنک سیگنال ورودی لایه‌ی قبل یعنی $q-1$ بوده و با بکار بردن تبدیل f که ما آن را تابع ادغام بیشینه اختیار نموده‌ایم، بدست می‌آید. همچنین α میزان تأثیر خطای بازسازی و سازگاری لیبیل را کنترل می‌کند.

در فرمول فوق $Z^{(q)} = \{z_1^{(q)}, z_2^{(q)}, \dots, z_n^{(q)}\} \in \mathbb{R}^{m \times n}$ کد مای تنک متمایز سیگنال ورودی $y_k^{(q)}$ در لایه‌ی q ام برای طبقه‌بندی هستند؛ و نیز $z_i^{(q)} = \{0, \dots, 1, 1, \dots, 0\}^t \in \mathbb{R}^m$.

مقدار یک برای $z_i^{(q)}$ زمانی اتفاق می‌افتد که مؤلفه‌ی $y_k^{(q)}$ و متناظر با آن، اتم $d_k^{(q)}$ از دیکشنری D دارای یک برچسب باشند و در غیر این صورت مقدار صفر برای آن در نظر گرفته می‌شود.

$$\nabla_{\hat{x}_{k\Lambda}^{(q)}} L(l_k, \hat{x}_k^{(q)}, W) = 0 \quad (20)$$

در بخش بعد، ما رویکرد بهینه‌سازی این روش را توضیح می‌دهیم؛ به این ترتیب که در طول آموزش A ، D و X ابتدا محاسبه می‌شوند و سپس ماتریس W برای طبقه‌بندی آموزش داده می‌شود.

2.4.4 رویکرد بهینه‌سازی

به منظور پیدا کردن راه‌حل بهینه برای تمام پارامترها به صورت هم‌زمان، مبنای عملکرد ما الگوریتم K-SVD قرار داده شده است، لذا معادله‌ی (11) را به صورت زیر بازنویسی می‌کنیم:

$$\begin{aligned} & \arg\min_{x,A,D} \left\| \left(\frac{y_k^{(q)}}{\sqrt{\alpha Z^{(q)}}} \right) - \left(\frac{D}{\sqrt{\alpha A}} \right) x^{(q)} \right\|_2^2 \\ & s.t \quad \forall i. \|x_i\|_0 < T \end{aligned} \quad (21)$$

حال اگر فرض کنیم، $y_{new} = \left(\frac{y_k^{(q)}}{\sqrt{\alpha Z^{(q)}}} \right)$ ، $c = \left(\frac{D}{\sqrt{\alpha A}} \right)$ و $X = x^{(q)}$ حل مسئله بهینه‌سازی فوق به حل معادله‌ی زیر منجر می‌گردد:

$$\begin{aligned} & (X, D) = \arg\min_{X,D} \|y_{new} - D_{new} X\|_2^2 \\ & s.t \quad \forall i. \|x_i\|_0 < T \end{aligned} \quad (22)$$

و این دقیقاً مسئله‌ی است که K-SVD آن را حل می‌کند [4][21] و [29]. با توجه به این الگوریتم، d_k و ضرایب نظیرش یعنی k امین سطر در X که به صورت x_R^k معرفی می‌شود، در هر لحظه به روز می‌شوند.

با در نظر گرفتن $E_k = (Y - \sum_{j \neq k} d_j x_R^j)$ ، \tilde{x}_R^k و نتیجه‌ی حذف عناصر صفر از x_R^k و E_k خواهند بود. همچنین d_k و \tilde{x}_R^k نیز به صورت زیر محاسبه می‌گردند:

$$(\tilde{x}_R^k, d_k) = \arg\min_{\tilde{x}_R^k, d_k} \|\tilde{E}_k - d_k \tilde{x}_R^k\|_F^2 \quad (23)$$

عملگر SVD روی \tilde{E}_k اجرا می‌شود یعنی: $U \Sigma V^t = SVD(\tilde{E}_k)$ سپس d_k و \tilde{x}_R^k به صورت زیر محاسبه می‌شوند:

$$\begin{aligned} d_k &= U(:,1) \\ \tilde{x}_R^k &= \Sigma(1,1)V(:,1) \end{aligned} \quad (24)$$

و نهایتاً \tilde{x}_R^k ، به جای مقادیر غیر صفر x_R^k ، استفاده می‌شود. از جمله ویژگی‌های مهم این روش آن است که D ، X و W به طور هم‌زمان یاد گرفته می‌شوند، این روش از مینیمم سازی محلی

2.4.3 محاسبه‌ی گرادیان‌ها

برای استفاده از الگوریتم پس انتشار خطا یعنی محاسبه‌ی هر گرادیان با استفاده از قانون مشتق‌گیری زنجیره‌ای، گرادیان به همان شیوه‌ای که در معادله‌ی (8) و (9) از بخش 1-2 انجام شد و با جانشینی $\hat{x}_k^{(q)}$ به جای x_k و $y_k^{(q)}$ به جای y_k محاسبه می‌گردد.

برای دوتایی (Y, l) که معرف سیگنال ورودی به همراه برجسب آن بود، گرادیان L با توجه به دیکشنری لایه‌ی آخر یعنی $D^{(Q)}$ طبق فرمول (8) و (11) محاسبه می‌گردد و به صورت زیر داریم:

$$\begin{aligned} & \nabla_{D^{(Q)}} L(l_k, \hat{x}_k^{(Q)}, W) \\ &= D^{(Q)} \beta \hat{x}_k^{(Q)T} + (y_k^{(Q)} - D^{(Q)} \hat{x}_k^{(Q)}) \beta^T \end{aligned} \quad (15)$$

که β برای اندیس مای مشمول در Λ به صورت زیر تعریف می‌شود:

$$\beta_\Lambda = \left(D^{(Q)\Lambda T} D^{(Q)\Lambda} \right)^{-1} \nabla_{\hat{x}_{k\Lambda}^{(Q)}} L(l_k, \hat{x}_k^{(Q)}, W) \quad (16)$$

و همچنین اگر $j \notin \Lambda$ ، $\beta_j = 0$.

برای محاسبه‌ی گرادیان L نسبت به دیکشنری D در لایه‌ی q یعنی $D^{(q)}$ داریم:

$$\begin{aligned} & \nabla_{D^{(q)}} L(l_k, \hat{x}_k^{(q)}, W) = \\ & D^{(q)} \beta \hat{x}_k^{(q)T} + (y_k^{(q)} - D^{(q)} \hat{x}_k^{(q)}) \beta^T \end{aligned} \quad (17)$$

$$\beta_\Lambda = \left(D^{(q)\Lambda T} D^{(q)\Lambda} \right)^{-1} \nabla_{\hat{x}_{k\Lambda}^{(q)}} L(l_k, \hat{x}_k^{(q)}, W) \quad (18)$$

تأکید می‌کنیم که فقط لایه‌ی آخر یعنی Q در مرحله‌ی طبقه‌بندی دخالت دارد چراکه در این مرحله داریم $L(l_k, \hat{x}_k^{(Q)}, W)$ و نه $L(l_k, \hat{x}_k^{(q)}, W)$ و با توجه به این نکته، خروجی لایه‌ی آخر بردار سیگنال کد می‌باشد. برای محاسبه‌ی گرادیان تابع هزینه‌ی طبقه‌بندی یعنی $L(l_k, \hat{x}_k^{(q)}, W)$ نسبت به دیکشنری D در لایه‌ی q یعنی $D^{(q)}$ ، محاسبه‌ی $\nabla_{\hat{x}_{k\Lambda}^{(q)}} L(l_k, \hat{x}_k^{(q)}, W)$ لازم است. محاسبه‌ی این گرادیان می‌تواند به صورت زیر تجزیه شود:

$$\frac{\partial L}{\partial \hat{x}^{(q)}} = \frac{\partial L}{\partial \hat{x}^{(q+1)}} \frac{\partial \hat{x}^{(q+1)}}{\partial y^{(q+1)}} \frac{\partial y^{(q+1)}}{\partial \hat{x}^{(q)}} \quad (19)$$

درجایی که:

$$\frac{\partial \hat{x}^{(q)}}{\partial y^{(q)}} = \left(D_\Lambda^{(q)T} D_\Lambda^{(q)} \right)^{-1} D_\Lambda^{(q)}$$

و در غیر این صورت،

برای تصویر تست y_i ابتدا با حل مسئله‌ی بهینه‌سازی زیر، ضرایب تنک x_i را محاسبه می‌نمائیم:

$$x_i = \operatorname{argmin}_{x_i} \|y_i - \bar{D}x_i\|_2^2 \quad \text{s.t.} \quad \forall i, \|x_i\|_0 < T \quad (28)$$

و سپس طبقه‌بندی‌کننده‌ی خطی \hat{W} کد مای تنک را وزن دهی نموده و آن‌ها را به کلاس‌ها منصوب می‌کند و نهایتاً برچسب موردنظر یعنی کلاسی که y_i به آن تعلق دارد، به‌صورت زیر برای آن تخمین زده می‌شود:

$$\text{class} = \operatorname{argmax}_{x_i} \hat{W}x_i \quad (29)$$

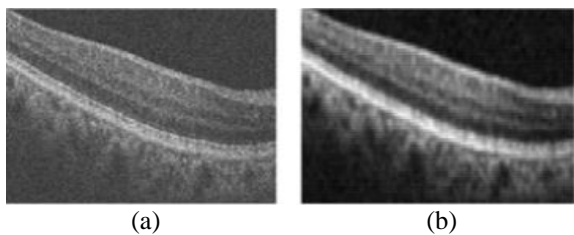
3. آزمایش‌ها و نتایج

در این بخش به‌منظور بررسی نتایج و اعتبارسنجی الگوریتم پیشنهادی از مجموعه داده‌های دام استفاده شده است. این مجموعه داده توسط اسرینیواسان و همکاران [30] منتشر شده که جهت تحلیل تصاویر OCT در دانشگاه مای دام، هاروارد و میشیگان مورد استفاده قرار گرفته است. مجموعه داده‌ی دام دربردارنده‌ی اسکن‌های حجمی شامل 45 نوع حجم می‌باشد که 15 مورد طبیعی، 15 مورد بیمار AMD و 15 مورد بیمار DME تعداد اسکن‌های OCT در هر حجم بین 36 تا 97 تصویر متغیر است.

در فاز پیش‌پردازش تصاویر به‌منظور سنجش عملکرد روش حذف نویز مورد بررسی، معیار مای نسبت کنتراست به نویز (CNR)، حفاظت از بافت (TP)، حفاظت از لبه (EP) و تعداد مساوی جستجو (ENL) محاسبه شدند. نتایج حاصله در جدول 1 و تحلیل بصری شکل 9 بیانگر کارایی بالای روش انجام شده می‌باشد:

جدول 1. مقادیر اندازه‌گیری شده مربوط به روش حذف نویز پیشنهادی
Table 1. values of proposed denoising method

	Original image	Noise removed image
CNR	3.09	24.24
	1.35	10.63
EP	1	0.88
	0	0.01
TP	1	0.18
	0	0.03
ENL	28.69	2045.08
	0.49	± 367.62
Time of calculation	-	24.78
		± 2.62



شکل 9. (a) تصویر نویز دار، (b) تصویر حذف نویز شده

Fig 9. a) Noisy Image b) Noise Removed Image

اجتناب می‌کند و قابل اجرا در مسائلی با تعداد کلاس بالا می‌باشد. به‌علاوه این روش به ما اجازه می‌دهد که به‌راحتی بخش متمایزکننده‌ی دیگری، یعنی خطای کدگذاری تنک متمایز را در تابع هدف ترکیب نمائیم که این امر بدون توجه به سایز دیکشنری، منجر به بازنمایی تنک متمایز می‌شود.

همچنین ما نیاز داریم که پارامتر مای D_0 و A_0 برای الگوریتم مان مقداردهی اولیه شوند. برای دیکشنری اولیه، در روش پیشنهادی، ابتدا در هر کلاس با استفاده از الگوریتم خوشه‌بندی k-means [43]، عملیات خوشه‌بندی را انجام داده و مراکز خوشه‌ها را به‌عنوان نماینده‌های آن کلاس در ستون مای دیکشنری اولیه‌مان قرار می‌دهیم. بنابراین ستون k ام دیکشنری یعنی d_k مرکز خوشه مای ایجاد شده در کلاس با برچسب k می‌باشد. برچسب هر ستون دیکشنری متناظر با کلاسی است که از عناصر آن تولید شده است و در طول فرایند یادگیری ثابت می‌ماند و این درحالی‌که است که عناصر d_k یعنی ستون مای دیکشنری در طول فرایند به‌روز می‌شوند. عناصر دیکشنری به‌صورت یکنواخت به هر کلاس اختصاص دارند و تعداد آن‌ها وابسته به اندازه‌ی دیکشنری می‌باشد. در مورد مقداردهی اولیه به A_0 ، مدل رگرسیون چند متغیره استفاده شده است [37]، که در آن:

$$A = \operatorname{argmin}_A \|z^{(q)} - Ax^{(q)}\|_2^2 + \gamma \|A\|_2^2 \quad (25)$$

و منجر به جواب زیر می‌شود:

$$A = \left(x^{(q)}(x^{(q)})^t + \gamma I \right)^{-1} x^{(q)}(z^{(q)})^t \quad (26)$$

2.4.5 طبقه‌بندی

ما $D = (d_1, d_2, \dots, d_m)$ و $W = (w_1, w_2, \dots, w_m)$ و $A = (a_1, a_2, \dots, a_m)$ را از D_{new} با به کار بردن الگوریتم K-SVD به دست آورده‌ایم، ولی به‌سادگی برای طبقه‌بندی داده‌های تست به کار نمی‌آیند لذا D ، W و A با استفاده از نرم 2 باهم در D_{new} نرمال می‌شوند، یعنی $\forall k, \|d_k^t, \sqrt{\alpha} a_k^t\|_2 = 1$. دیکشنری مطلوب \hat{D} ، پارامترهای تبدیل \hat{A} و پارامترهای طبقه‌بندی‌کننده‌ی \hat{W} به‌صورت زیر محاسبه می‌شوند:

$$\begin{aligned} \hat{D} &= \left\{ \frac{d_1}{\|d_1\|_2}, \frac{d_2}{\|d_2\|_2}, \dots, \frac{d_m}{\|d_m\|_2} \right\} \\ \hat{W} &= \left\{ \frac{w_1}{\|w_1\|_2}, \frac{w_2}{\|w_2\|_2}, \dots, \frac{w_m}{\|w_m\|_2} \right\} \\ \hat{A} &= \left\{ \frac{a_1}{\|a_1\|_2}, \frac{a_2}{\|a_2\|_2}, \dots, \frac{a_m}{\|a_m\|_2} \right\} \end{aligned} \quad (27)$$

با توجه به جداول فوق بررسی نتایج نشان می‌دهد، در شرایط مناسب افزایش تعداد تکرار آزمایش و نیز افزایش تعداد داده‌های آموزشی انتخابی باعث بالا رفتن دقت طبقه‌بندی می‌شود. همچنین به منظور اعتبارسنجی الگوریتم طبقه‌بندی، در جدول 4 روش پیشنهادی در این مقاله با سه روش CNN و SVM و SRC مقایسه شده است.

جدول 3. مقایسه نرخ طبقه‌بندی با داده‌های آموزشی متفاوت با تکرار مای گوناگون بر روی مجموعه داده‌ی دام

Table 3. Comparison of classification rates with different training data with different iteration on the Duke dataset

Class/Iteration	1/2 dataset training			2/3 dataset training		
	5	10	15	5	10	15
Normal	96.45%	97.30%	100%	97.64%	99.50%	100%
ADM	94.20%	95.46%	97.02%	96.89%	97.61%	98.32%
DME	99.77%	100%	99.87%	99.90%	100%	99.90%

روش اول تشخیص خودکار بیماری بر اساس طبقه‌بندی کننده ماشین بردار پشتیبان (SVM) چندلایه [38]، روش دوم مبتنی بر تصاویر خام و لایه مای تقسیم شده توسط شبکه مای عصبی کانولوشنی [39] و روش سوم طبقه‌بندی مبتنی بر نمایش تنک یا SRC [36] می‌باشند.

جدول 4. مقایسه عملکرد الگوریتم طبقه‌بندی پیشنهادی با سه روش مبتنی بر SRC, SVM, CNN

Table 4. Comparison of proposed algorithm performance with three methods based on SRC, SVM, CNN.

Method	Dataset	Accuracy
Proposed	45	98.61%
CNN	70	98.5%
SVM	90	97.77%
SRC	64	95.1%

توجه داریم که با توجه به متفاوت بودن مجموعه داده‌های مورد آزمایش و همچنین تفاوت در تعداد داده‌های آموزشی انتخابی و تکرار فرایند طبقه‌بندی در هر یک از روش‌ها مقادیر مندرج در جدول قابل قیاس نیست.

4. نتیجه‌گیری

هدف از این مطالعه بهبود عملکرد طبقه‌بندی روش‌های نوین، با استفاده از مدلی چندلایه بود. مدل چندلایه‌ای که از الگوریتم پیشرفته K-SVD چندلایه، برای آموزش ماتریس دیکشنری و الگوهای پایه استفاده می‌کرد تا بتواند با الگوپذیری از مدل‌هایی با معماری چندلایه، ویژگی‌های بهتری را در تصاویر OCT شبکیه بیاموزد. همچنین علاوه بر استفاده از برجسب‌های کلاس داده‌های

برای بررسی پیچیدگی روش پیشنهادی در این مقاله متوسط و انحراف از معیار زمان موردنیاز برای تحلیل هر اسلایس OCT بر روی یک کامپیوتر شخصی با مشخصات ویندوز 32 بیتی و پنتیوم اینتل دو هسته‌ای با رم 4 گیگابایت و پردازنده‌ی 3 گیگاهرتز گزارش شده است.

در فاز اصلی مقاله که مربوط به طبقه‌بندی تصاویر OCT و تفکیک و تشخیص تصاویر نرمال و غیرنرمال می‌باشد، اعتبارسنجی کراس چندین بار تکرار شد؛ در هر تکرار، طبقه‌بندی کننده‌ی خطی چند کلاس روی 42 حجم (به‌غیر از یک حجم از هر کلاس) آموزش داده و روی سه حجم باقیمانده تست شد. از آنجایی که هر حجم شامل تعداد زیادی از تصاویر OCT مربوط به یک شخص خاص است، ما یک حجم را به یک کلاس (نرمال، AMD و DME) اختصاص دادیم تا بیشترین کارایی را در طبقه‌بندی داشته باشیم. نتایج اعتبارسنجی کراس در جدول 2 نشان می‌دهد، صد درصد از 30 حجم تصاویر OCT کلاس مای AMD و DME دقیقاً درست طبقه‌بندی شدند و 95.85% از تصاویر نرمال طبقه‌بندی صحیح داشتند که نتایج بسیار خوبی در مقایسه با روش‌های مشابه موجود می‌باشد.

جدول 2 نتایج اعتبارسنجی کراس روی مجموعه داده‌ی دام

Table 2. Cross-validation results on Doc dataset

Class	Proposed method
Normal	95.85%
ADM	100%
DME	100%

در ادامه ما آزمایش‌های بیشتری با مجموعه داده‌های آموزشی و تعداد تکرار مای مختلف انجام دادیم؛ برای به دست آوردن نتایج مناسب فرایند آزمایش را در سه نوبت، 5، 10 و 15 بار تکرار کردیم تصاویر آموزشی متفاوتی را به صورت تصادفی از مجموعه داده‌ها انتخاب کرده و مابقی را داده‌ی تست در نظر گرفتیم. در هر بار آزمایش نرخ طبقه‌بندی تصاویر نرمال و بیمار که شامل AMD و DME بودند محاسبه و نتایج نهایی بر اساس میانگین و انحراف از معیار نرخ مای طبقه‌بندی تعیین شدند. در آزمایش‌ها، ما ابتدا نیمی از داده‌ها را به‌عنوان داده‌ی آموزشی در نظر گرفتیم و در مرحله‌ی بعد دوسوم داده‌ها به‌عنوان داده‌ی آموزشی و بقیه داده‌ی تست در نظر گرفته شدند. نتایج آزمایش‌ها در جدول 3 آورده شده است.

مراجع

- [1] Schirmacher F, Köhler T, Husvogt L, (2017). Quantile sparse image prior for spatio-temporal denoising of retinal OCT data. *Medical Image Computing and Computer-Assisted Intervention*, 12(4), 83-91.
- [2] Sarbjeet Kaur, V., & Banga, K. (2014). Review paper of character recognition using image processing. *International Journal of Computer Science & Engineering Technology*, 5(1), 254-268.
- [3] Fan Meng et al, (2017). A sparse dictionary learning-based adaptive patch in painting method for thick Clouds removal from high-spatial resolution remote sensing imagery. *Sensors*, 17(9), 123-135.
- [4] Zheng H., Zhu, J., Yang, Z., & Jin, Z. (2017). Effective micro-expression recognition using relaxed K-SVD algorithm. *International Journal of Machine Learning*, 8(6), 2043-2049.
- [5] Chan Wai Tim, S., Rombaut, M., & Pellerin, D, (2015). Rejection-based classification for action recognition using a spatio-temporal dictionary. *Advanced Concepts for Intelligent Vision Systems*, 15(8), 522-533.
- [6] Zhang, Z. et al. (2016). Discriminative-dictionary-learning-based multilevel point-cluster features for ALS point cloud classification. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12), 334-347.
- [7] Zhang, Z. et al. (2016). Sparse codes auto-extractor for classification: A joint embedding and dictionary learning framework for representation. *IEEE Signal Processing Society*, 64(14), 553-561.
- [8] de Moura J., Novo, J., & Rouco, J. (2017). Automatic Detection of Blood Vessels in Retinal OCT Images. *Biomedical Applications Based on Natural and Artificial Computing*, 10(4), 3-10.
- [9] Oguz, I. et al. (2016). Optimal retinal cyst segmentation from OCT images. *Medical Imaging*, 15(9), 245-256.
- [10] Sun, Y., Lim S., Sun, Z. (2017). Fully automated macular pathology detection in retina optical coherence tomography images using sparse coding and dictionary learning. *Journal of biomedical optics*, 22(1), 214-225
- [11] Rickman, C. B. et al. (2013). Dry age-related macular degeneration: mechanisms, therapeutic targets, and imaging dry AMD mechanisms, targets, and imaging. *Invest. Ophthalmol. Visual Sci*, 54(14), 68- 80.
- [12] Antonym B.J et al. (2013). A combined machine-learning and graph-based framework for the segmentation of retinal surfaces in SD-OCT volumes. *Biomed. Opt. Express*, 4(12), 2712-2728.
- آموزشی، اطلاعات برچسب نیز در یادگیری اتم مای ماتریس دیکشنری به کار گرفته شد تا تمایز بیشتری را در فرایند کدگذاری تنک اعمال کند و این منجر به موفقیت مراحل کدگذاری تنک در هر تکه از تصویر و نهایتاً کل تصویر، باهدف پیدا کردن نمایش مؤثرتری از داده به منظور طبقه‌بندی شد. ایده اصلی روش پیشنهادی این بود که لایه کدگذاری تنک را تکرار کرده و از K-SVD سازگار با برچسب استفاده نماید تا تبعیض پذیری ویژگی‌ها را افزایش دهد. هدف دیگر روش ارائه‌شده این بود که اندازه ابعاد داده ورودی را با کاهش دادن کدگذاری تنک تصویر اصلی کنترل کند. به این ترتیب که به جای کدگذاری تصویر اصلی از کدگذاری تنک تکه مای کوچک از تصویر اصلی که می‌تواند عملیات پردازش و تحلیل را بسیار مؤثرتر نماید، استفاده کند. سومین ایده‌ی این روش پیدا کردن دیکشنری بهینه جهت ارتقا تأثیرگذاری فرایند طبقه‌بندی بود که این مهم با استفاده از الگوریتم مای خوشه‌بندی و پیدا کرده مراکز کلاس‌ها محقق شد. ویژگی مهم دیگر این روش در مقایسه با روش‌های پیچیده‌تر، عدم اتکالی آن به فرایند تقطیع لایه مای شبکه‌ی بود چراکه بیماری مای شبکه‌ی، لایه‌ها را مدام تغییر می‌دهند و کار تقطیع لایه‌ها را پیچیده می‌کنند.
- نتیجه‌ی این تحقیق ارائه‌ی سیستمی اتوماتیک به منظور تشخیص برخی ناهنجاری مای شبکه‌ی بود به طوری که با تحلیل و آنالیز بر روی داده‌های دام توانست با دقت بسیار خوبی در مقایسه با روش‌های نوین این حوزه در شناسایی الگوهای ظریف بیماری در تصاویر OCT موفق عمل نموده، تصاویر نرمال و بیمار را در غالب دو بیماری تباهی ماکولای وابسته به سن یعنی AMD و ماکولای دیابتی یعنی DME را تفکیک و در تشخیص پاتولوژی شبکه‌ی با دقت بسیار بالایی به پزشک معالج کمک کند. به عنوان یک پیشنهاد تحقیقاتی و کاربردی برای متخصصین و آیندگان با تعمیم این روش به کلاس مای بیشتر می‌توان تمام بیماری مای شبکه‌ی را پوشش داد و از آن به عنوان ابزاری بالقوه مؤثر در تشخیص کامپیوتری و غربالگری بیماری‌های شبکه‌ی و یا در حوزه مای وسیع‌تر چشم استفاده نمود.

پی‌نوشت

- ¹ Dictionary Learning and Sparse Representation
- ² Optical Coherence Tomography
- ³ Diabetic Macular Edema
- ⁴ Age-related Macular Degeneration
- ⁵ Singular value decomposition
- ⁶ Principal Component Analysis
- ⁷ Support Vector Machine

- [26] Hijazi, M.H.A, Coenen, F., & Zheng, Y. (2012). Data mining techniques for the screening of age-related macular degeneration. *Knowled e-Based Syst.* 29(1), 83–92.
- [27] Kafieh, R., & Rabbani, H. (2013). Optical coherence tomography noise reduction over learned dictionaries with introduction of complex wavelet for noise reduction. *Proceedings, Wavelets and Sparsity*, 8858(26), 238-247.
- [28] Albarrak, A. et al. (2012). Volumetric image mining based on decomposition and graph analysis: an application to retinal optical coherence tomography. *Computational Intelligence and Informatics*, 1344(64), 245-256.
- [29] Wang, Q. et al, (2017). Synthesis K-SVD based analysis dictionary learning for pattern classification. *Multimedia Tools and Applications*, 12(4), pp1-24.
- [30] Srinivasan, P.P et al, (2014). Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomed. Opt. Express*, 5(10), 3568–3577.
- [31] Esmaili, M., Dehnavi, A. M., & Rabbani, H. (2017). Speckle noise reduction in optical coherence tomography using two-dimensional curvelet-based dictionary learning. *Journal of medical*, 17(2), 86-91.
- [32] Yankelevsky, Y. et al, (2016). Structure-aware classification using supervised dictionary learning. *Computer Vision and Pattern Recognition*, 10(5), 679-692.
- [33] Sun, Y., Li, S.H, Sun, Z.H. (2017). Fully automated macular pathology detection in retina optical coherence tomography images using sparse coding and dictionary learning. *J. Biomed. Opt.*, 22(1), 160-171.
- [34] Hassan, T. et al, (2015). Review of OCT and fundus images for detection of Macular Edema. *Imaging Systems and Techniques*, 155(45), 234-245.
- [35] Shalev-Shwartz, S., & Shamir, O. (2017). Failures of gradient-based deep learning. *International Journal of Machine Learning*, 17(3), 456-466.
- [36] Javidi, M., Pourreza, H.R., & Harati, A. (2017). Vessel segmentation and microaneurysm detection using discriminative dictionary learning and sparse representation. *Computer methods and programs in biomedicine*, 139(1), 93-108.
- [37] Geimer, T., et al, (2017). A Kernel Ridge regression model for respiratory motion estimation in radiotherapy. *Bildverarbeitung für die Medizin*, 13(4), 155-160.
- [38] Khalid, S., et al, (2017). Fully automated robust system to detect retinal edema, central serous chorioretinopathy, and age related macular degeneration
- [13] Carass, A. et al. (2014). Multiple-object geometric deformable model for segmentation of macular OCT. *Biomed. Opt. Express*, 5(4), 1062–1074.
- [14] Chium S. J. et al. (2012). Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images. *Invest. Ophthalmol. Visual Sci*, 53(1), 53–61.
- [15] Sun, Y. et al. (2016). 3D automatic segmentation method for retinal optical coherence tomography volume data using boundary surface enhancement. *J. Innovative Opt. Health Sci*, 9(2), 165-177.
- [16] Vermeer, K. et al. (2011), Automated segmentation by pixel classification of retinal layers in ophthalmic OCT images. *Biomed. Opt. Express*, 2(6), 1743–1756.
- [17] Yang, Q. et al. (2011), Automated segmentation of outer retinal layers in macular OCT images of patients with retinitis pigmentosa, *Biomed. Opt. Express* 2(9), 2493–2503.
- [18] Lang, A. et al. (2013), Retinal layer segmentation of macular OCT images using boundary classification. *Biomed. Opt. Express*, 4(7), 1133–1152.
- [19] Garcia-Allende, P. B et al. (2011). Morphological analysis of optical coherence tomography images for automated classification of gastrointestinal tissues. *Biomed. Opt. Express*, 2(10), 2821–2836.
- [20] Chan, S. et al, (2016), Multi-layer Dictionary Learning for Image Classification. *Advanced Concepts for Intelligent Vision Systems*, 4(15), 132-143.
- [21] Jiang, Z., Lin, Z., & Davis, L., (2011). Learning a discriminative dictionary for sparse coding via label consistent K-SVD. *IEEE computer society conference on computer vision and pattern recognition*, 6(12), 1697-1704.
- [22] Pande, P. et al. (2014), Automated classification of optical coherence tomography images for the diagnosis of oral malignancy in the hamster cheek pouch. *J. Biomed. Opt.* 19(8), 356-368.
- [23] Alsaih, K., et al, (2017). Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images. *Bio Medical Engineering*, 16(1), 447-455.
- [24] Liu, Y. et al. (2011). Automated macular pathology diagnosis in retinal OCT images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding. *Med. Image Anal*, 15(5), 748–759.
- [25] Zheng, Y., Hijazi, M.H.A, & Coenen, F. (2012). Automated ‘disease/no disease’ grading of age-related macular degeneration by an image mining approach. *Invest. Ophthalmol. Visual Sci*, 53(13), 8310–8318.

[44] Kafieh, R. (2014). *Combination of graph based and space-frequency methods in analysis of Optical coherence Tomography (OCT) images*, (Unpublished doctoral thesis), Isfahan University, Isfahan. [in Persian].

Supervised Dictionary Learning
and Sparse Representation

Sparse Representation-
based Classification

Hyper-reflective complex

contrast to noise ratio

Texture Preservation

Edge Preservation

Equivalent Number of Looks

$$\|X\|_F^2 = \sum_{i=1}^n \sum_{j=1}^k x_{ij}^2$$

پیوست‌ها

S-DLSR

SRC

HRC

CNR

TP

EP

ENL

نرم فروبنیوس

from optical coherence tomography images. *J. Biomed. Opt.*, 71(48), 278-293.

[39] Lu, D. et al. (2017). Retinal fluid segmentation and detection in optical coherence tomography images using fully convolutional neural network. *Hindawi BioMed Research International*, 17(6), 477- 489.

[40] Gangeh, M. et al, (2015). Supervised dictionary learning and sparse representation-A review. *Computer Vision and Pattern Recognition*, 20(3), 145-158.

[41] Guan, J., et al. (2018). Polynomial dictionary learning algorithms in sparse representations. *Signal Processing*, 142(7), 492-503

[42] Su, H., et al, (2018), Multifeature dictionary learning for collaborative representation classification of Hyperspectral imagery. *IEEE Transactions on Geoscien*, 56(4), 345-357.

[43] Aggarwal, C. C, & Reddy, C.K. (2013). *Data clustering, algorithms and applications*. New York: Wiley.

How to cite this paper:

Azadeh Montazeri, Mahboubeh Shamsi, Rouhollah Dianat, (2019). Automatic recognition of retinal diseases using mathematical models of image processing, based on multilayer-dictionary learning. *Journal of Technology of Education*, 13(3), 371-386.

DOI: 10.22061/jte.2018.3434.1872

URL: http://jte.sru.ac.ir/?_action=showPDF&article=850

