



ORIGINAL RESEARCH PAPER

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

A. Montazeri¹, M. Shamsi^{2,*}, R. Dianat³¹Department of Information Technology, Faculty of Computer, University of Qom, Iran²Department of Computer, Faculty of Computer and Electrical, Qom University of Technology, Iran³Department of Computer, Faculty of Computer, University of Qom, Iran

ABSTRACT

Submitted: 17 June 2018
 Reviewed: 16 July 2018
 Revised: 25 August 2018
 Accepted: 26 September 2018

KEYWORDS:

Multi-layer Dictionary Learning
 Sparse Representation
 K-SVD Algorithm
 Optical Coherence Tomography
 Classification

* Corresponding author
 ✉ shamsi@qut.ac.ir

Background and Objectives: Image processing is one of the most important issues in the field of artificial intelligence, which is used in various industrial, medical, military, and security systems. One of the most important applications of image processing is the extraction of different types of classification in the field of medical sciences. By using powerful algorithms in this field, intelligent systems can be invented that automatically understand and interpret the medical characteristics of individuals without the need to the physician supervision can discover useful information to help experts make good judgments. When the necessary parameters for the diagnosis of the disease increase, the diagnosis and prognosis of the disease becomes very difficult even for an expert, which is why computer diagnostic tools have been used in recent decades to help the physicians. This has led to a reduction in possible errors due to fatigue or inexperience of the specialist, and to provide the required medical data to the physician in less time and with more detail and accuracy. The purpose of this study is to improve the classification of new methods using a multi-layered model to address retinal diseases diagnosis.

Methods: 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.

Findings: 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.

Conclusion: The results of this study were to provide an automatic system for the diagnosis of some retinal abnormalities in a way that it could do data analysis with high accuracy in comparison to other modern methods to diagnosis delicate patterns of OCT, separate images of normal and patient the normal and in two age-related macular degeneration diseases (AMD), and diabetic macular degeneration (DME), and help the physician to diagnose retinal pathology with great care. As a suggestion for professionals and future research, by generalizing this method to the more classes, we can cover the entire retinal myopia and use it as a potentially effective tool in computerized diagnosis and screening for retinal disease or in the wider eye area.



NUMBER OF REFERENCES

44



NUMBER OF FIGURES

9



NUMBER OF TABLES

4

مقاله پژوهشی

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

آزاده منتظری^۱، محبوبه شمسی^{۲*}، روح‌الله دیانت^۳^۱ گروه مهندسی فناوری اطلاعات، دانشکده کامپیوتر، دانشگاه قم، قم، ایران^۲ گروه کامپیوتر، دانشکده برق و کامپیوتر، دانشگاه صنعتی قم، قم، ایران^۳ گروه کامپیوتر، دانشکده کامپیوتر، دانشگاه قم، قم، ایران

چکیده

پیشینه و اهداف: پردازش تصویر یکی از مسائل مهم در حوزه ی هوش مصنوعی می باشد که در موارد مختلف صنعتی، پزشکی، نظامی، سیستم های امنیتی و... کاربرد دارد. از مهم ترین زمینه های کاربردی پردازش تصویر استخراج قواعد طبقه بندی در حیطه علم پزشکی است. با به کارگیری الگوریتم های قدرتمند این حوزه می توان سیستم های هوشمندی ابداع کرد که به شکل خودکار و بدون نیاز به نظارت پزشک قادر به فهم و تفسیر ویژگی های 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 می شود، تشخیص و پیش بینی بیماری حتی برای یک متخصص خبره پزشکی نیز به سختی ممکن می گردد 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 بی تجربگی فرد متخصص تا حدی کاهش پیدا کند و داده های پزشکی مورد نیاز، طی مدت زمان کمتر و با جزئیات و دقت بیشتر در اختیار پزشک قرار گیرد. هدف از این مطالعه بهبود عملکرد طبقه بندی روش های نوین، با استفاده از مدلی چندلایه به منظور کمک به تشخیص بیماری های شبکه ی چشم است.

دریافت: ۲۷ خرداد ۱۳۹۷

داوری: ۲۵ تیر ۱۳۹۷

اصلاح: ۰۳ شهریور ۱۳۹۷

پذیرش: ۰۴ مهر ۱۳۹۷

واژگان کلیدی:

یادگیری دیکشنری چندلایه

نمایش تنک

الگوریتم K-SVD

مقطع نگاری همدوسی شبکه

طبقه بندی

روش ها: این مدل از الگوریتم K-SVD پیشرفته، برای یادگیری ماتریس دیکشنری و الگوهای پایه استفاده می کند تا بتواند با الگوپذیری از معماری چندلایه، ویژگی های بهتری را در تصاویر OCT شبکه ی بیاموزد. همچنین در این معماری، علاوه بر استفاده از برجسب های کلاس داده های آموزشی، اطلاعات برجسب نیز در هر ستون پایه در ماتریس دیکشنری ترکیب می شود تا در کدگذاری تنک در طی فرآیند یادگیری دیکشنری بیشترین تبعیض اعمال شود که این منجر به موفقیت مراحل کدگذاری تنک و جمع بندی، در پیدا کردن نمایش مؤثرتری از داده به منظور طبقه بندی می گردد. برای اعتبارسنجی الگوریتم، از مجموعه داده های داک استفاده شده است. این مجموعه داده شامل اسکن های حجمی از ۴۵ نوع، که ۱۵ نوع حجم از آن طبیعی، ۱۵ نوع بیمار تخریب ماکولای وابسته به سن (AMD) و ۱۵ نوع بیمار ورم ماکولای دیابتی (DME) است که تعداد اسکن های OCT در هر حجم بین ۳۶ تا ۹۷ تصویر متغیر می باشد.

* نویسنده مسئول

shamsi@qut.ac.ir

یافته ها: نتایج تجربی نشان می دهد که الگوریتم پیشنهادی این مقاله توانسته است با پیشی گرفتن از بسیاری از مدل های جدید یادگیری دیکشنری و نمایش تنک، بسیار خوب عمل نماید و با دقت خوبی منجر به طبقه بندی صحیح ۹۵٫۸۵٪ برای تصاویر نرمال و صد درصد برای تصاویر بیمار (AMD و DME) شود.

نتیجه گیری: نتیجه ی این تحقیق ارائه ی سیستمی اتوماتیک به منظور تشخیص برخی ناهنجاری های شبکه ی بود به طوری که با تحلیل و آنالیز بر روی داده های دام توانست با دقت بسیار خوبی در مقایسه با روش های نوین این حوزه در شناسایی الگوهای ظریف بیماری در تصاویر OCT موفق عمل نموده، تصاویر نرمال و بیمار را در دو بیماری تباهی ماکولای وابسته به سن یعنی AMD ورم ماکولای دیابتی یعنی DME را تفکیک و در تشخیص پاتولوژی شبکه 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 های بیشتر می توان تمام بیماری های شبکه را پوشش داد و از آن به عنوان ابزاری بالقوه مؤثر در تشخیص کامپیوتری و غربالگری بیماری های شبکه و یا در حوزه های وسیع تر چشم استفاده نمود.

مقدمه

برای تعبیر و تحلیل انسان باشد. برای نمونه در پزشکی به کمک مدل ها و روش های موجود وضوح و کیفیت تصویر را ارتقا می دهند یا این که برای تعبیر آسان تر تصاویر اشعه ایکس یا سایر تصاویر پزشکی، سطوح شدت روشنایی را بارنگ، رمز می کنند. متخصصان جغرافیایی نیز از این روش ها یا روش های مشابه برای مطالعه الگوهای آلودگی هوا که با

پردازش تصویر یکی از مسائل مهم در حوزه ی هوش مصنوعی می باشد که در موارد مختلف صنعتی، پزشکی، نظامی، سیستم های امنیتی و... کاربرد دارد. گر چه اغلب این مسائل باهم نامرتبط هستند، اما عموماً نیازمند روش هایی هستند که قادر به ارتقای اطلاعات تصویری

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 به‌سختی 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 قابل‌توجهی به کیفیت تصمیمات بالینی نمایند [۱-۳].

در این میان تلاش‌ها و پیشرفت‌های مهمی در عرصه پردازش تصاویر شبکیه نیز به‌منظور ارائه دستگاه‌هایی اتوماتیک جهت تشخیص بیماری‌های مختلف روی آن انجام شده است. چنین سیستم‌هایی علاوه بر اینکه امکان پردازش تصاویر شبکیه 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 هم‌دوسی نوری) از روی شبکیه و فراهم شدن امکان اخذ اطلاعات و تصاویر با دقت و کیفیت بالا از لایه‌های مختلف شبکیه باعث شده است تا باکار بر روی این تصاویر، خطای الگوریتم‌های ارائه‌شده جهت تشخیص و استخراج اتوماتیک الگوهای مرتبط با بیماری‌های شبکیه را کمتر نمود، و روش‌ها و الگوریتم‌های نوینی در جهت آنالیز اتوماتیک این تصاویر معرفی کرد. تصویربرداری OCT چشم یک روش تصویربرداری از لایه‌های 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 [۱۱].

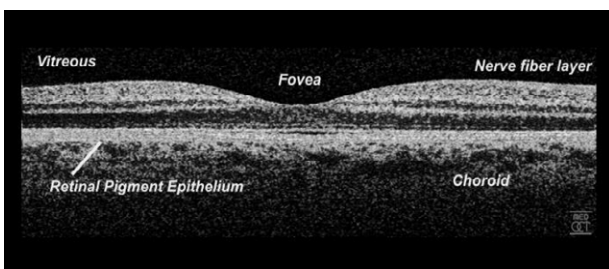
در طول دو دهه گذشته اکثر آثار محققین در زمینه پردازش تصاویر OCT، به دو بخش اصلی اختصاص یافته که بخش اول روش‌های بخش‌بندی و تقطیع لایه‌های شبکیه [۱۸-۱۲] است، که در این مقاله مورد بحث قرار نگرفته و بخش دوم از مقالات این حوزه، که هدف ما است به روش‌های گوناگون طبقه‌بندی تصاویر OCT پرداخته است که عموماً مبتنی بر بازسازی تصویر هستند [۳۱-۱۹]. در سال ۲۰۱۱ لیو همکارانش به‌منظور تشخیص آسیب‌های ناحیه‌ی ماکولا، از جمله AMD و DME، روشی را پیشنهاد دادند که از الگوهای دودویی محلی

تصویربرداری هوایی و ماهواره‌ای به‌دست‌آمده 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 پرنرژی 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 [۵-۱].

به‌این 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 DLSR^۱ 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 OCT^۲ شبکیه دارد [۹-۳].

شبکیه داخلی‌ترین لایه‌ی چشم است و شامل سلول‌های گیرنده‌ی نور و نورون‌ها می‌باشد. بافت شبکیه 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 سیگنال‌های عصبی برای تحلیل نهایی توسط مغز می‌باشد. لکه‌ی زرد یا ماکولا بخشی از شبکیه است که بیشترین حساسیت به نور را دارد و موجب دیدمستقیم 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 شناخته‌شده متابولیکی که شیوع بالایی هم دارد، دیابت است. دیابت یا مرض قند می‌تواند باعث ایجاد ورم در ناحیه ماکولای شبکیه چشم شود. هرگونه آسیب یا بیماری شبکیه، نوعی رتینوپاتی تلقی می‌شود اما زمانی که به‌واسطه بیماری قند خون رگ‌هایی که سلول‌های شبکیه چشم را تغذیه می‌کنند دچار آسیب می‌شوند به آن رتینوپاتی دیابتی گفته می‌شود. ورم ماکولای دیابتی یا DME^۳ شکلی از رتینوپاتی دیابتی است که در آن از عروق خونی مایعاتی به داخل ماکولا نشت می‌کند. تورم این بخش از شبکیه موجب تاری دید می‌شود. همچنین استحاله ماکولای وابسته به سن یا AMD^۴ که در افراد سالمند ظاهر می‌یابد از جمله بیماری‌های شایع مربوط به شبکیه چشم است. این بیماری با کاهش دید، بیمار را دچار اختلال بینایی می‌کند [۱۰].

از مهم‌ترین زمینه‌های کاربردی پردازش تصویر استخراج قواعد طبقه‌بندی در حیطه علم پزشکی است. با به‌کارگیری 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



شکل ۱: اسکن OCT در بافت شبکیه چشم در ۸۰۰ نانومتر و رزولوشن محوری ۳ میکرون
Fig. 1: OCT scan in retinal tissue at 800 nm and axial resolution of 3 microns

یادگیری دیکشنری

انتقال سیگنال‌ها (فضای بردارهای ویژگی) به یک فضای جدید با خواص مطلوب، یک رویکرد رایج برای بهبود کار آبی، 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 PCA^۲ و نمایش تنک، بردارهای پایه‌ی وابسته به زمینه دارند. 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.

به‌عنوان 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 مطرح‌شده در این حوزه سعی دارند که این الگوریتم ساده را بهبود دهند و به‌جای 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 [۲۹] یک الگوریتم تکرارشونده‌ی موفق در حوزه‌ی یادگیری دیکشنری است که با به حداقل رساندن خطای بازسازی در فرمول (۱)، 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.

$$\langle D, X \rangle = \operatorname{argmin}_{D, X} \|Y - DX\|_F^2 \quad \forall i, x_{i0} < T \quad (1)$$

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

استفاده نموده و با استفاده از یک هرم فضایی چندبعدی و تحلیل مؤلفه‌های اصلی (PCA) برای کاهش ابعاد، تصاویر را بازسازی کردند [۲۴].

در سال ۲۰۱۲، مدلی برای بازسازی تصاویر بر اساس گراف پیشنهاد شد [۲۵ و ۲۶] 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، زیرگرافی‌های مشترک برای بازسازی تصاویر انتخاب شدند. در نهایت با استفاده از بردارهای ویژگی، طبقه‌بندی کننده، آموزش داده شد و یک دسته‌بندی دوتایی بر روی تصاویر معمولی و بیمار ADM شکل گرفت. به دنبال آن، آلبراک و همکاران در [۲۸] مدلی بر روی تصاویر سه‌بعدی OCT ارائه نمودند که ابتدا تصویر سه‌بعدی اولیه را به تصاویر با ابعاد کمتر تجزیه نموده سپس آن‌ها را به‌وسیله‌ی ساختار درختی بازسازی می‌کرد و نهایتاً با ترکیب ویژگی‌های استخراج‌شده از زیرگراف‌ها، 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 [۴-۶]. 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 [۴، ۵، ۷]. 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 با ناظر انجام می‌گردد.

در این مقاله مدلی با ناظر و چندلایه مبتنی بر الگوریتم K-SVD^۵ 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 تجزیه شده، 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 تکرارشونده منجر به تمایز بیشتر در ویژگی‌های به‌دست‌آمده و مینیمم سازی خطای بازسازی و کدگذاری تنک در خروجی لایه‌ی آخر می‌شود.

ماتریس D را آغاز می‌گردد:

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

که E^i اختلاف بین سیگنال ورودی و ترکیب بازسازی شده‌ی سیگنال ورودی با استفاده از عناصر دیکشنری است. در معادله‌ی (۵)، x^j j امین 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 $1 \leq h \leq n$ ، $x_{j,h} = 0$ ، 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_i و ضریب تنک نظیرش به v_p تبدیل می‌شود. این روند برای عناصر بعدی دیکشنری نیز با اضافه کردن i و محاسبه مجدد معادله‌ی (۵) و اجرای دوباره‌ی SVD ادامه می‌یابد؛ بنابراین تمام دیکشنری به‌روزرسانی می‌شود. با محاسبه مجدد ضرایب تنک به دفعات، 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 کلاس مای 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 الگوریتم مای قدرتمند طبقه‌بندی مانند SVM^۲ 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.

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 می‌کنند.

آموزش دیکشنری با ناظر

همان طوری 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 برجسب هر کلاس و داده‌های خاص آن، در یادگیری اتم مای دیکشنری و ضرایب نمایش تنک استفاده کنند. همین امر منجر به تشکیل دسته جدیدی از فن‌ها شده که اطلاعات برجسب

$X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$ مجموعه ضرایب تنک می‌باشد و m تعداد عناصر دیکشنری و n هم تعداد سیگنال مای 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 ستون مای ماتریس دیکشنری است. فرض کنید در مرحله‌ی اول دیکشنری D ثابت و فرمول (۱) نسبت به ماتریس X حل شود. بنابراین نرم فروبنیوس را به صورت زیر می‌توان بازنویسی کرد:

$$Y - DX_F^2 = \sum_{i=1}^k y_i - Dx_{i2}^2 \quad (2)$$

در نتیجه معادله‌ی (۱) به k مسئله به شکل زیر تقسیم می‌شود:

$$\langle D, X \rangle = \operatorname{argmin}_{D, X} y_i - Dx_{i2}^2 \quad \forall i, x_{i0} < T \quad (3)$$

T به اندازه‌ی کافی کوچک انتخاب می‌شود تا تنک بودن x_i ‌ها حفظ شود و تعداد کمی از ضرایب آن غیر صفر باشند. معادله‌ی (۳) 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 الگوریتم مای تعقیبی و یا حریصانه قابل انجام می‌باشد. این الگوریتم‌ها حریصانه اند و اتم مای دیکشنری پی‌درپی انتخاب می‌گردند. این روش‌ها شامل محاسبه حاصل ضرب درونی سیگنال و اتم‌های دیکشنری اند و گاهی نیز راه‌حل مای 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 درست بیشتر می‌باشد اما مسئله غیر محدب می‌گردد که ممکن است در بهینه‌های محلی گرفتار شده و به راه‌حل مای غیرقابل اعتمادی در فضای جستجو منجر گردد. در این دسته از روش‌ها از ضرب کننده مای لاگرانژ به منظور تبدیل قیود به تابع هدف و از روشی تکراری نتیجه شده بر اساس ایده حداقل مربعات دوباره وزن دهی شده تکرار شونده که نرم l_p را به عنوان یک نرم l_2 وزن داده شده کنترل می‌نماید، استفاده می‌گردد [۴۰] و [۴۲] و [۴۴].

در نهایت با حل معادله‌ی (۳) ماتریس ضرایب X را به صورت زیر خواهد بود:

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

ستون مای 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

را در محاسبه دیکشنری، ضرایب تنک یا هر دو آن‌ها به کار می‌گیرند و یادگیری دیکشنری و نمایش تنک با ناظر نامیده می‌شود [۲]، [۳۲]، [۳۴].

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 به وضوح قابل مشاهده 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 پیش‌پردازشی بر روی تصاویر OCT می‌باشد [۲۷].

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 تنک" 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 [۲۷]. 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 و فقی، روش‌های ماشین بردار پشتیبان و تخمین مای بی‌زین.

روش‌های مبتنی بر نمایش تنک 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 تنک 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 خاطر نشان کرد که تعداد زیادی از کارهای تحقیقاتی پیش از مرحله تقطیع تصویر از هیچ پیش‌پردازش خاصی کمک نمی‌گیرند.

در بین روش‌های مبتنی بر نمایش تنک با توجه به اینکه مبنای

به‌منظور تفسیر بهتر به‌وسیله‌ی مدل‌سازی ریاضی، فرض کنید دوتایی (y_k, l_k) موجود باشد به‌طوری‌که $[y_1, y_2, \dots, y_n] \in \mathbb{R}^{2 \times n}$ سیگنال ورودی دوبعدی و $l = [l_1, l_2, \dots, l_N]$ برچسب نظیر آن است. اگر L تابع هزینه‌ی طبقه‌بندی و W پارامترهای آن باشد، مسئله‌ی یادگیری دیکشنری با ناظر را به‌صورت زیر می‌توان نوشت:

$$\hat{x}_k = \operatorname{argmin}_{x_k} y_k - D x_k^2 \quad \text{s.t.} \quad \forall i \quad x_{i0} < T \quad (6)$$

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

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 Y و T حد آستانه و فاکتور تنظیم تنکی می‌باشند. $Y - DX^2$ خطای بازسازی نامیده می‌شود و بامی نیمم 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 سپس دیکشنری D و پارامترهای W برای انطباق با مسئله‌ی طبقه‌بندی به روز می‌شوند. هدف فرمول (۷) به حداقل رساندن تابع هزینه L با توجه به پارامترهای D و W است که به‌روزرسانی شدند. به‌منظور می‌نیمم سازی تابع هزینه می‌توان از روشی مشابه با الگوریتم پس انتشار خطا [۳۵] 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. یک راه غلبه بر این مشکل محاسبه‌ی گرادیان تابع هزینه L بر اساس دیکشنری D مفروض است؛ این موضوع در [۲] و [۴] 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:

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

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

می‌رود [۲۷].

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) $\text{Var}_{\text{old}} = \text{var}_{\text{new}}, K = k + 1$

(d) برای $i = 1:10$

۱. تعداد k داده را به صورت تصادفی انتخاب کن

۲. دیکشنری ارتقا یافته را برای k داده بساز

۳. از دیکشنری ارتقا یافته برای حذف نویز یک داده مشخص (مثلاً داده شماره ۱) استفاده کن.

۴. نسبت کنتراست به نویز CNR را برای نتیجه ی حاصل

محاسبه کن

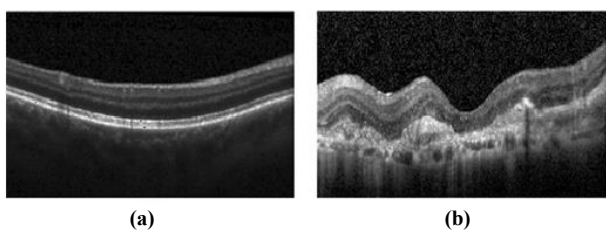
۵. $\text{Temp}(i) = \text{CNR}$

(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 سلول مای دارای بازتابندگی بالا یعنی HRC، هدف اصلی در الگوریتم های مسطح سازی به شمار می آید. این مجموعه سلول ها به بخش های داخلی و خارجی سلول های نور پذیر به انضمام سلول های رنگ دانه ای شش ضلعی اطلاق می گردد. 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 می سازد. در روش نخست اسیرینی و همکارانش در [۳۰] با سوار کردن لایه ی RPE بر روی چندجمله ای مرتبه ی دوم نواحی شبکیه را هم تراز و سپس شبکیه را مسطح کردند، که اگر لایه ی RPE شبکیه در حد زیادی آسیب دیده باشد (شکل ۲ قسمت b) به علت عدم تطابق با یک منحنی مناسب دچار اشکال می شود.



شکل ۲: تصاویر OCT از شبکیه: (a) نمایی از لایه ی RPE واضح (b) لایه ی RPE تخریب شده
Fig. 2: OCT images of the retina: (a) View of the clear RPE layer (b) the RPE layer damaged

اصلی کار ما در طبقه بندی تصاویر OCT، الگوریتم KSVD است، در پیش پردازش نیز بنای کار را بر اساس این الگوریتم قرار می دهیم و به تشریح آن می پردازیم.

از آنجایی که عموماً در کاربردهای پردازش تصویر تعداد پیکسل ها بسیار زیاد می باشد و اعمال الگوریتم روی تمام پیکسل ها به روش ذکر شده بسیار زمان بر بوده و به دلیل از دست رفتن اطلاعات محلی، غیردقیق نیز می باشد، معمولاً در پردازش تصاویر از تکه های کوچک تری با اندازه از پیش تعیین شده استفاده می گردد که با/ بدون همپوشانی از تصویر انتخاب می گردند.

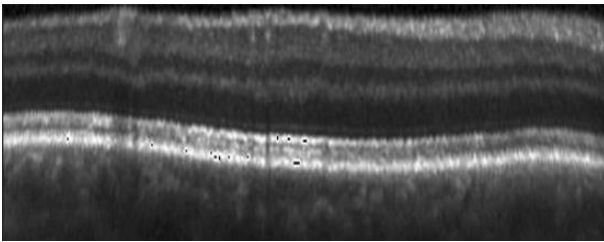
اگر هر تکه ی کوچک را $x_{ij} = R_{ij}X$ با اندازه ی $\sqrt{n} \times \sqrt{n}$ در نظر بگیریم، ماتریس سازنده ی تکه مای کوچک یعنی R_{ij} دارای اندازه ی $n \times N$ بوده و با اعمال روی تصویر با اندازه ی $\sqrt{N} \times \sqrt{N}$ ، بلوک ij را از آن استخراج می کند.

در این صورت الگوریتم K-SVD را می توان به صورت زیر تعمیم داد. در مرحله ی اول فرض می کنیم دیکشنری D نامعلوم است و مسئله به صورت زیر مطرح می شود:

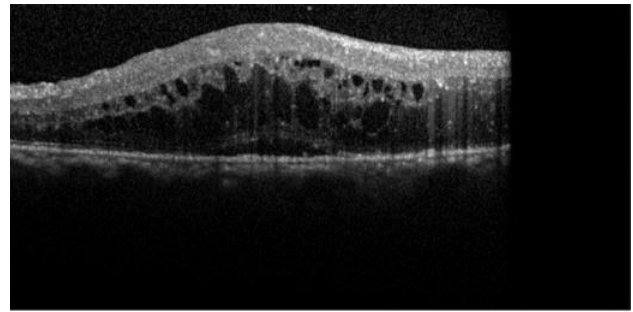
$$\left\{ \hat{D}, \hat{a}_{ij}, \hat{X} \right\} = \arg \min_{\hat{D}, \hat{a}_{ij}, \hat{X}} \gamma X - Y_2^2 + \sum_{ij} \mu_{ij} a_{ij0} + \sum_{ij} D a_{ij} - R_{ij} X_2^2 \quad (10)$$

که در آن عبارت اول تضمین کننده ی نزدیکی بین نسخه ی نویزی تصویر Y و نسخه ی بعد از حذف نویز (نامعلوم) X می باشد. عبارات دوم و سوم نمایانگر تنگ بودن ضرایب و نزدیکی پاسخ بازنمایی شده با دیکشنری به هر تکه ی کوچک می باشد. با در نظر گرفتن هر تکه ی کوچک به صورت $x_{ij} = R_{ij}X$ ، ضرایب تنگ در بازنمایی آن تکه ی کوچک بر روی دیکشنری D می باشد. در مرحله ی دوم فرض می شود D و X ثابت هستند و بازنمایی تنگ محاسبه می گردد. با در دست داشتن مقادیر ضرایب بازنمایی، دیکشنری با استفاده از روش K-SVD به روز می شود.

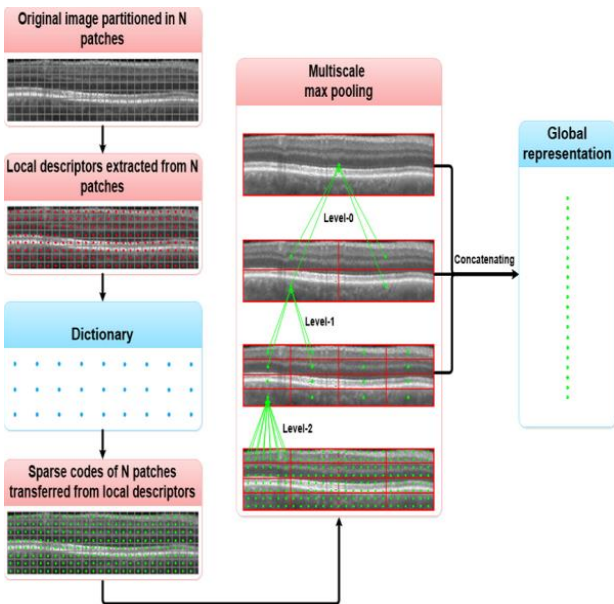
برای ساخت دیکشنری پایه تخصیص یافته، می توان از دو روش استفاده نمود. روش اول، انتخاب تصادفی یک نمونه تصویر از بین نمونه های موجود و اعمال آموزش دیکشنری بر روی آن نمونه خاص می باشد. نتایج حاصل نشان می دهد که نمونه های مختلف، منجر به نتایج حذف نویز مختلفی می شود، این موضوع ممکن است تعبیر صورت گرفته از عملکرد 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 آن ها صورت می گیرد و تعداد مورد انتخاب به نحوی می باشد که کمترین وابستگی به انتخاب داده های مجزا در آن وجود داشته باشد. برای این منظور، ابتدا یک دیکشنری اولیه انتخاب شده، آموزش دیکشنری روی داده اول صورت می پذیرد. دیکشنری حاصل، سپس به عنوان دیکشنری اولیه برای آموزش یک دیکشنری دیگر (دیکشنری به روز شده) بر روی داده بعدی به کار می رود. این روند تا رسیدن به یک نسخه دیکشنری ارتقا یافته برای استفاده در حذف نویز تمام داده های متعلق به آن کلاس به کار



شکل ۵: تصویر هم‌تراز شده‌ی OCT
Fig. 5: Aligned OCT image



شکل ۳: تصویر OCT شبکیه که در آن بخش آسیب‌دیده متورم شده است.
Fig. 3: The retinal OCT image in which the damaged part swells

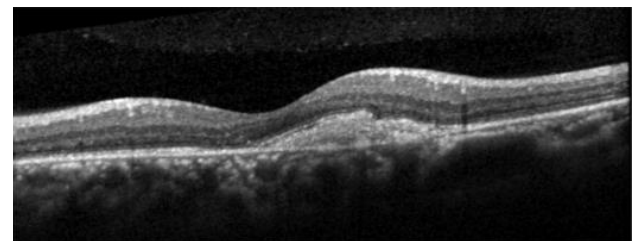


شکل ۶: بازنمایی تصویر با استفاده از عملیات تجزیه به تکه‌های کوچک و جمع‌آوری با تابع ماکزیمم [۱۰]

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

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. باین‌حال 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 تنک 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 [۱۷]. آن‌ها پیشنهاد دادند مرحله‌ی کدگذاری تنک در سطح پیچ‌ها با سایز کوچک‌تر از تصویر انجام شود و به دنبال آن نیز مراحل جمع‌آوری



شکل ۴: تصویر OCT شبکیه که در آن لایه‌ی RPE شبکیه 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

درروش دوم، لیو همکاران [۲۴] با استفاده از انطباق کل تصویر OCT شبکیه با یک چندجمله‌ای مرتبه دوم نواحی شبکیه را هم‌تراز می‌کنند. اگرچه روش هم‌ترازی آن‌ها می‌تواند برای تصاویر شبکیه‌ای که RPE آن‌ها تخریب‌شده کارساز باشد ولی نقطه‌ضعف روش موجود این است که اثربخشی آن بر روی نوع خاصی از بیماری شبکیه که در آن بخش آسیب‌دیده متورم شده است (شکل ۳)، تا حد زیادی کاهش می‌یابد.

سوم، تصاویر OCT زیادی وجود دارند که در آن‌ها لایه‌ی RPE شبکیه مستقیم است اما با یک زاویه مشخص با خط افقی قرار گرفته است (شکل ۴). در این حالت، انطباق شبکیه با خط مستقیم ممکن است بهتر از یک سهمی باشد.

برای حل مشکلات سه‌گانه ذکر شده در بالا و هم‌ترازی شبکیه با استفاده از روشی پایدارتر، در این مقاله بر مبنای روشی که یانگویی و همکاران در سال (۲۰۱۷) [۳۳] ارائه دادند عملیات هم‌ترازی انجام می‌شود به این ترتیب که، ابتدا تصویر شبکیه 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 شبکیه است، انتخاب می‌گردد. نهایتاً به‌طور خودکار بین یک چندجمله‌ای مرتبه دوم و یک خط مستقیم جهت انطباق با مجموعه نقاط انتخاب‌شده، تصمیم‌گیری می‌شود.

تجزیه پیچ و جمع‌آوری

روش‌های یادگیری دیکشنری برای بازسازی یا طبقه‌بندی تصاویر

کدهای تنک \hat{x}_k را می‌توان به‌عنوان بردار ۳ بعدی X با عمق برابر با تعداد اتم‌ها در دیکشنری D در نظر گرفت.

۳. بردار ۳ بعدی حاصل به‌عنوان ورودی لایه بعدی مورد استفاده قرار می‌گیرد.

(۱) و (۲) را می‌توان برای هر تعدادی از لایه‌ها تکرار کرد.

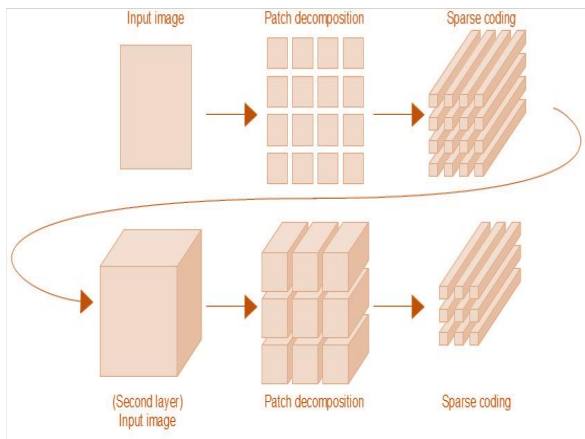
برای بهینه‌سازی معماری چندلایه باهدف طبقه‌بندی، پیدا کردن دیکشنری مای بهینه در هر لایه لازم است و جهت افزایش دقت طبقه‌بندی و کار آیی آن اهمیت زیادی دارد. در این معماری برای به دست آوردن بهترین دیکشنری در اولین لایه، با استفاده از الگوریتم K -maens در هر کلاس عملیات خوشه‌بندی را انجام داده و مراکز خوشه‌ها به‌عنوان اتم‌های دیکشنری انتخاب می‌شوند. در آخر با تجمیع 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 استفاده‌ی بهینه از برچسب کلاس‌ها و انتخاب دیکشنری مناسب، فرایند طبقه‌بندی را تا حد زیادی تأثیرگذار می‌نماید؛ شکل ۷ نمای کلی از این فرایند است.

مدل‌سازی ریاضی روش پیشنهادی

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

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

$$\begin{aligned} (\hat{x}_k^{(q)}, A, D) = \operatorname{argmin}_{x, A, D} y_k^{(q)} - Dx^{(q)2} \\ + \alpha Z^{(q)} - Ax^{(q)2} \quad s.t \quad \forall i, x_{i0} < T \end{aligned} \quad (11)$$



شکل ۷. معماری دولایه‌ی روش پیشنهادی

Fig. 7: Two-layered architecture of the proposed method

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 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 تنک مای همپوشانی محلی y_k تجزیه می‌کنیم که به‌وسیله دیکشنری D و با محاسبه ضرایب \hat{x}_k تنک کدگذاری می‌شوند. این تنک‌ها به‌طور مساوی بر روی یک توری مشبک محلی سازی می‌شوند، به‌طوری‌که موقعیت نسبی از تمام تنک‌ها ایجاد گردد. سپس کدهای \hat{x}_k در یک یا چند مرحله، روی گروه کوچکی از تنک‌ها به‌وسیله‌ی تابعی ترکیب می‌شوند (مثلاً تابع ادغام بیشینه) به‌طوری‌که در هر مرحله منجر به کاهش تعداد تنک مای کوچک می‌گردد (به شکل ۶ نگاه کنید).

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 تعیض 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. به‌این‌ترتیب که به‌جای کدگذاری تصویر اصلی از کدگذاری تنک 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 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 فرایند را می‌توان چندین بار تکرار کرد که در هر بار تکرار، لایه‌ی جدید با دیکشنری جدید طراحی می‌شود.

چارچوب کلی روش پیشنهادی

روش پیشنهادی چندلایه را می‌توان به شرح زیر توصیف کرد:

۱. ابتدا هر تصویر ورودی به مجموعه‌ای از تنک مای همپوشانی کوچک y_k با حفظ محلی سازی مکانی خود، تجزیه می‌شوند.
۲. هر پیچ y_k با استفاده از اولین دیکشنری D کدگذاری می‌شود. از آنجایی که محلی سازی مکانی این تنک‌ها حفظ شده است، مجموعه‌ی

تابع را انتخاب نموده‌ایم و طبقه‌بندی کننده ی منتخب برای این خروجی را تابع سافت مکس در نظر می‌گیریم. اگر ما مسئله‌ی طبقه‌بندی را با C کلاس فرض کنیم، تابع کراس آن‌روپی به‌صورت زیر محاسبه می‌گردد:

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

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

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

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 (۸) و (۹) از بخش ۲-۱ انجام شد و با جانشینی $\hat{x}_k^{(q)}$ به‌جای $x_k^{(q)}$ و $y_k^{(q)}$ به‌جای y_k محاسبه می‌گردد. برای دوتایی (Y, l) که معرف سیگنال ورودی به همراه برچسب آن بود، گرادیان L با توجه به دیکشنری لایه‌ی آخر یعنی $D^{(q)}$ طبق فرمول (۸) و (۱۱) محاسبه می‌گردد و به‌صورت زیر داریم:

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

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

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

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

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

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

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

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 هزینه ی 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 به‌صورت زیر تجزیه شود:

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

$$\min_{D^{(0)}, D^{(2)}, \dots, D^{(Q)}, W} \sum_{k=1}^n L(l_k, \hat{x}_k^{(Q)}, W) \quad (10)$$

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

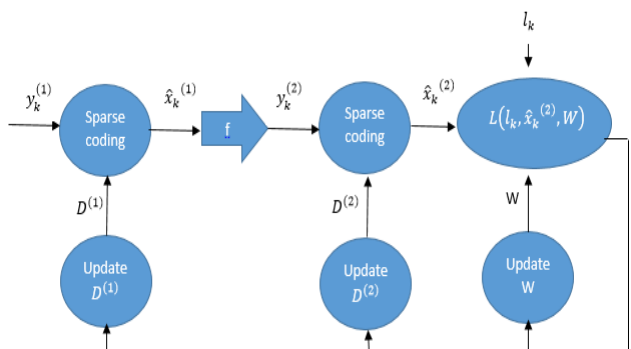
در فرمول فوق $Z^{(q)} = \{z_1^{(q)}, z_2^{(q)}, \dots, z_n^{(q)}\} \in \mathbb{R}^{m \times n}$ کد مای تنک متمایز سیگنال ورودی $y_k^{(q)}$ ، در لایه‌ی q ام برای طبقه‌بندی هستند؛ و نیز $z_i^{(q)} = \{0, \dots, 1, \dots, 0\}^T \in \mathbb{R}^m$

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. در این فرمول A نیز یک ماتریس تبدیل خطی می‌باشد؛ ما تبدیل خطی $g(A, X) = AX$ را در نظر می‌گیریم که کد مای تنک معمولی X را به کد مای تنک متمایز در فضای \mathbb{R}^m منتقل می‌کند.

عبارت $Z^{(q)} - AX^{(q)2}$ در فرمول (۱۱) 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 هزینه ی طبقه‌بندی با توجه به دیکشنری های هر لایه، ما از الگوریتم پس انتشار خطا استفاده نموده‌ایم [۳۵]؛ بنابراین لازم است گرادیان نسبت به دیکشنری مای $\{D^{(1)}, D^{(2)}, \dots, D^{(Q)}\}$ محاسبه گردد. ما فقط از خروجی لایه‌ی آخر یعنی $\hat{x}_k^{(Q)}$ به‌عنوان معیارها و شاخص مای طبقه‌بندی استفاده کرده ایم و تابع هزینه ی طبقه‌بندی روی مجموعه‌ی داده‌های آموزشی ورودی از n تصویر کمینه می‌شود (شکل ۸ را ببینید).

همچنین برای میزان تلفات طبقه‌بندی با توجه به نتایج خوبی که در طبقه‌بندی مای چند کلاسه از تابع کراس آن‌روپی دیده‌شده است، این



شکل ۸: معماری پیشنهادی با دو لایه
Fig. 8: Proposed structure with 2 layers

درجایی که:

$$\tilde{x}_R^k = \sum(1,1) V(:,1) \quad (24)$$

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 به‌طور هم‌زمان یاد گرفته می‌شوند، این روش از مینیمم سازی محلی اجتناب می‌کند و قابل اجرا در مسائلی با تعداد کلاس بالا می‌باشد. به‌علاوه این روش به ما اجازه می‌دهد که به راحتی 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 پارامترهای D_0 و A_0 برای الگوریتم ما مقداردهی اولیه شوند. برای دیکشنری اولیه، در روش پیشنهادی، ابتدا در هر کلاس با استفاده از الگوریتم خوشه‌بندی k -means [۴۳]، عملیات خوشه‌بندی را انجام داده و مراکز خوشه‌ها را به‌عنوان نماینده‌ی آن کلاس در ستون مای دیکشنری اولیه ما قرار می‌دهیم. بنابراین ستون k ام دیکشنری یعنی d_k مرکز خوشه مای ایجادشده در کلاس با برجسب k می‌باشد. برجسب هر ستون دیکشنری متناظر با کلاسی است که از عناصر آن تولیدشده 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 استفاده‌شده است [۳۷]، که در آن:

$$A = \operatorname{argmin}_{A,Z} z^{(q)} - Ax^{(q)2} + \gamma A_2^2 \quad (25)$$

و منجر به جواب زیر می‌شود:

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

طبقه‌بندی

ما $D = (d_1, d_2, \dots, d_m)$ و $W = (w_1, w_2, \dots, w_m)$ با به کار بردن الگوریتم K -SVD به دست آورده ایم، ولی به سادگی برای طبقه‌بندی داده‌های تست به کار نمی‌آیند لذا D و W با استفاده از نرم ۲ با هم در D_{new} نرمال می‌شوند، یعنی $\forall k, \|d_k^t, \sqrt{\alpha \alpha_k}\|_2 = 1$. دیکشنری مطلوب \hat{D} ، پارامترهای تبدیل \hat{A} و پارامترهای طبقه‌بندی کننده \hat{W} به‌صورت زیر محاسبه می‌شوند:

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

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

و در غیر این صورت،

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

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.

رویکرد بهینه‌سازی

به‌منظور پیدا کردن راه‌حل بهینه برای تمام پارامترها به‌صورت هم‌زمان، مبنای عملکرد ما الگوریتم K -SVD قرار داده‌شده است، لذا معادله (۱۱) را به‌صورت زیر بازنویسی می‌کنیم:

$$\begin{aligned} & (\hat{x}_k^{(q)}, A, D) \\ & = \operatorname{argmin}_{x,A,D} \left\| \left(\frac{y_k^{(q)}}{\sqrt{\alpha Z^{(q)}}} \right) - \left(\frac{D}{\sqrt{\alpha A}} \right) x^{(q)} \right\|_2^2 \\ & \text{s.t. } \forall i, x_{i0} < T \end{aligned} \quad (21)$$

حال اگر فرض کنیم، $y_{new} = \left(\frac{y_k^{(q)}}{\sqrt{\alpha Z^{(q)}}} \right)$ و $c = \left(\frac{D}{\sqrt{\alpha A}} \right)$ ،

$x^{(q)} = X$ ، حل مسئله بهینه‌سازی فوق به حل معادله‌ی زیر منجر می‌گردد:

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

و این دقیقاً مسئله‌ی است که K -SVD آن را حل می‌کند [۴]، [۲۱] و [۲۹]. با توجه به این الگوریتم، d_k و ضرایب نظیرش یعنی k امین سطر در X که به‌صورت x_R^k معرفی می‌شود، در هر لحظه به‌روز می‌شوند.

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

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

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

$$d_k = U(:,1)$$

برای بررسی پیچیدگی روش پیشنهادی در این مقاله متوسط و انحراف از معیار زمان مورد نیاز برای تحلیل هر اسلایس OCT بر روی یک کامپیوتر شخصی با مشخصات ویندوز ۳۲ بیتی و پنتیوم اینتل دو هسته‌ای با رم ۴ گیگابایت و پردازنده‌ی ۳ 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 کلاسه روی ۴۲ 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 ۲ نشان می‌دهد، صد درصد از ۳۰ حجم تصاویر OCT کلاس مای AMD و DME دقیقاً درست طبقه‌بندی شدند و ۹۵٫۸۵٪ از تصاویر نرمال طبقه‌بندی صحیح داشتند که نتایج بسیار خوبی در مقایسه با روش‌های مشابه موجود می‌باشد.

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، ۵، ۱۰ و ۱۵ بار تکرار کردیم تصاویر آموزشی متفاوتی را به‌صورت 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 ۳ آورده شده است.

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 به‌منظور اعتبارسنجی الگوریتم طبقه‌بندی، در جدول ۴ روش پیشنهادی در این مقاله با سه روش CNN و SVM و SRC مقایسه شده است.

روش اول تشخیص خودکار بیماری بر اساس طبقه‌بندی کننده ماشین بردار پشتیبان (SVM) چندلایه [۳۸]، روش دوم مبتنی بر تصاویر خام و لایه مای تقسیم‌شده توسط شبکه مای عصبی کانولوشنی [۳۹] و

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 طبقه‌بندی کننده‌ی 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:

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

آزمایش‌ها و نتایج

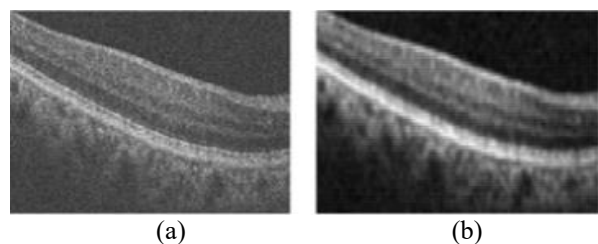
در این بخش به‌منظور بررسی نتایج و اعتبارسنجی الگوریتم پیشنهادی از مجموعه داده‌های دام استفاده شده است. این مجموعه داده توسط اسرینواسان و همکاران [۳۰] منتشر شده که جهت تحلیل تصاویر OCT در دانشگاه مای دام، هاروارد و میشیگان مورد استفاده قرار گرفته است. مجموعه داده‌ی دام دربردارنده‌ی اسکن‌های حجمی شامل ۴۵ نوع حجم می‌باشد که ۱۵ مورد طبیعی، ۱۵ مورد بیمار AMD و ۱۵ مورد بیمار DME و تعداد اسکن‌های OCT در هر حجم بین ۳۶ تا ۹۷ تصویر متغیر است.

در فاز پیش‌پردازش تصاویر به‌منظور سنجش عملکرد روش حذف نویز موردبررسی، 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 ۱ و تحلیل بصری شکل ۹ بیانگر کارایی بالای روش انجام‌شده می‌باشد:

جدول ۱: 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 ۹: (a) تصویر نویز دار، (b) تصویر حذف نویز شده
Fig. 9: a) Noisy Image b) Noise Removed Image

جدول ۲: 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%

جدول ۳: 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%

نتیجه‌ی این تحقیق ارائه‌ی سیستمی اتوماتیک به‌منظور تشخیص برخی ناهنجاری مای شبکیه بود به‌طوری‌که 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 ورم ماکولای دیابتی یعنی DME را تفکیک و در تشخیص پاتولوژی شبکیه با دقت بسیار بالایی به پزشک معالج کمک کند. به‌عنوان یک پیشنهاد تحقیقاتی و کاربردی برای متخصصین و آیندگان با تعمیم این روش به کلاس مای بیشتر می‌توان تمام بیماری مای شبکیه را پوشش داد و از آن به‌عنوان ابزاری بالقوه مؤثر در تشخیص کامپیوتری و غربالگری بیماری‌های شبکیه و یا در حوزه مای وسیع‌تر چشم استفاده نمود.

بی‌نوشت

- 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

مشارکت نویسندگان

تمام نویسندگان به نسبت سهم برابر در این پژوهش مشارکت داشتند.

تشکر و قدردانی

از تمام کسانی که ما را در انجام این پژوهش یاری رساندند تشکر و قدردانی داریم.

تعارض و منافع

«هیچ گونه تعارض منافع توسط نویسندگان بیان نشده است.»

منابع و مآخذ

[1] Schirmacher F, Köhler T, Husvogt L. Quantile sparse image prior for spatio-temporal denoising of retinal OCT data. *Medical Image Computing and Computer-Assisted Intervention*. 2017; 12(4): 83-91.

جدول ۴: 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%

روش سوم طبقه‌بندی مبتنی بر نمایش تنک یا SRC [۳۶] می‌باشند. 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.

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 می‌کند 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 آموزشی، اطلاعات برجسب 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 به‌منظور طبقه‌بندی شد. ایده اصلی روش پیشنهادی این بود که لایه کدگذاری تنک را تکرار کرده و از K-SVD سازگار با برجسب 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. به‌این‌ترتیب که به‌جای 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 اتکال آن به فرایند تقطیع لایه مای شبکیه بود چراکه بیماری مای شبکیه، لایه‌ها را مدام تغییر می‌دهند و کار تقطیع لایه‌ها را پیچیده می‌کنند.

- ers in macular OCT images of patients with retinitis pigmentosa. *Biomed. Opt. Express*. 2011; 2(9): 2493–2503.
- [18] Lang A. et al. Retinal layer segmentation of macular OCT images using boundary classification. *Biomed. Opt. Express*. 2013; 4(7): 1133–1152.
- [19] Garcia-Allende PB et al. Morphological analysis of optical coherence tomography images for automated classification of gastrointestinal tissues. *Biomed. Opt. Express*. 2011; 2(10): 2821–2836.
- [20] Chan S. et al. Multi-layer Dictionary Learning for Image Classification. *Advanced Concepts for Intelligent Vision Systems*. 2016; 4(15): 132-143.
- [21] Jiang Z, Lin Z, Davis L. Learning a discriminative dictionary for sparse coding via label consistent K-SVD. *IEEE Computer Society Conference on Computer Cision and Pattern Recognition*. 2011; 6(12), 1697-1704.
- [22] Pande P. et al. Automated classification of optical coherence tomography images for the diagnosis of oral malignancy in the hamster cheek pouch. *J. Biomed. Opt.* 2014; 19(8), 356-368.
- [23] Alsaih K, et al. Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images. *Bio Medical Engineering*. 2017; 16(1): 447-455.
- [24] Liu Y. et al. 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.* 2011; 15(5): 748–759.
- [25] Zheng Y, Hijazi MHA, Coenen F. Automated ‘disease/no disease’ grading of age-related macular degeneration by an image mining approach. *Invest. Ophthalmol. Visual Sci.* 2012; 53(13): 8310–8318.
- [26] Hijazi MHA, Coenen F, Zheng Y. Data mining techniques for the screening of age-related macular degeneration. *Knowled e-Based Syst.* 2012; 29(1): 83–92.
- [27] Kafieh R, Rabbani H. Optical coherence tomography noise reduction over learned dictionaries with introduction of complex wavelet for noise reduction. *Proceedings,Wavelets and Sparsity*. 2013; 8858(26): 238-247.
- [28] Albarrak A. et al. Volumetric image mining based on decomposition and graph analysis: an application to retinal optical coherence tomography. *Computational Intelligence and Informatics*. 2012; 1344(64): 245-256.
- [29] Wang Q et al. Synthesis K-SVD based analysis dictionary learning for pattern classification. *Multimedia Tools and Applications*. 2017; 12(4): pp1-24.
- [30] Srinivasan PP, et al. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomed. Opt. Express*. 2014; 5(10): 3568–3577.
- [31] Esmaeili M, Dehnavi AM, Rabbani H. Speckle noise reduction in optical coherence tomography using two-dimensional curvelet-based dictionary learning. *Journal of Medical*. 2017; 17(2), 86-91.
- [32] Yankelevsky Y. et al. Structure-aware classification using super-
 [2] Sarbjeet Kaur V, Banga K. Review paper of character recognition using image processing. *International Journal of Computer Science & Engineering Technology*. 2014; 5(1): 254-268.
- [3] Fan Meng et al. A sparse dictionary learning-based adaptive patch in painting method for thick Clouds removal from high-spatial resolution remote sensing imagery. *Sensors*. 2017; 17(9): 123-135.
- [4] Zheng H, Zhu J, Yang Z, Jin Z. Effective micro-expression recognition using relaxed K-SVD algorithm. *International Journal of Machine Learning*. 2017; 8(6): 2043-2049.
- [5] Chan Wai Tim S, Rombaut M, Pellerin D. Rejection-based classification for action recognition using a spatio-temporal dictionary. *Advanced Concepts for Intelligent Vision Systems*. 2015; 15(8), 522-533.
- [6] Zhang Z. et al. Discriminative-dictionary-learning-based multi-level point-cluster features for ALS point cloud classification. *IEEE Transactions on Geoscience and Remote Sensing*. 2016; 54(12): 334-347.
- [7] Zhang Z. et al. Sparse codes auto-extractor for classification: A joint embedding and dictionary learning framework for representation. *IEEE Signal Processing Society*. 2016; 64(14): 553-561.
- [8] de Moura J, Novo J, Rouco J. Automatic Detection of Blood Vessels in Retinal OCT Images. *Biomedical Applications Based on Natural and Artificial Computing*. 2017; 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. Fully automated macular pathology detection in retina optical coherence tomography images using sparse coding and dictionary learning. *Journal of Biomedical Optics*. 2017; 22(1): 214-225
- [11] Rickman CB. et al. Dry age-related macular degeneration: mechanisms, therapeutic targets, and imagingdry AMD mechanisms, targets, and imaging. *Invest. Ophthalmol. Visual Sci.* 2013; 54(14): 68– 80.
- [12] Antonym BJ et al. A combined machine-learning and graph-based framework for the segmentation of retinal surfaces in SD-OCT volumes. *Biomed. Opt. Express*. 2013; 4(12): 2712–2728.
- [13] Carass A. et al. Multiple-object geometric deformable model for segmentation of macular OCT. *Biomed. Opt. Express*. 2014; 5(4): 1062–1074.
- [14] Chium S J. et al. Validated automatic segmentation of AMD pathology including drusen and geographic atrophy in SD-OCT images. *Invest. Ophthalmol. Visual Sci.* 2012; 53(1): 53–61.
- [15] Sun Y. et al. 3D automatic segmentation method for retinal optical coherence tomography volume data using boundary surface enhancement. *J. Innovative Opt. Health Sci*. 2016; 9(2): 165-177.
- [16] Vermeer K. et al. Automated segmentation by pixel classification of retinal layers in ophthalmic OCT images. *Biomed. Opt. Express*. 2011; 2(6): 1743–1756.
- [17] Yang, Q. et al. Automated segmentation of outer retinal lay-

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

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

[43] Aggarwal CC, Reddy CK. *Data clustering, algorithms and applications*. New York: Wiley; 2013.

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

vised dictionary learning. *Computer Vision and Pattern Recognition*. 2016; 10(5): 679-692.

[33] Sun Y, Li SH, Sun ZH. Fully automated macular pathology detection in retina optical coherence tomography images using sparse coding and dictionary learning. *J. Biomed. Opt.* 2017; 22(1): 160-171.

[34] Hassan T, et al. Review of OCT and fundus images for detection of Macular Edema. *Imaging Systems and Techniques*. 2015; 155(45): 234-245.

[35] Shalev-Shwartz S, Shamir O. Failures of gradient-based deep learning. *International Journal of Machine Learning*. 2017; 17(3): 456-466.

[36] Javidi M, Pourreza HR, Harati A. (2017). Vessel segmentation and microaneurysm detection using discriminative dictionary learning and sparse representation. *Computer Methods and Programs in Biomedicine*. 2017; 139(1): 93-108.

[37] Geimer T, et al. A Kernel Ridge regression model for respiratory motion estimation in radiotherapy. *Bildverarbeitung für die Medizin*. 2017; 13(4): 155-160.

[38] Khalid S, et al. Fully automated robust system to detect retinal edema, central serous chorioretinopathy, and age related macular degeneration from optical coherence tomography images. *J. Biomed. Opt.* 2017; 71(48): 278-293.

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

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

پیوست‌ها

Supervised Dictionary Learning and Sparse Representation

S-DLSR

Sparse Representation-based Classification

SRC

Hyper-reflective complex

HRC

Contrast to noise ratio

CNR

Texture Preservation

TP

Edge Preservation

EP

Equivalent Number of Looks

ENL

$$X_F^2 = \sum_{i=1}^n \sum_{j=1}^k x_{ij}^2$$

نرم فروبنیوس

Citation: (Vancouver): Montazeri A, Shamsi M, Dianat R. [Automatic recognition of retinal diseases using mathematical models of image processing, based on multilayer-dictionary learning]. *Tech. Edu. J.* 2019; 13(2): 234-248.

 <http://dx.doi.org/10.22061/jte.2018.3434.1872>



COPYRIGHTS

©2019 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.