

بهبود روش درهم‌تنیدگی تصویر مبتنی بر یادگیری با در نظر گرفتن وزن‌های مختلف برای

زمینه و پیش‌زمینه

محمود امین‌طوسی

دانشکده ریاضی و علوم کامپیوتر، دانشگاه حکیم سبزواری، سبزوار، ایران

چکیده

استخراج دقیق پیش‌زمینه در تصویر، درهم‌تنیدگی تصویر نامیده می‌شود. از جمله روش‌های معروف در این حوزه روش درهم‌تنیدگی مبتنی بر یادگیری است. در این شیوه داده‌های آموزشی دو طبقه در قالب یک نقشه سه‌گانه که شامل نقاط آموزشی بر چسب‌دار (زمینه و پیش‌زمینه) و نقاط بدون برچسب هستند، به سامانه داده می‌شوند. میزان تعلق هر پیکسل به زمینه یا پیش‌زمینه بر اساس داده‌های آموزشی و کمینه‌سازی یک تابع هدف مشخص می‌شود. در تابع هدف تمامی کارهای منتشرشده، میزان جریمه تخطی از مقادیر درست برای داده‌های آموزشی - خواه متعلق به زمینه باشند یا پیش‌زمینه - یکسان در نظر گرفته شده است. در این مقاله با در نظر گرفتن وزن متفاوت برای داده‌های آموزشی دو طبقه، این شیوه بهبود داده شده و کارایی آن در دو کاربرد متفاوت، نشان داده شده است. کاربرد نخست، دقیق‌تر کردن جداسازی متن از تصویر و کاربرد دوم، دقیق‌تر کردن خروجی روش‌های استخراج رگ‌های خونی شبکه چشم در کناره‌های رگ‌های شناسایی شده است. در کاربرد نخست، متنی فارسی که بر روی زمینه دارای بافت ناهموار درج شده با یک روش معمول آستانه‌گذاری استخراج شده و سپس خروجی قطعه‌بندی‌شده با روش پیشنهادی دقیق‌تر شده است. در کاربرد دوم، ابتدا با یک روش موجود شناسایی رگ، قسمت‌هایی از تصویر که به احتمال زیاد متعلق به دو دسته رگ و غیر رگ هستند، بر چسب‌گذاری می‌شوند. تعیین دقیق‌تر تعلق پیکسل‌های مرز رگ‌های استخراج‌شده به هر یک از دو دسته رگ یا غیر آن توسط روش پیشنهادی انجام می‌شود. نتایج کمی و دیداری، کارایی شیوه پیشنهادی را نشان داده است.

واژگان کلیدی: قطعه‌بندی تصویر، درهم‌تنیدگی تصویر، استخراج رگ، یادگیری ماشین، استخراج متن از تصویر

Enhancement of Learning Based Image Matting Method with Different Background/Foreground Weights

Mahmood Amintoosi

Faculty of Mathematics and Computer Science, Hakim Sabzevari University, Sabzevar, Iran

Abstract

The problem of accurate foreground estimation in images is called Image Matting. In image matting methods, a map is used as learning data, which is produced by those pixels that are definitely foreground, definitely background, and unknown. This three-level pixel map is often referred to as a trimap, which is produced manually in alpha matte datasets. The true class of unknown pixels will be estimated by minimizing of an objective function. Several methods for image matting has been proposed. The learning-based method is one the pioneering works which is the basis of many other approaches in the field of image matting. In this method it is assumed that each pixel's alpha value is a linear combination of its associated neighboring pixels. A Laplacian matrix in the objective function shows the similarity of the pixels. The coefficients of the linear combination are estimated with a local learning process by minimizing a quadratic cost function. The method of Lagrange multiplier and ridge regression technique are used for estimation of alpha values. In this objective

* Corresponding author

* نویسنده عهده‌دار مکاتبات

function the violation of the predefined training pixels' alpha values from their true values is controlled by a penalty term. Considering this coefficient as infinity, forces the matte (alpha) value to be 1 for the labeled foreground pixels and 0 for background. The weight of this penalty term still was taken equal for all training samples. In this paper the performance of the matting method is increased by considering different weights for different learning pixels. The good performance of the proposed method is demonstrated in two applications. The first application is improving the quality of a text extraction method and the second application is enhancement of an eye retinal segmentation system. In the first application, a Persian text which is fused onto a textured background is extracted by a thresholding method. After that the segmented output is enhanced by the proposed matting method. In the second application, segmentation is done with an existing vessel extraction method. The edges' pixels of detected vessels that may be classified inaccurately are classified by the proposed image matting method. Subjective and objective comparisons show the better performance of the proposed method.

Keywords: Image Matting, Image Segmentation, Machine Learning, Text Extraction, Retinal Vessel Segmentation

پیش‌زمینه و میزان تیرگی پیش‌زمینه (alpha matte) در آن نقطه برآورد شود. این مسأله یک مسأله بدطرح^۴ است که پژوهش‌های متعددی را در سالیان اخیر به خود معطوف نموده است [2-8].

فرض کنید مجموعه پیکسل‌های I با $\Omega = \{1, \dots, n\}$ نشان داده شوند، که n تعداد کل پیکسل‌هاست. برای زیرمجموعه‌ای از پیکسل‌های تصویر، تعلق قطعی آنها به دو زیرمجموعه زمینه و پیش‌زمینه مشخص است که به‌عنوان داده‌های برچسب‌گذاری شده $\Omega_l \subset \Omega$ در نظر گرفته می‌شوند. مقدار α برای این دو دسته صفر و یک است که با نواحی سیاه و سفید در قالب نقشه سه‌گانه^۵ به الگوریتم داده می‌شوند. برآورد آلفا به‌عنوان محاسبه مقدار α برای مجموعه‌ای از پیکسل‌های نامعلوم (برچسب‌گذاری نشده) $\Omega_u = \Omega - \Omega_l$ تعریف می‌شود که از آن با عبارات alpha matting یا digital matting یا image matting نام برده می‌شود [7]، و در این نوشتار از آن با نام «درهم‌تیدگی تصویر» یاد خواهد شد.

در یک فرآیند تعاملی، کاربر قسمت‌هایی از هر دو زیرمجموعه را مشخص و الگوریتم مربوطه تصویر را به دو قسمت زمینه و شیء، کاربر از طریق ایجاد یک نقشه سه‌گانه، سامانه را آموزش می‌دهد. نقشه به‌نحوی ایجاد می‌شود که آن نواحی از تصویر که مطمئناً متعلق به زمینه هستند با رنگ سیاه، نواحی متعلق به شیء (پیش‌زمینه) با رنگ سفید و نواحی دارای عدم قطعیت با رنگ خاکستری مشخص شوند. نقاط متناظر با نواحی سیاه و سفید این نقشه، در تصویر اصلی به‌عنوان داده‌های آموزشی در نظر گرفته می‌شوند و هدف این روش تخمین برچسب سایر نقاط (پیکسل‌های متناظر با نقاط خاکستری) با استفاده از این داده‌های آموزشی است.

۱- مقدمه

شناسایی اشیای موجود در تصویر یکی از معروف‌ترین مسائل در حوزه پردازش تصویر و بینایی ماشین است که کاربردهای عملی بسیاری داشته و یک مرحله اصلی در بسیاری از سامانه‌های مبتنی بر پردازش تصویر محسوب می‌شود. از جمله می‌توان به کاربرد آن در سامانه‌های مراقبت بینایی، کاربردهای پزشکی و شناسایی بصری حروف اشاره کرد.

شناسایی اشیای درواقع یک مسأله قطعه‌بندی است که روش‌های متعددی برای آن تاکنون عرضه شده است؛ یکی از روش‌های بسیار کارای مرتبط با این حوزه، روش درهم‌تیدگی تصویر است^۱ [1]. در این شیوه هدف، برآورد دقیق میزان تعلق هر پیکسل تصویر به دو دسته زمینه^۲ و پیش‌زمینه^۳ است. این شیوه یکی از روش‌های کلیدی در ویرایش و مونتاژ فیلم و تصویر است که با استفاده از آن، شیئی از زمینه جدا شده و در تصویر دیگری با زمینه‌ای متفاوت قرار می‌گیرد. هر چه برآورد ناحیه پیش‌زمینه دقیق‌تر باشد، عمل ترکیب شیئی با زمینه جدید، بهتر خواهد بود. فرض کنید تصویر I ، ترکیب خطی از زمینه B و پیش‌زمینه F با میزان مشارکت α باشد:

$$I = \alpha F + (1 - \alpha)B \quad (1)$$

که α در بازه صفر و یک است. هدف پیدا کردن α ، F و B با داشتن فقط مقدار شدت روشنایی پیکسل (I) است. به عبارت دیگر با داشتن رنگ هر پیکسل، باید مقادیر رنگ‌های زمینه،

معنی فرهنگ‌لغتی “Matting” «بوریا‌بافی» است که در برخی اسناد فارسی هم استفاده شده است. از جمله مترادف‌های Matting در زبان انگلیسی، عبارت Tangle است که معانی «درهم و برهم کردن»، «ژولیدن»، «درهم پیچ کردن» و «گیرافتادن» برای آن آمده است. با توجه به مفهوم Image Matting، نگارنده، عبارت «درهم‌تیدگی تصویر» را معادل بهتری نسبت به بوریا‌بافی برای آن می‌داند.

^۴ Ill-posed

^۵ Trimap

^۲ Background

^۳ Foreground

چهار نتایج آزمایش‌های انجام‌شده را بر روی دو کاربرد متفاوت ملاحظه خواهیم کرد. مثال نخست مرتبط با دقیق‌تر کردن استخراج متن از تصویر و دومین مثال مرتبط با شناسایی بهتر رگ‌های چشم است. آخرین بخش به نتیجه‌گیری اختصاص داده شده است.

۲- روش درهم‌تندیدگی تصویر مبتنی بر یادگیری

همان‌گونه که اشاره شد، منظور از درهم‌تندیدگی تصویر، تخمین دقیق پیش‌زمینه در تصویر و ویدئو است. در این بخش روش مبتنی بر یادگیری [7] که مبنای روش پیشنهادی است، بیان خواهد شد.

در روش مبتنی بر یادگیری برای هر پیکسل $i \in \Omega$ فرض می‌شود α_i می‌تواند با ترکیب خطی از مقادیر آلفای پیکسل‌های مجاورش، برآورد شود. مقدار α برای همه پیکسل‌ها از طریق کمینه‌سازی تابع هزینه درجه دوم، به‌طور هم‌زمان برآورد می‌شود. فرض کنید $N_i = \{T_1, \dots, T_m\}$ مجموعه همسایگی پیکسل i و $\alpha_i = \{\alpha_{T_1}, \dots, \alpha_{T_j}, \dots, \alpha_{T_m}\}$ بردار مقادیر آلفای N_i باشد که اگر $T_j \in N_i$ $f_i = \{f_{iT_1}, \dots, f_{iT_j}, \dots, f_{iT_m}\}$ برای نمایش بردار ضرایب ترکیب خطی در نظر گرفته شود، آنگاه ترکیب خطی ضرایب را برای پیکسل i ، می‌توان به‌صورت زیر نوشت:

$$\alpha_i = f_i^T \alpha_i \quad (4)$$

می‌توان α_i را در معادله (۴) به‌صورت ترکیب خطی از مقادیر آلفای همه پیکسل‌ها بازنویسی کرد. اگر مقادیر آلفای همه پیکسل‌ها با بردار $\alpha = [\alpha_1, \dots, \alpha_n]$ و ضرایب با $\xi_i = [f_{i1}, \dots, f_{im}]$ نمایش داده شوند، خواهیم داشت: $\alpha_i = \xi_i^T \alpha$ با تعریف ماتریس جدید F به‌صورت $F = [\xi_1, \dots, \xi_n]$ که F $\{\xi_i\}_{i \in \Omega}$ ، معادله (۴) را می‌توان به‌صورت $\alpha = F^T \alpha$ بازنویسی کرد. اگر F معلوم باشد، آلفا از طریق کمینه‌سازی تابع درجه دوم زیر برآورد می‌شود:

$$\operatorname{argmin}_{\alpha} \|\alpha - F^T \alpha\|^2 + c \|\alpha - \alpha_i^*\|^2 \quad (5)$$

که در آن α_i^* برچسب از قبل مشخص برای داده‌های آموزشی است (نواحی سیاه و سفید در تصویر آموزشی) و α_i تخمین همین نقاط در α کمینه است. $C_{n \times n}$ ماتریس قطری است که زامین عنصر قطری، مقدار ثابت c را اختیار می‌کند، اگر $j \in \Omega_i$ باشد و بقیه عناصر قطری صفر هستند. اگر بردار α^* به‌طول n در نظر گرفته شود، که زامین عنصر آن مقدار معلوم

در [8] از روش فرم بسته^۱ برای حل مسأله استفاده شده است. در این شیوه فرض می‌شود F و B در پنجره کوچکی اطراف هر پیکسل به‌طور تقریبی ثابت هستند. این فرض اجازه می‌دهد تا رابطه (۱) به‌صورت زیر بازنویسی شده و α به‌عنوان یک تابع خطی از تصویر I به‌صورت زیر در نظر گرفته شود:

$$\alpha_i = aI_i + b \quad \forall i \in w \quad (2)$$

بنا به رابطه (۱) برای هر پیکسل i در پنجره w رابطه $I_i = \alpha_i F_i + (1 - \alpha_i) B_i$ برقرار است. داریم:

$$\alpha_i(F_i - B_i) = I_i - B_i \\ \Rightarrow \alpha_i = \frac{1}{F_i - B_i} I_i + \frac{-B_i}{F_i - B_i}$$

لذا $a = \frac{1}{F_i - B_i}$ ، $b = \frac{-B_i}{F_i - B_i}$ است. در روش فرم بسته، حل مسأله از طریق کمینه‌سازی تابع زیر انجام می‌شود:

$$J(a, b) = \sum_{k \in I} (\sum_{i \in w_k} (\alpha_i - aI_i - b)^2 + \varepsilon a^2) \quad (3)$$

که w_k پنجره کوچکی اطراف پیکسل k می‌باشد. در مرجع اخیر و همچنین در [9] تابع هدف، نرم دوی خطای تقریب است. در مرجع [5] با هدف حفظ بهتر لبه‌های تصویر از نرم یک استفاده شده است. در مرجع [3] از روش نزدیک‌ترین همسایگی استفاده شده و یک ویژگی آن توانایی قطعه‌بندی تصویر به بیشتر از دو قطعه است. پژوهش‌های متعددی در حوزه درهم‌تندیدگی تصویر انجام شده است؛ تا آنجا که نگارنده مطلع است، در کارهای انجام‌شده، نمونه‌های آموزشی متعلق به دو طبقه زمینه و پیش‌زمینه وزن یکسانی در فرآیند حل مسأله دارند؛ اما در همه مسائل ممکن است، تخطی از نمونه‌های زمینه و پیش‌زمینه اهمیت یکسانی نداشته باشند.

آنچه در این نوشتار مدنظر است، بهبود روش مرجع [7] است. در این روش نیز وزن یکسانی برای نمونه‌های آموزشی در نظر گرفته شده است. ایده اصلی این مقاله اصلاح روش [7] به نحوی است که بتوان برای هر یک از دو طبقه ضرایب متفاوتی را لحاظ کرد. روش درهم‌تندیدگی تصویر مبتنی بر یادگیری از روش‌های پایه‌ای این حوزه است که شیوه‌های نوین مختلفی بر اساس این روش پیشنهاد شده‌اند [3, 10, 11]. با توجه به اینکه روش پیشنهادی، مبتنی بر بهبود این روش پایه‌ای است، می‌توان از آن در سایر پژوهش‌های برگرفته از این روش نیز استفاده کرد.

در بخش بعد روش [7] به‌صورت جزئی‌تر بیان شده و سپس روش پیشنهادی بر پایه آن ذکر خواهد شد. در بخش

^۱ Closed Form

α_j باشد (اگر $j \in \Omega_l$)، می‌توان معادله (۵) را به صورت زیر فرمول‌بندی کرد:

$$\underset{\alpha \in \mathbb{R}^n}{\operatorname{argmin}} \alpha^T (I_n - F)(I_n - F)^T \alpha + (\alpha - \alpha^*)^T C (\alpha - \alpha^*) \quad (۶)$$

که I ماتریس همانی است. با مشتق‌گیری نسبت به α و برابر صفر قراردادن آن جواب بهینه از رابطه زیر حاصل خواهد شد:

$$\alpha = ((I_n - F)(I_n - F)^T + C)^{-1} C \alpha^*$$

۳- روش پیشنهادی

در این بخش ایده اصلی این نوشتار که چگونگی لحاظ کردن ضرایب متفاوت وزنی برای پیکسل‌های آموزشی روش [7] است، بیان خواهد شد. در کاربردهای درهم‌تنیدگی تصویر، نقشه سه‌گانه‌ای برای مشخص کردن نقاط آموزشی مورد نیاز است. در مثال‌های معمول این حوزه، این نقشه به صورت دستی تولید می‌شود. در کاربردهای مورد اشاره در بخش بعد، بهبود نتایج روش‌های قبلی استخراج شیئی از تصویر مدنظر است. در ادامه این بخش شیوه‌ای خودکار برای ایجاد نقشه سه‌گانه مبتنی بر یک روش قطعه‌بندی موجود، یاد خواهد شد که در بخش کاربردها مورد استفاده قرار می‌گیرد.

۳-۱- بهبود روش درهم‌تنیدگی تصویر مبتنی بر یادگیری برای دقیق‌تر کردن نتیجه

قطعه‌بندی

در بخش قبل اشاره شد که ضریب تخطی مقادیر برآورد شده آلفا برای پیکسل‌های با طبقه نامعلوم در روش [7] یکسان در نظر گرفته شده است. ضریب c در رابطه (۵) که در قالب ماتریس C در رابطه (۶) ظاهر شده بود، وزن مربوط به تخطی آلفای تخمین زده شده نقاط آموزشی را از مقادیر واقعی آنها نشان می‌دهد. در این فرمول‌بندی این وزن برای همه نقاط آموزشی - خواه متعلق به پس‌زمینه باشند و خواه متعلق به پیش‌زمینه - یکسان است.

در حوزه کاری «درهم‌تنیدگی تصویر» به‌طور عمومی نقاط آموزشی یقینی هستند و در قالب نقشه سه‌گانه به صورت دقیق مشخص می‌شود که هر نقطه آموزشی متعلق به کدام دسته است؛ اما در صورت تولید خودکار نقشه سه‌گانه - به‌خصوص در مرز نواحی سه‌گانه - اطمینان صددرصدی از صحت برچسب‌گذاری وجود نخواهد داشت. پیشنهادی که در این خصوص ارائه می‌شود، گرفتن وزن متفاوت تخطی از نقاط آموزشی، برای نقاط زمینه و پیش‌زمینه است.

اگر اسکالره‌های c_b و c_f به ترتیب ضریب تخطی از زمینه و پیش‌زمینه را مشخص کنند، ماتریس‌های C_b و C_f (ماتریس حاوی ضرایب تخطی از نقاط آموزشی زمینه و پیش‌زمینه) را بر اساس آنها و به صورتی مشابه با C ایجاد می‌کنیم. به این ترتیب رابطه (۶) به صورت زیر در خواهد آمد:

$$\underset{\alpha \in \mathbb{R}^n}{\operatorname{argmin}} [\alpha^T (I_n - F)(I_n - F)^T \alpha + (\alpha - \alpha^*)^T C_b (\alpha - \alpha^*) + (\alpha - \alpha^*)^T C_f (\alpha - \alpha^*)] \quad (۷)$$

با ترکیب مناسب ماتریس‌های C_b و C_f و ایجاد یک ماتریس واحد، رابطه (۷) را می‌توان به صورت ساده‌تر بازنویسی کرد. ترکیب C_b و C_f به این صورت است که برای مقادیر روی قطر اصلی C متناظر با پیکسل‌های آموزشی متعلق به زمینه در α^* از C_b و برای مقادیر متناظر با پیکسل‌های پیش‌زمینه از C_f استفاده می‌شود. سایر عناصر هم صفر هستند. با محاسبه C بر اساس ماتریس‌های C_b و C_f ، و با همان فرض معادله (۴) و با مشتق‌گیری از (۷) خواهیم داشت:

$$\alpha = ((I_n - F)(I_n - F)^T + C)^{-1} C \alpha^* \quad (۸)$$

اگر ضرایب ترکیب خطی در معادله (۴) معلوم باشند، مقدار α برآورد می‌شود. برای تعیین ضرایب ترکیب خطی در معادله (۴) به صورت زیر عمل می‌شود: برای بردار معلوم x $x' = [x^T 1]$ در نظر گرفته شده، سپس یک مدل خطی آلفارنگ به صورت زیر انتخاب می‌شود:

$$\alpha = x^T \beta + \beta_0 = x'^T \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix} \quad (۹)$$

که $\beta = [\beta_1 \dots \beta_d]$ و β_0 ضرایب مدل هستند. با فرض $X_i = [x'_{T1} \dots x'_{Tm}]$ که یک ماتریس $m \times (d+1)$ است، می‌توان با روش رگرسیون رنج، برآوردی از β و β_0 را توسط حل مسأله بهینه‌سازی درجه دوم زیر به دست آورد [7]:

$$\underset{\beta, \beta_0}{\operatorname{argmin}} \left\| \alpha_i - X_i \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix} \right\|^2 + \lambda_r \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix}^T \begin{bmatrix} \beta \\ \beta_0 \end{bmatrix}$$

جواب بهینه معادله بالا به صورت زیر حاصل می‌شود:

$$\begin{bmatrix} \hat{\beta} \\ \hat{\beta}_0 \end{bmatrix} = (X_r^T X_i + \lambda_r I)^{-1} X_r^T \alpha \quad (۱۰)$$

با جای‌گذاری معادله (۱۰) در معادله (۹)، f_i به صورت $f_i = (X_i X_i^T + \lambda_r I_{(m)})^{-1} X_i X_i^T \alpha$ حاصل می‌شود که مستقل از $\{\alpha_j\}_{j \in \Omega_i}$ بوده و فقط به $\{x_j\}_{j \in \Omega_i}$ وابسته است. حال کافیست f_i حاصل، در معادله (۴) جایگزین شود تا ضرایب خطی مورد نیاز معادله (۶) که فرم ساده شده معادله (۷) بود، به دست آیند. آلفای به دست آمده از حل (۷) برچسب قطعه‌بندی پیکسل‌های تصویر است.

این شیوه آن است که با تغییر اندازه و شکل عنصر ساختاری^۴ می‌توان اندازه نواحی سفید، خاکستری و سیاه در نقشه سه‌گانه را به آسانی تغییر داد.

۴- کاربردها

بهبود پیشنهادی در این مقاله بر روی روش [7]، مناسب حالاتی است که میزان اطمینان از نقشه سه‌گانه در مناطق مربوط به زمینه و پیش‌زمینه متفاوت باشد. مقایسه‌ها و تصاویر معمول مورد استفاده در حوزه درهم‌تنیدگی تصویر فاقد چنین فرضی هستند؛ به همین دلیل برای نمایش برتری شیوه پیشنهادی، نیازمند کاربردهای خاصی هستیم که فرض این نوشتار برقرار باشد. در این بخش با یاد دو کاربرد مناسب با ایده این مقاله، توانایی شیوه اصلاحی پیشنهادشده در عمل نشان داده می‌شود. کاربرد نخست، دقیق‌تر کردن استخراج متن از تصویر (شکل ۱) و کاربرد دوم دقیق‌تر کردن شناسایی رگ‌های خونی شبکیه چشم است.

۴-۱- دقیق‌تر کردن نتایج استخراج متن از

تصویر

در نخستین کاربرد، مسأله دقیق‌تر کردن استخراج متن از تصویر با استفاده از درهم‌تنیدگی تصویر مورد بررسی قرار خواهد گرفت. استخراج نواحی متعلق به متن از زمینه، مرحله پیش‌پردازشی در همه سامانه‌های شناسایی بصری حروف است [12]. تا آنجا که نگارنده مطلع است، تاکنون شیوه درهم‌تنیدگی تصویر برای دقیق‌تر کردن خروجی یک سامانه جداسازی متن از زمینه به کار نرفته است. برای داشتن مجموعه داده‌ای که خروجی آن قابل ارزیابی باشد، قطعه شعری از حافظ بر روی زمینه‌ای دارای بافت ناهموار درج شده و در آزمایش‌ها به کار برده شده است (شکل ۱). رنگ متن و میزان شفافیت زمینه و متن متفاوت در نظر گرفته شده بوده است.

به‌عنوان یک روش قطعه‌بندی تصاویر دارای متن، در این بخش از روش مرسوم آستانه‌گذاری و مبتنی بر شیوه آستانه‌یابی اوتسو [13] استفاده خواهد شد. شکل (۱-ج) خروجی قطعه‌بندی تصویر ۱۱ (الف) با روش آستانه‌گذاری اوتسو را نشان می‌دهد (یک عمل ساده حذف پیکسل‌های جدا افتاده نیز بعد از قطعه‌بندی انجام شده است). اگر این سطح آستانه، θ نامیده شود، دو تصویر بالا و پایین آن، خروجی حاصل از قطعه‌بندی با سطوح آستانه $\theta - 0.1$ و $\theta + 0.1$ را نشان می‌دهند.

با متمایز کردن C_b و C_f می‌توان ضریب تخطی جداگانه‌ای برای نقاط آموزشی دو طبقه در نظر گرفت و انتظار نتایجی متفاوت داشت. نتایج آزمایش‌های انجام‌شده این حدس را تأیید کرده است.

در ادامه از این شیوه پیشنهادی - که از روش درهم‌تنیدگی تصویر مبتنی بر یادگیری استفاده می‌شود - با عنوان «روش درهم‌تنیدگی تصویر مبتنی بر یادگیری وزن‌دهی‌شده» و یا به اختصار 'WLbM' یاد خواهد شد.

۲-۳- تولید خودکار نقشه سه‌گانه

نتایج روش‌های درهم‌تنیدگی تصویر نشان می‌دهند که با فرض داشتن نقشه سه‌گانه مناسب، این شیوه‌ها قابلیت خوبی در جداسازی مرز دقیق شیئی و زمینه در تصاویر طبیعی دارند. در پژوهش‌های معمول در حوزه درهم‌تنیدگی تصویر، نقشه سه‌گانه به صورت دستی تولید می‌شود. در این بخش یک روش خودکار برای تولید نقشه سه‌گانه بیان می‌شود که در بخش کاربرد شیوه پیشنهادی بخش قبل مورد استفاده قرار خواهد گرفت.

کاربردهای مدنظر این نوشتار، دقیق‌تر کردن خروجی قطعه‌بندی روش‌های پیشین است. با فرض داشتن روشی پیشین در قطعه‌بندی، برای تولید نقشه سه‌گانه، از عملیات مورفولوژی روی نواحی استخراج‌شده، استفاده می‌شود. یک بار عمل سایش (خوردگی)^۲ و یک بار عمل گسترش (اتساع)^۳ روی نواحی حاصل از قطعه‌بندی استخراج‌شده، اعمال شده و نتایج با هم ترکیب می‌شوند. اگر نتیجه این عملیات را به ترتیب تصاویر E و D بنامیم، در تصویر E ناحیه قطعه‌بندی‌شده کوچک‌تر و در تصویر D ناحیه قطعه‌بندی‌شده بزرگ‌تر خواهد بود. برای تولید نقشه سه‌گانه، کاری که انجام شده آن است که تصویر حاصل از عملیات خوردگی (تصویر E) به‌عنوان نقاطی که به حتم متعلق به پیش‌زمینه هستند، در نظر گرفته می‌شود؛ اختلاف تصویر D و تصویر E به‌عنوان مناطق با برچسب نامشخص (خاکستری) و سایر نقاط به‌عنوان زمینه (سیاه) در نظر گرفته می‌شوند. نمایش مراحل کار در بخش بعد، نحوه کار را بهتر نشان خواهد داد.

مرز نواحی قطعه‌بندی‌شده، نقاطی هستند که اطمینان صدها درصد نسبت به تعلق آنها به دو طبقه وجود ندارد؛ تعیین تعلق آنها به دو طبقه با روش‌های درهم‌تنیدگی انجام خواهد شد. به جز فرسایش و گسترش می‌توان از روش‌های دیگر مرتبط با استخراج دور ناحیه هم استفاده کرد؛ اما برتری

^۱ WLbM: Weighted Learning based Matting

^۲ Erosion

^۳ Dilation

^۴ Structure Element

همان‌گونه که از شکل (۱) مشخص است، در تصویر خروجی حاصل از سطح آستانه پایین‌تر از بهینه، متن، حالت خورده شده دارد و در تصویر حاصل از سطح آستانه بالاتر از بهینه، نقاطی از زمینه نیز به اشتباه به‌عنوان متن بخش‌بندی شده‌اند. بهترین حالت، همان سطر میانی است که حاصل از سطح آستانه بهینه است؛ با این حال این خروجی نیز خروجی مناسبی نیست. کیفیت نامناسب خروجی در لبه‌های حروف مشهود است که در ادامه با استفاده از روش‌های درهم‌تنیدگی سعی در بهبود آن خواهیم داشت.

بهبود لبه‌ها حتی با افزایش اندازه تصویر با روش bicubic با هزینه ضد شکستگی^۱ که لبه‌ها را نرم‌تر می‌کند و افزایش شدت نور تصویر شکل (۱-ه) نیز موفقیت‌آمیز نیست (برای مقایسه نتایج، از نسخه الکترونیکی این نوشتار و در صورت نیاز، بزرگ‌نمایی استفاده شود). در شکل‌های بعد خروجی حاصل از دیگر روش‌ها با این خروجی مقایسه خواهد شد.



(الف) تصویر اصلی

عاشق که شد که یار به حالش نظر نکرد

(ب) قطعه‌بندی با سطح آستانه کمتر از سطح آستانه بهینه

عاشق که شد که یار به حالش نظر نکرد

(ج) با سطح آستانه بهینه با روش اوتسو [۱۳]

عاشق که شد که یار به حالش نظر نکرد

(د) با سطح آستانه بیشتر از سطح بهینه

عاشق که شد که یار به حالش نظر نکرد

(ه) افزایش اندازه با روش bicubic و افزایش شدت نور

(شکل ۱-۱): (الف) تصویر اصلی. (ب) تا (ه) خروجی قطعه‌بندی

تصویر (الف) با روش آستانه‌گذاری.

(Figure-1): Image segmentation of (a) with thresholding method.

۴-۱-۱- تولید نقشه سه‌گانه

برای تولید نقشه سه‌گانه از روش یادشده در بخش ۲-۳

استفاده شده است. خروجی روش قطعه‌بندی حاصل از آستانه‌گذاری بخش پیش - که لبه‌های بخش‌های آن مورد اطمینان نیستند - برای بهبود بیشتر ملاک عمل خواهد بود. با عملیات سایش و گسترش این قطعه‌بندی، نقشه سه‌گانه به‌صورت نمایش داده‌شده در شکل (۲-ه) تولید می‌شود.

۴-۱-۲- آزمایش‌های انجام‌شده

شکل‌های (۲ تا ۸) نتایج دیداری و کمی آزمایش‌های مربوط به بهبود استخراج متن از تصویر با روش‌های مختلف را نشان می‌دهند که در ادامه توضیح داده خواهند شد. شکل‌های (۲ و ۳) دو مصرع از یک بیت شعر حافظ را نشان می‌دهند که بر روی زمینه‌های مختلف ترکیب شده‌اند. زیرشکل‌های (الف) و (ب) در هر دو شکل نمایان‌گر لایه زمینه و لایه پیش‌زمینه در فوتوشاپ هستند. زیرشکل (ج) نتیجه ترکیب دولایه را نشان می‌دهد که تصویری است که برای استخراج متن مورد پردازش قرار خواهد گرفت. در ترکیب دو لایه میزان شفافیت و رنگ هم مقداری تغییر داده شده است.

زیرشکل (د) از شکل‌های (۲ و ۳) خروجی قطعه‌بندی با روش آستانه‌گذاری اوتسو و زیرشکل (ه) نقشه سه‌گانه را نشان می‌دهد. تفاوت نقشه سه‌گانه شکل (۲-ه) با نقشه سه‌گانه شکل (۳-ه) آن است که در اولی، عنصر ساختاری عمل گسترش بزرگ‌تر بوده و لذا ناحیه خاکستری بزرگ‌تری تولید کرده است. زیرشکل (و) خروجی روش درهم‌تنیدگی تصویر [7] را نشان می‌دهد که زیرشکل (ج) به‌عنوان تصویر ورودی و زیرشکل (ه) به‌عنوان نقشه سه‌گانه ورودی‌های آن بوده‌اند. مقدار c در رابطه (۵) برای این آزمایش، هشتصد در نظر گرفته شده که همان مقدار مورد استفاده در آزمایش‌های نویسندگان [7] است.

زیرشکل‌های (ز) و (ح) خروجی روش پیشنهادی در بخش ۱-۳ است که در اولی جریمه مربوط به تخطی از نقاط آموزشی زمینه و در دومی جریمه مربوط به تخطی از نقاط آموزشی پیش‌زمینه بیشتر بوده است. از همان مقدار هشتصد مورد استفاده نویسندگان [7] برای ضریب تخطی زیاد و از مقدار یک برای ضریب تخطی کم استفاده شده است.

بررسی مقدار مناسب ضرایب تخطی در بخش بعد انجام خواهد شد. همان‌گونه که مشاهده می‌شود، حالت آخر در شکل (۲)، خروجی‌های بهتری نسبت به سایر روش‌ها تولید کرده است.

یک دلیل برتری شیوه درهم‌تنیدگی تصویر در این کاربرد آن است که برای نمایش بهتر قلم‌ها در نمایش‌گر و جلوگیری از نمایش دندان‌دانه حروف، از ضد شکستگی

^۱ Anti-Aliasing

گرچه حالت نمایش داده‌شده در زیر شکل (ز) (شکل‌های ۲ و ۳) خروجی مناسبی از منظر قطعه‌بندی نیست و به نظر می‌رسد که خروجی بدتر شده و متن حالتی شبیه به پخش شدن جوهر پیدا کرده است؛ اما این حالت می‌تواند به عنوان یک جلوه ویژه در نرم‌افزارهای پردازش تصویر بکار گرفته شود.



(الف) تصویر زمینه

ای خوابه درد نیست وگرنه طیب هست

(ب) تصویر پیش‌زمینه



(ج) تصویر مورد پردازش

ای خوابه درد نیست وگرنه طیب هست

(د) خروجی روش آستانه‌گذاری

ای خوابه درد نیست وگرنه طیب هست

(هـ) نقشه سه‌گانه با گسترش کمتر ناحیه مجهول

ای خوابه درد نیست وگرنه طیب هست

(و) خروجی روش [7] با $c = 800$

ای خوابه درد نیست وگرنه طیب هست

(ز) خروجی روش پیشنهادی با $c_b = 1$ و $c_f = 800$

ای خوابه درد نیست وگرنه طیب هست

(ح) خروجی روش پیشنهادی با $c_b = 800$ و $c_f = 1$

(شکل-۳): مراحل و نتایج کار بر روی تصویر مصرع دوم

(Figure-3): Intermediate and the final result of the experiments on 2nd hemistich.

استفاده می‌شود که در نتیجه لبه‌ها هموارتر نمایش داده می‌شوند. نمایش هموار لبه‌ها با تصویر دودویی حاصل از قطعه‌بندی معمول امکان‌پذیر نیست؛ اما روش‌های درهم‌نیدگی تصویر که میزان مشارکت زمینه و پیش‌زمینه را به صورت آلفا (در بازه صفر تا یک) می‌دهند، برای این منظور مناسب هستند؛ لذا خروجی این شیوه‌ها در کاربرد یادشده، چشم‌نوازتر است.



(الف) تصویر زمینه

عاشق که شد که یار به حالش نظر نکرد

(ب) تصویر پیش‌زمینه



(ج) تصویر مورد پردازش

عاشق که شد که یار به حالش نظر نکرد

(د) خروجی روش آستانه‌گذاری

عاشق که شد که یار به حالش نظر نکرد

(هـ) نقشه سه‌گانه

عاشق که شد که یار به حالش نظر نکرد

(و) خروجی روش [7] با $c = 800$

عاشق که شد که یار به حالش نظر نکرد

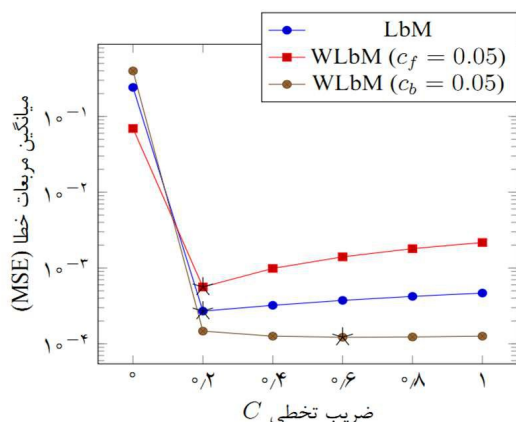
(ز) خروجی روش پیشنهادی با $c_b = 1$ و $c = 800$

عاشق که شد که یار به حالش نظر نکرد

(ح) خروجی روش پیشنهادی با $c_b = 800$ و $c_f = 1$

(شکل-۲): مراحل و نتایج کار بر روی تصویر مصرع اول

(Figure-2): Intermediate and the final result of the experiments on 1st hemistich.



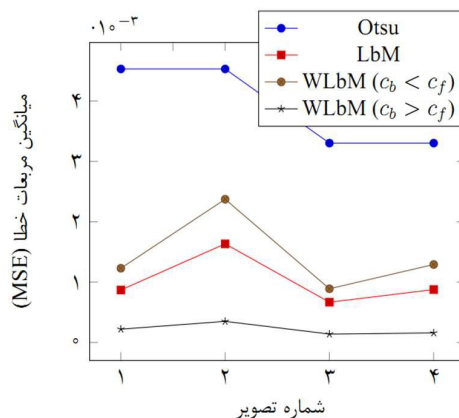
(شکل-۵): مقایسه میانگین مربعات خطا بین خروجی روش‌های مختلف و تصویر اصلی متن مورد استخراج در مقادیر مختلف مقادیر تخطی، زمینه و پیش‌زمینه. مقدار کمیته خطای هر روش با ستاره مشخص شده است. بهترین نتیجه متعلق به روش پیشنهادی در حالتی است که $(c_b = 0.05, c_f = 0.6)$. (توضیح بیشتر در متن، بخش ۴-۱-۳).

(Figure-5): Comparison of the MSE results on text extraction example. The best results is obtained by the proposed approach (when $c_b = 0.05, c_f = 0.6$).

همان‌گونه که از نتایج آزمایش‌ها مشخص است، تغییر ضرایب می‌تواند خروجی‌های متفاوتی را سبب شود. دقیق‌تر شدن کیفیت خروجی متن یا حالت نمایشی از پخش شدگی جوهر روی متن می‌تواند با تغییر ضرایب حاصل شود.

۳-۱-۴- مقادیر بهینه ضریب تخطی

آزمایش‌های بخش قبل مشخص کردند که روش پیشنهادی با تغییر ضرایب تخطی می‌تواند نتایج دیداری و کمی بهتری نسبت به روش پایه‌ای [7] به‌دست دهد. در آزمایش‌های بخش قبل سعی شد از مقدار مورد استفاده توسط نویسندگان [7] استفاده شود. شکل‌های (۵ تا ۸) نتایج آزمایش‌های این بخش را برای پیدا کردن مقادیر بهینه ضرایب تخطی در روش پیشنهادی نشان می‌دهند که در ادامه توضیح داده می‌شود. از آنجا که در روش پیشنهادی دو ضریب c_b و c_f را داریم که ضریب میزان تخطی مقدار آلفای خروجی روش روی پیکسل‌های آموزشی زمینه و پیش‌زمینه را مشخص می‌کنند، بازه مقادیر هر کدام بازه $[0, 1]$ در نظر گرفته می‌شود. برای پیدا کردن مقادیر بهینه دو ضریب، هنگامی که تعداد مقادیر ممکن زیاد نباشند، می‌توان از جستجوی سراسری استفاده کرد که زمان‌بر است.



(شکل-۴): مقایسه میانگین مربعات خطا بین خروجی روش‌های مختلف و تصویر اصلی متن مورد استخراج. بهترین نتیجه متعلق به روش پیشنهادی در حالتی است که $c_b > c_f$. (توضیح بیشتر در متن).

(Figure-4): Comparison of the MSE results on text extraction example. The best results is obtained by the proposed approach (when $c_b > c_f$).

با در نظر گرفتن شکل (ب) به‌عنوان خروجی قطعه‌بندی درست، خطای چهار خروجی زیر شکل‌های (د)، (و)، (ز) و (ح) قابل محاسبه است. شکل (۴) میانگین مربعات خطا (رابطه ۱۱) بین خروجی دو شیوه نمایش داده‌شده در شکل‌های ۲ و ۳ و جواب درست را نشان می‌دهد. محور x نمایانگر تصویر مورد نظر است. تصاویر ۲ (ج) و ۳ (ح) شماره‌های ۲ و ۳ هستند. شماره‌های ۱ و ۴ همان تصاویر بالا هستند که نقشه سه‌گانه با میزان «گسترش» کمتر/بیشتر از دو تصویر دیگر تولید شده است. میانگین مربعات خطا از رابطه زیر حاصل می‌شود که در آن Y_i مقدار جواب درست و \hat{Y}_i خروجی روش مورد مقایسه برای پیکسل i ام است:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (11)$$

در کادر راهنمای شکل (۴)، Otsu مشخص‌کننده روش اوتسو، LbM (Learning based Matting) مشخص‌کننده روش [7] و WLbM مشخص‌کننده شیوه پیشنهادی است که دو حالت دارد: حالتی که جریمه میزان تخطی نقاط آموزشی زمینه از پیش‌زمینه بیشتر است $(c_b > c_f)$ و برعکس $(c_b < c_f)$. همان‌گونه که مشخص است، حالت دوم روش پیشنهادی، خطای کمتری نسبت به سایر روش‌ها دارد که منطبق بر نتایج نشان داده‌شده به‌صورت بصری در شکل‌های (۲ و ۳) است.

ماشک‌شد که یار به حالش نظر نکرد

عاشق که شد که یار به حالش نظر نکرد
عاشق که شد که یار به حالش نظر نکرد
عاشق که شد که یار به حالش نظر نکرد

(شکل-۸): از بالا به پایین: خروجی (مناسب) روش پیشنهادی با مقدارهای ضریب تخطی $c_f = 0.05$ و c_b به ترتیب برابر با ۰، ۰/۲، ۰/۴ و ۰/۸؛ ۱؛ متناظر با روش WLbM ($c_b = 0.05$) از شکل (۵).
(Figure-8): From top to bottom: The results of the proposed method with $c_f = 0.05$ and c_b equals to 0, 0.2, 0.6 and 1, respectively.

برای روش پیشنهادی، مقدار یک پارامتر ۰/۰۵ و برای پارامتر بعدی مقادیر ۰، ۰/۲، ۰/۴، ۰/۶، ۰/۸ و ۱ در نظر گرفته شد. بر طبق نتایج حاصل‌شده، کمترین خطا مربوط به حالت $(c_b, c_f) = (0.6, 0.05)$ است. استفاده از سایر روش‌های بهینه‌سازی برای پیدا کردن مقادیر بهینه این پارامترها از جمله کارهای آینده این پژوهش است. روش ساده یادشده، نشان داد که با انتخاب مقادیر مناسب برای ضرایب تخطی زمینه و پیش‌زمینه، روش پیشنهادی می‌تواند خطای کمتری نسبت به روش اصلی به‌دست دهد.

تصاویر خروجی متناظر با مقادیر تخطی برابر با ۰، ۰/۲، ۰/۴ و ۰/۸ و ۱ از شکل (۵) در شکل‌های (۶ تا ۸) آمده است که به‌ترتیب مربوط به حالت‌های LbM، WLbM ($c_f = 0.05$) و WLbM ($c_b = 0.05$) از شکل (۵) هستند. کمینه خطای مربوط به هر حالت نمودار با ستاره مشخص شده است. هر سه مورد در وضعیت مقدار تخطی برابر با صفر، بیشترین خطا را داشته‌اند. در دو حالت نخست، مقدار خطا با افزایش مقدار ضریب تخطی افزایش پیدا کرده است (به جز در مقدار صفر). اما در حالت سوم - WLbM ($c_b = 0.05$) - از صفر تا ۰/۶ مقدار خطا سیر نزولی دارد و پس از آن با افزایش مواجه می‌شود. نتایج دیداری نمایش داده‌شده در شکل‌های (۶ تا ۸) مؤید نتایج کمی است. در حالت مقدار تخطی صفر، هر سه روش ضعیف‌ترین خروجی‌ها را دارند. سطر سوم نتایج روش پیشنهادی در شکل (۸) که متناظر با حالت $(c_b, c_f) =$

عاشق که شد که یار به حالش نظر نکرد
عاشق که شد که یار به حالش نظر نکرد
عاشق که شد که یار به حالش نظر نکرد

(شکل-۶): از بالا به پایین: خروجی روش [7] (LbM) با مقدارهای ضریب تخطی برابر با ۰، ۰/۲، ۰/۴ و ۱؛ متناظر با روش LbM از شکل (۵).

(Figure-6): From top to bottom: The results of the LbM method [7] with C equals to 0, 0.2, 0.6 and 1.

با ثابت در نظر گرفتن یک ضریب و کمتر کردن تعداد پارامترها، هم زمان مورد نیاز، کمتر خواهد شد و هم نتایج به‌راحتی با روش اصلی - که یک ضریب تخطی دارد - قابل مقایسه خواهد بود. در آزمایش‌های نشان داده‌شده در شکل (۵) برای روش [7] مقادیر ۰، ۰/۲، ۰/۴، ۰/۶، ۰/۸ و ۱ در نظر گرفته شد.

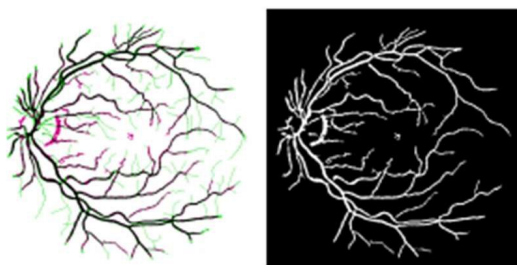
عاشق که شد که یار به حالش نظر نکرد
عاشق که شد که یار به حالش نظر نکرد
عاشق که شد که یار به حالش نظر نکرد

(شکل-۷): از بالا به پایین: خروجی (نامناسب) روش پیشنهادی با مقدارهای ضریب تخطی $c_b = 0.05$ و c_f به ترتیب برابر با ۰، ۰/۲، ۰/۴ و ۰/۸؛ متناظر با روش WLbM ($c_f = 0.05$) از شکل (۵).
(Figure-7): From top to bottom: The results of the proposed method with $c_b = 0.05$ and c_f equals to 0, 0.2, 0.6 and 1, respectively.

که بخش داخلی تر رگ به درستی شناسایی شده است. دسته‌ای از اشتباهات شناسایی بیش تر در کناره‌های رگ‌ها اتفاق افتاده است که در تصویر (۱۰-ب) با رنگ بنفش نشان داده شده‌اند.



(الف) تصویر اول از مجموعه DRIVE (ب) تصویر مرجع (شکل-۹): یک تصویر شبکیه و تصویر مرجع (رگ‌های استخراج شده به صورت دستی) آن به عنوان جواب درست (Ground Truth) از مجموعه دادگان DRIVE.
(Figure-9): (a) First retinal image from DRIVE dataset, (b) the hand labelled "truth" image.



(الف) خروجی روش [17] (ب) خطای روش [17] (شکل-۱۰): یک روش استخراج رگ و خطای شناسایی آن. (الف) خروجی طبقه‌بندی با روش استخراج موجک [17] بر روی شکل ۹ (الف). (ب) نمایش خطای روش، نواحی سیاه و سفید رنگ، پیکسل‌هایی هستند که به درستی طبقه‌بندی شده‌اند (نقاط رگ و غیررگ). پیکسل‌های بنفش رنگ، نواحی غیررگی هستند که به اشتباه به عنوان رگ برچسب خورده‌اند؛ و نواحی سبز رنگ، رگ‌هایی هستند که به اشتباه تشخیص داده نشده‌اند.
(Figure-10): (a) Vessel segmented image of [17]. In (b) the black and white pixels are those pixels that are classified correctly; Purple pixels show misclassified pixels labeled as 'vessel' and Green pixels demonstrate vessel regions that labeled as 'non vessel'.

نحوه کار، مشابه روش مورد استفاده در بخش استخراج متن از تصویر است. ابتدا با یکی از روش‌های استخراج رگ، عمل قطع‌بندی انجام شده، نقشه سه‌گانه بر اساس این قطع‌بندی تولید و با روش درهم‌تنیدگی تصویر، روی نواحی لبه‌های رگ‌ها، آشکارسازی دقیق‌تری انجام می‌شود.

(0.6.0.05) است، دارای کمترین خطاست که مقایسه بصری این شکل با شکل‌های (۶ و ۷) همین موضوع را نشان می‌دهد؛ لذا شیوه پیشنهادی در این کاربرد می‌تواند نتایج بهتری نسبت به شیوه اصلی تولید کند.

۲-۴- دقیق‌تر کردن نتایج استخراج رگ‌های

چشم

به عنوان کاربرد دوم، مسأله دقیق‌تر کردن استخراج رگ‌های خونی چشم را مورد بررسی قرار خواهیم داد. استخراج درست بافت‌های خاص در تصاویر پزشکی همچون عروق خونی شبکه چشم و شریان‌های کرونری قلب در تصاویر آنژیوگرافی، کمک مؤثری در تشخیص درست پزشک بوده و پژوهش‌های متعددی را به خود معطوف کرده‌اند [14-16].

روش‌های مرسوم که در حوزه استخراج رگ‌های خونی شبکه چشم وجود دارند [17-21]، به طور عمومی یا جزو روش‌های بدون ناظر هستند که از طریق خوشه‌بندی و آستانه‌یابی، اقدام به جداسازی ناحیه رگ می‌کنند و یا جزو روش‌های با ناظر هستند که از طبقه‌بندی‌هایی همچون ماشین بردار پشتیبان [18] برای قطع‌بندی استفاده می‌کنند. استفاده از روش‌های معمول پردازش تصویر همچون فیلترینگ، لبه‌یابی، متعادل‌سازی هیستوگرام، عملیات مورفولوژی و تبدیل موجک [17] هم در این حوزه مرسوم است. با توجه به اینکه هدف این مقاله، پیشنهاد یک روش استخراج رگ نیست، به مرور مفصل روش‌های استخراج رگ‌های خونی چشم پرداخته نخواهد شد و تنها یک روش استخراج رگ به کار گرفته شده و نتایج آن بهبود داده خواهد شد.

در آزمایش‌ها از مجموعه دادگان DRIVE^۱ استفاده شده است. تصویر شکل (۹-الف) نخستین تصویر از مجموعه دادگان یادشده و تصویر (۹-ب) جواب دانسته (برچسب‌گذاری شده) آن را نشان می‌دهد. شکل (۱۰-الف) که خروجی روش [17] را نشان می‌دهد، ملاحظه فرمایید؛ در مقایسه خروجی این شیوه با تصویر مرجع متناظر با آن شکل (۹-ب) مشاهده می‌شود که بیشتر خطای روش یادشده در عدم استخراج رگ‌های ریز، در کناره‌های رگ‌های استخراج‌شده و یا ناحیه صفحه نوری (Optic Disk) بوده است. شکل (۱۰-ب) نتیجه مقایسه را نشان می‌دهد. نقاط سیاه و سفید، نقاطی هستند که به درستی طبقه‌بندی شده‌اند. نقاط بنفش/سبز نقاطی هستند که به اشتباه رگ تشخیص داده شده‌اند/نشده‌اند. برای رگ‌های استخراج‌شده با اطمینان به نسبت زیادی می‌توان گفت

^۱ <http://www.isi.uu.nl/Research/Databases/DRIVE/>

۴-۲-۱- تولید نقشه سه‌گانه

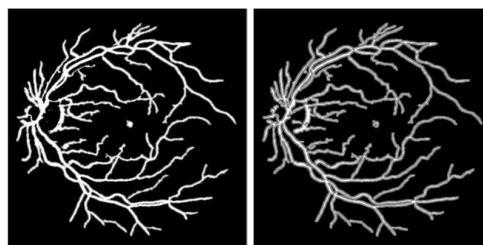
نقشه سه‌گانه مطابق شیوه‌ی یادشده در بخش ۳-۲ و بر اساس روش استخراج رگ [17] تولید می‌شود. همان‌گونه که پیش‌تر اشاره شد، برای تولید نقشه سه‌گانه، از عملیات موفورلوژی فرسایش و گسترش روی نواحی استخراج‌شده استفاده می‌شود. نقشه سه‌گانه تولیدشده برای تصویر (۹-الف) و با خروجی قطعه‌بندی روش [17] شکل (۱۰-الف) در شکل (۱۱-الف) نشان داده شده است.

۴-۲-۲- آزمایش‌های انجام‌شده

برای هر تصویر از مجموعه دادگان، نقشه سه‌گانه ایجادشده، به همراه کانال سبز تصویر ورودی [20] به الگوریتم درهم‌تنیدگی تصویر داده می‌شود.

از آنجا که صفحه‌ی نوری جزو ناحیه رگ‌ها نیست، برای هر تصویر ناحیه مربوط به دیسک نوری با یک روش خوشه‌بندی ساده (روش K-means) - به‌صورت تقریبی - شناسایی و از خروجی حذف شده است. به‌منظور پرهیز از تفاوت خروجی خوشه‌بندی در اجراهای متعدد، این عمل یک‌بار برای هر تصویر انجام شده و ناحیه مربوط به صفحه‌ی نوری برای هر تصویر ذخیره و از خروجی هر روش حذف شده است.

عملیات بالا برای هر تصویر از مجموعه دادگان مورد اشاره انجام شده است. تصویر (۱۰-الف) خروجی روش [17] در استخراج رگ‌های شکل (۹-الف) را نشان می‌دهد. بر اساس این خروجی و مطابق با روش پیشنهادی یادشده در بخش قبل نقشه سه‌گانه به‌صورت خودکار تولید می‌شود که نتیجه آن در شکل (۱۱-الف) نشان داده شده است.



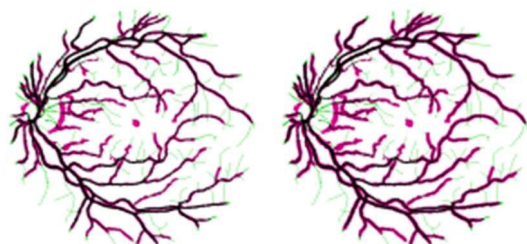
(الف) نقشه سه‌گانه (ب) آلفای به‌دست آمده

(شکل-۱۱): (الف) نقشه سه‌گانه تولیدشده برای شکل (۱۰-الف)، مطابق شیوه‌ی پیشنهادی در بخش ۳-۲. (ب) آلفای به‌دست آمده با روش درهم‌تنیدگی تصویر مبتنی بر یادگیری [7] از شکل (۹-الف) و با نقشه سه‌گانه شکل (۱۱-الف).

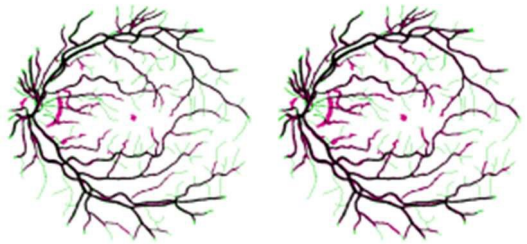
(Figure-11): Trimap and the alpha image produced by the matting method [7].

اگر این نقشه سه‌گانه به‌عنوان α^* در رابطه (۶) به همراه سطح سبز^۱ تصویر ورودی به الگوریتم درهم‌تنیدگی تصویر [7] داده شده و α بهینه محاسبه شود، می‌توان امیدوار بود که تعلق نواحی خاکستری شکل (۱۱-الف) به طبقه رگ یا غیررگ با دقت بیشتری مشخص شود. آلفای محاسبه‌شده که در شکل (۱۱-ب) نشان داده شده است، مقادیر بین صفر و یک را دارد و برای تعیین دو طبقه یک آستانه‌یابی نیاز دارد. برای سطوح آستانه ۰/۰۲، ۰/۰۴ و ... و ۰/۹۸ این کار انجام شده و بهترین سطح آستانه برای هر تصویر مشخص شده است.

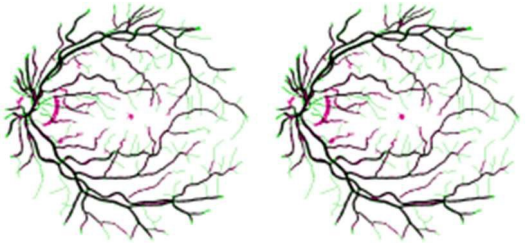
شکل‌های (۱۲ و ۱۳) نتیجه دقیق‌ترکردن روش استخراج رگ حاصل از روش [17] را نشان می‌دهند.



(الف) سطح آستانه: ۰/۱۰ (ب) سطح آستانه: ۰/۳۰



(ج) سطح آستانه: ۰/۵۰ (د) سطح آستانه: ۰/۷۰



(ه) سطح آستانه: ۰/۹۰ (و) سطح آستانه: ۰/۹۸

(شکل-۱۲): خروجی بر اساس سطوح مختلف آستانه روی آلفای به‌دست آمده از روش LbM [7] و با نقشه سه‌گانه (۱۱-الف) (تولیدشده بر اساس [17]).

(Figure-12): Enhancement of retinal vessel segmentation [17] with matting method [7] on different thresholding values of alpha.

مربوط به شبکه‌ی، نسبت به دو کانال رنگی قرمز و آبی، تمایز بیشتری بین رگ و زمینه وجود دارد.

^۱ در بسیاری از پژوهش‌های انجام شده در حوزه شناسایی رگ و ازجمله مرجع [20] اشاره شده است که در کانال سبز تصویر

(جدول-۱): ارزیابی مبتنی بر معیار صحت. ستون دوم: روش [17].

ستون سوم: روش LbM [7]. ستون‌های چهارم و پنجم: روش پیشنهادی WLbM (بخش ۳-۱) با ضرایب مختلف برای زمینه و

پیش‌زمینه

(Table-1): Performance comparisons based on Accuracy. First column: Image Number in DRIVE dataset, second column: [17], third column: LbM [7], the last two columns are the proposed WLbM method with different values for (C_b, C_f) . (Sec. 1-3).

روش مرجع [17]	ضرایب جرمه زمینه و پیش‌زمینه (1, 800)	800	(800, 1)	پیش‌زمینه
۱	0.943	0.943	0.946	0.944
۲	0.949	0.950	0.951	0.949
۳	0.941	0.942	0.942	0.941
۴	0.946	0.946	0.948	0.946
۵	0.952	0.952	0.952	0.953
۶	0.938	0.938	0.938	0.938
۷	0.945	0.946	0.948	0.946
۸	0.929	0.930	0.933	0.930
۹	0.938	0.938	0.938	0.939
۱۰	0.947	0.947	0.950	0.948
۱۱	0.941	0.941	0.946	0.941
۱۲	0.946	0.946	0.949	0.946
۱۳	0.931	0.931	0.930	0.931
۱۴	0.942	0.943	0.948	0.943
۱۵	0.933	0.934	0.933	0.935
۱۶	0.943	0.944	0.946	0.944
۱۷	0.947	0.948	0.951	0.948
۱۸	0.948	0.947	0.949	0.947
۱۹	0.944	0.943	0.944	0.943
۲۰	0.936	0.936	0.937	0.936
میانگین	0.942	0.942	0.944	0.942

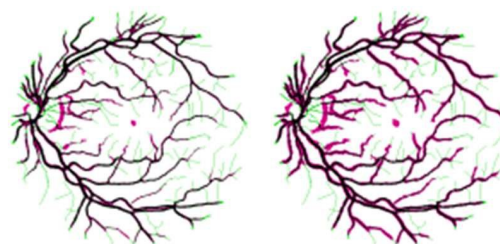
همان‌گونه که انتظار می‌رود، هنگامی که سطح آستانه

کم است، نرخ مثبت کاذب (نواحی بنفش‌رنگ) زیاد است و با افزایش سطح آستانه، نرخ مثبت کاذب کم شده و نرخ منفی کاذب (نواحی سبزرنگ) زیاد می‌شود.

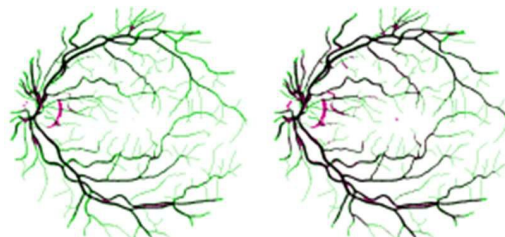
به‌صورت مشابه، شکل (۱۳) نتایج استفاده از روش پیشنهادی (بخش ۳-۱) را به‌عنوان روش درهم‌تنیدگی نشان می‌دهد.

گرچه نحوه عملکرد الگوریتم با این شکل‌ها تا حدودی مشخص می‌شود، ولی برای نتیجه‌گیری دقیق باید به نتایج کمی مقایسه خروجی با جواب واقعی متوسل شد. به این منظور خروجی حاصل از آستانه‌یابی در سطوح مختلف با جواب واقعی با معیارهای «صحت»^۳، «حساسیت»^۴، «نرخ مثبت کاذب» و «نرخ منفی کاذب» مورد سنجش قرار گرفته است (جداول ۱، ۲، ۳ و ۴).

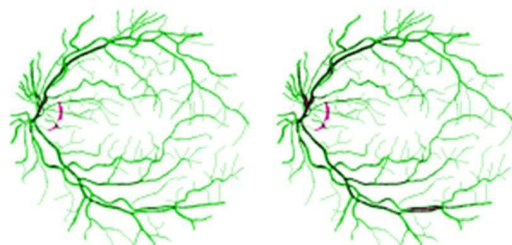
بهترین جواب در هر سطر با زمینه خاکستری رنگ



(الف) سطح آستانه: ۰/۱۰ (ب) سطح آستانه: ۰/۳۰



(ج) سطح آستانه: ۰/۵۰ (د) سطح آستانه: ۰/۷۰



(ه) سطح آستانه: ۰/۹۰ (و) سطح آستانه: ۰/۹۸

(شکل-۱۳): خروجی بر اساس سطوح مختلف آستانه با استفاده از روش پیشنهادی WLbM (بخش ۳-۱) و با نقشه سه‌گانه (۱۱-الف) (تولید شده بر اساس [17]).

(Figure-13): Enhancement of retinal vessel segmentation [17] with the proposed matting method on different thresholding values of alpha.

در این دو شکل، نقشه سه‌گانه مبتنی بر روش [17] ایجاد شده و نتیجه با روش‌های درهم‌تنیدگی تصویر [7] و اصلاح‌شده آن در این نوشتار، بهبود داده شده است. شکل (۱۲) نتیجه روش درهم‌تنیدگی مبتنی بر یادگیری [17] (LbM، بخش ۴-۲) برای سطوح آستانه ۰/۱، ۰/۲، ... و ۰/۹۸ را نشان می‌دهد. نخستین تصویر که نمایش داده شده، مربوط به سطح آستانه ۰/۱ است.

در هر تصویر، نواحی سیاه و سفیدرنگ، پیکسل‌هایی هستند که به‌درستی طبقه‌بندی شده‌اند (نقاط رگ و غیررگ). پیکسل‌های بنفش‌رنگ، نواحی غیررگی هستند که به‌اشتباه به‌عنوان رگ، برچسب خورده‌اند (مثبت کاذب)^۱ و نواحی سبزرنگ، رگ‌هایی هستند که به‌اشتباه تشخیص داده نشده‌اند (منفی کاذب)^۲.

^۳ Accuracy

^۴ Sensitivity, true positive rate, or the recall

^۱ False Positive

^۲ False Negative

(جدول-۳): ارزیابی مبتنی بر معیار نرخ مثبت کاذب.

ستون دوم: روش [17]، ستون سوم: روش LbM [7]، ستون‌های چهارم و پنجم: روش پیشنهادی WLbM (بخش ۳-۱) با ضرایب مختلف برای زمینه و پیش‌زمینه.

(Table-3): Performance comparisons based on False Positive Rate. First column: Image Number in DRIVE dataset, second column: [17], third column: LbM [7] the last two columns are the proposed WLbM method with different values for (C_b, C_f) .

روش مرجع	ضرایب جریمه زمینه و پیش‌زمینه	800	(800, 1)	(1, 800)
[17]				
۱	0.029	0.031	0.023	0.031
۲	0.015	0.018	0.017	0.017
۳	0.023	0.027	0.025	0.024
۴	0.022	0.022	0.016	0.023
۵	0.022	0.023	0.019	0.024
۶	0.021	0.021	0.019	0.021
۷	0.032	0.032	0.021	0.033
۸	0.047	0.048	0.029	0.049
۹	0.027	0.028	0.023	0.029
۱۰	0.031	0.031	0.021	0.032
۱۱	0.028	0.029	0.021	0.029
۱۲	0.028	0.028	0.021	0.029
۱۳	0.021	0.023	0.023	0.022
۱۴	0.036	0.037	0.020	0.037
۱۵	0.039	0.040	0.025	0.041
۱۶	0.027	0.028	0.021	0.029
۱۷	0.028	0.029	0.021	0.030
۱۸	0.018	0.021	0.020	0.020
۱۹	0.019	0.021	0.019	0.020
۲۰	0.022	0.023	0.021	0.024
میانگین	0.027	0.028	0.021	0.028

ستون سوم نتیجه استفاده از روش درهم‌تندگی تصویر [7] روی جواب [17] را نشان می‌دهد. ستون‌های چهارم و پنجم نتیجه اعمال شیوه اصلاح‌شده [7] که در بخش ۳-۱ پیشنهاد شده بود و با مقادیر $(800, 1) = (C_b, C_f)$ و $(1, 800) = (C_b, C_f)$ به‌عنوان ضرایب زمینه و پیش‌زمینه را نشان می‌دهند.

در مورد معیارهای صحت و حساسیت، مقادیر بیشتر و در دو معیار نرخ مثبت کاذب و نرخ منفی کاذب مقادیر کمتر، به‌تربودن نتیجه را مشخص می‌کنند. همان‌گونه که مشاهده می‌شود با این مجموعه دادگان، در بیشتر معیارهای سنجش، در حالت میانگین، یکی از نسخه‌های روش پیشنهادی مبتنی بر درهم‌تندگی تصویر حائز بهترین عملکرد بوده است. در معیارهای صحت، نرخ مثبت کاذب و نرخ منفی کاذب، روش پیشنهادی نسبت به روش [7] کارا تر بوده است.

مشخص شده است. سطر آخر هر جدول نیز میانگین نتایج برای هر ستون را نشان می‌دهد که در اینجا نیز بهترین مورد با زمینه خاکستری مشخص شده است.

(جدول-۲): ارزیابی مبتنی بر معیار حساسیت. ستون دوم: روش

[17]، ستون سوم: روش LbM [7]، ستون‌های چهارم و پنجم:

روش پیشنهادی WLbM (بخش ۳-۱) با ضرایب مختلف برای

زمینه و پیش‌زمینه.

(Table-2): Performance comparisons based on Sensitivity. First column: Image Number in DRIVE dataset, second column: [17], third column: LbM [7], the last two columns are the proposed WLbM method with different values for (C_b, C_f) .

(Sec. 1-3).

روش مرجع	ضرایب جریمه زمینه و پیش‌زمینه	800	(800, 1)	(1, 800)
[17]				
۱	0.758	0.767	0.738	0.769
۲	0.742	0.765	0.766	0.755
۳	0.703	0.731	0.717	0.710
۴	0.726	0.727	0.701	0.728
۵	0.762	0.769	0.742	0.773
۶	0.671	0.671	0.659	0.673
۷	0.756	0.758	0.680	0.759
۸	0.717	0.720	0.580	0.724
۹	0.680	0.684	0.641	0.688
۱۰	0.763	0.767	0.700	0.769
۱۱	0.711	0.721	0.701	0.719
۱۲	0.746	0.751	0.725	0.754
۱۳	0.651	0.665	0.653	0.656
۱۴	0.768	0.772	0.685	0.774
۱۵	0.717	0.719	0.586	0.720
۱۶	0.739	0.745	0.712	0.747
۱۷	0.751	0.755	0.721	0.757
۱۸	0.718	0.745	0.748	0.738
۱۹	0.718	0.735	0.724	0.726
۲۰	0.669	0.675	0.664	0.677
میانگین	0.723	0.732	0.692	0.731

اگر (True Positive) TP، معرف تعداد مثبت صادق، (True Negative) TN، معرف تعداد منفی صادق، (False Positive) FP (False Negative) FN معرف تعداد مثبت کاذب و (False Negative) FN معرف تعداد منفی کاذب باشد، معیار صحت بر اساس رابطه $Acc = (TP + TN) / (TP + TN + FP + FN)$ و معیار حساسیت یا نرخ مثبت صادق به‌صورت $TPR = TP / (TP + FN)$ محاسبه می‌شود.

در این مقایسه‌ها بهترین حالت روش پیشنهادی با جواب اولیه مقایسه شده است. ستون نخست، شماره تصویر از مجموعه دادگان مورد استفاده است. ستون دوم مقدار معیار خروجی روش [17] در مقایسه با جواب دانسته را نشان می‌دهد.

(جدول-۴): ارزیابی مبتنی بر معیار نرخ منفی کاذب. ستون دوم: روش [17]، ستون سوم: روش LbM [7]، ستون‌های چهارم و پنجم: روش پیشنهادی WLbM (بخش ۳-۱) با ضرایب مختلف برای

زمینه و پیش‌زمینه

(Table-4): Performance comparisons based on False Negative Rate. First column: Image Number in DRIVE dataset, second column: [17], third column: LbM [7], the last two columns are the proposed WLbM method with different values for (C_b, C_f) .

روش مرجع	ضرایب جرمه زمینه و پیش‌زمینه	1	2	3
[17]	800	(800, 1)	(1, 800)	
۱	0.036	0.034	0.039	0.034
۲	0.045	0.041	0.040	0.042
۳	0.044	0.040	0.042	0.043
۴	0.040	0.040	0.043	0.040
۵	0.032	0.031	0.034	0.030
۶	0.051	0.051	0.053	0.051
۷	0.028	0.028	0.037	0.028
۸	0.031	0.030	0.045	0.030
۹	0.042	0.042	0.048	0.041
۱۰	0.028	0.027	0.035	0.027
۱۱	0.040	0.038	0.041	0.038
۱۲	0.033	0.033	0.036	0.032
۱۳	0.060	0.058	0.060	0.059
۱۴	0.028	0.028	0.039	0.028
۱۵	0.034	0.034	0.050	0.034
۱۶	0.037	0.036	0.041	0.036
۱۷	0.030	0.030	0.034	0.029
۱۸	0.043	0.039	0.040	0.039
۱۹	0.047	0.045	0.046	0.046
۲۰	0.051	0.050	0.052	0.050
میانگین	0.039	0.038	0.043	0.038

در معیار صحت هم فقط در چهار مورد، روش پایه، بهتر از روش‌های پیشنهادی بوده، اما در حالت میانگین با اختلاف ۰/۰۰۱، روش پایه بهتر بوده است. در مورد ضریب تخطی، مجدداً از همان مقدار هشتصد مورد استفاده نویسندگان [7] استفاده شده است.

مشابه با کاربرد قبل، مقدار بهینه را نیز می‌توان جستجو کرد؛ اما با همین مقدار نیز، برتری نسخه‌هایی از شیوه بهبودیافته پیشنهادی، مشخص است. از آنجا که مطابق بخش قبل، نقشه سه‌گانه مبتنی بر روش‌های شناسایی رگ موجود ایجاد می‌شود، این شیوه امکان آشکارسازی عروق پیداشده با شیوه مورد استفاده استخراج رگ را ندارد و تنها امکان دقیق‌تر کردن برجسب‌گذاری در ناحیه مرزی نقشه سه‌گانه را دارد.

در این بخش بهترین جواب روش‌های پیشنهادی از طریق مقایسه نتایج آستانه‌های مختلف ماتریس آلفا و جواب واقعی مشخص شده است؛ اما در عمل این شیوه مناسب نیست

و باید یک روش آستانه‌یابی خودکار استفاده شود. روش Leave-One-Out مورد بررسی قرار گرفت و نتایج تا حدود زیادی همانند نتایج یادشده بودند.

۵- نتیجه‌گیری

در سالیان اخیر، روش‌های درهم‌تنیدگی تصویر که با تعامل با کاربر به قطعه‌بندی ظریف تصویر می‌پردازند، توجهات زیادی را به خود معطوف کرده‌اند. یکی از روش‌های مطرح در این حوزه، روش درهم‌تنیدگی تصویر مبتنی بر یادگیری است. در این شیوه میزان تخطی جواب حاصل از داده‌های آموزشی یکسان در نظر گرفته می‌شود. نوآوری اصلی مقاله، در نظر گرفتن وزن متفاوت برای داده‌های آموزشی متعلق به دو طبقه مورد شناسایی، در روش درهم‌تنیدگی مبتنی بر یادگیری است. یک شیوه خودکار تولید نقشه سه‌گانه نیز معرفی شد که با عملیات مورفولوژی مناسب، نقشه سه‌گانه بر اساس یک روش قطعه‌بندی پیشین تولید می‌شود. روش‌هایی نیز برای پیدا کردن مقادیر بهینه ضرایب تخطی و حد آستانه مناسب بیان شد.

به‌منظور مقایسه شیوه اصلاحی با روش اصلی، هر دو شیوه در دو کاربرد مختلف برای بهبود روش‌های قطعه‌بندی قبلی به کار گرفته شدند. نتایج آزمایش‌های انجام‌شده کارایی شیوه پیشنهادی در حوزه بهبود نتایج «استخراج متن از تصویر» و «استخراج رگ‌های خونی شبکیه چشم» را نشان دادند. در کاربرد دقیق‌تر کردن کیفیت خروجی متن مستخرج از تصویر، ضرایب تخطی نامناسب، منجر به تولید حالت پخش‌شدگی جوهر روی متن شد که می‌تواند به‌عنوان جلوه‌های ویژه در نرم‌افزارهایی همچون فوتوشاپ به کار گرفته شود.

6- References

۶- مراجع

- [1] J. Wang and M. F. Cohen, "Image and video matting: A survey," *Found. Trends. Comput. Graph. Vis.*, vol.3, pp.97-175, Jan. 2007.
- [2] J. Johnson, D. Rajan, and H. Cholakkal, "Sparse codes as alpha matte," in *BMVC*, BMVA Press, 2014.
- [3] Q. Chen, D. Li, and C.-K. Tang, "KNN mat-ting," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol.35, pp.2175-2188, Sept 2013.
- [4] J. Gao, M. Paul, and J. Liu, "The image mat-ting method with regularized matte," *IEEE Computer Society in ICME*, , pp.550-555, 2012.

segmentation in retinal images,” in *Medical Image Computing and Computer-Assisted Intervention - MIC-CAI 2014*, vol. 8673 of *Lecture Notes in Computer Science*, pp.634–641, Springer, 2014.

- [19] V. M. Saffarzadeh, A. Osarch, and B. Shadgar, “Vessel segmentation in retinal images using multi-scale line operator and k-means clustering,” *Journal of Medical Signals & Sensors*, vol. 4, no.2, pp.122–129, 2014.
- [20] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov, “Trainable COSFIRE filters for vessel delineation with application to retinal images,” *Medical Image Analysis*, vol.19, no.1, pp.46–57, 2015.
- [۲۱] م. زردادی، ن. مهرشاد. “آشکارسازی عروق شبکیه چشم بر اساس مدل محاسباتی سلول ساده کورتکس اولیه بینایی.” فصل‌نامه پردازش علائم و داده‌ها، دوره ۱۳، شماره ۱، ص ۱۳۸–۱۲۷، ۱۳۹۵.
- [21] M. Zardadi, and N. Mehrshad, “A New Approach to Retinal Vessel Segmentation by Using Computational Model of Simple Cells in Primary Visual Cortex”. *JSDP*, vol. 13, no. 1, pp. 127–138, 2016.



محمود امین‌طوسی، دوره‌های کارشناسی و کارشناسی ارشد خود را در رشته‌های ریاضی و مهندسی کامپیوتر (نرم‌افزار) در دانشگاه فردوسی به اتمام رسانده و دوره دکترای خود را در رشته مهندسی کامپیوتر (هوش مصنوعی) در دانشگاه علم و صنعت ایران گذرانده است. علایق پژوهشی وی بینایی ماشین، یادگیری ماشین و بهینه‌سازی ترکیبیاتی است. نشانی رایانامه ایشان عبارت است از:

m.amintoosi@hsu.ac.ir

- [5] P. G. Lee and Y. Wu, “L1 matting,” *IEEE in ICIP*, pp.4665–4668, , 2010.
- [6] I. Choi, S. Kim, M. S. Brown, and Y. W. Tai, “A learning-based approach to reduce jpeg ar-tifacts in image matting,” in *2013 IEEE International Conference on Computer Vision*, pp.2880–2887, Dec 2013.
- [7] Y. Zheng and C. Kambhamettu, “Learning based digital matting,” in *12th International Conference on Computer Vision*, (Kyoto), pp.889–896, 2009.
- [8] A. Levin, D. Lischinski, and Y. Weiss, “A closed-form solution to natural image matting,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.30, no.2, pp.228–242, 2008.
- [9] J. Sun, J. Jia, C.-K. Tang, and H.-Y. Shum, “Poisson matting,” *ACM Trans. Graph.*, vol.23, pp.315–321, Aug. 2004.
- [10] Z. Zhang, Q. Zhu, and Y. Xie, “Learning based alpha matting using support vector regression,” in *2012 ,19th IEEE International Conference on Image Processing*, pp. 2109–2112, Sept 2012.
- [11] X. Li and Q. Cui, *Parallel Accelerated Matting Method Based on Local Learning*, pp.152–162 Cham: Springer International Publishing, 2017.
- [12] K. Jung, K. Kim, and A. Jain, “Text information extraction in images and video: a survey,” *Pattern Recognition*, vol.37, pp.977–997, 5 2004.
- [13] N. Otsu, “A Threshold Selection Method from Gray-level Histograms,” *IEEE Transactions on Systems, Man and Cybernetics*, vol.9, no.1, pp.62–66, 1979.
- [14] M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. Rudnicka, C. Owen, and S. Barman, “Blood vessel segmentation methodologies in retinal images - a survey,” *Comput. Methods Prog. Biomed.*, vol. 108, pp.407–433, Oct. 2012.
- [15] M. T. Dchkordi, S. Sadri, and A. Doosthoscini, “A review of coronary vessel segmentation algorithms,” *Journal of Medical Signals & Sensors*, vol.1, no.1, pp.49–54, 2011.
- [16] P. Talwar, M. D. Gupta, “Alpha-matting based retinal vessel extraction,” *United States Patent Application* 20160163041, June 2016.
- [17] P. Bankhead, C. N. Scholfield, J. G. McGeown, and T. M. Curtis, “Fast retinal vessel detection and measurement using wavelets and edge location refinement,” *PloS one*, vol.7, no.3, 2012.
- [18] J. I. Orlando, M. Blaschko, “Learning fullyconnected CRFs for blood vessel

