

Protein Spot Matching in Two Dimensional Gel Electrophoresis Images based on Probabilistic Models

S. Shamekhi^{1*}, M. H. Miran Baygi², A. Gooya³

¹ Assistant Professor, Bioelectric Department, Biomedical Engineering Faculty, Sahand University of Technology, Tabriz, Iran

² Associate Professor, Bioelectric Department, Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran

³ Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom

Receipt in the Online Submission System: 4 May 2018, Received in Revised Form: 30 July 2018, Accepted: 4 December 2018

Abstract

Matching of the protein spots in two dimensional gel electrophoresis (2DGE) images is a main process of analyzing these images. Due to the challenges of 2DGE images such as the presence of noise and artifacts, the matching of protein spots is performed under human supervision. This supervision involves human errors. Therefore, in this work a new automated algorithm has been proposed for spot matching in 2DGE images which is based on a probabilistic model. Due to the complexities of the proposed model, the Variational Bayes has been used to solve the equations of the model. The performance of the proposed algorithm has been evaluated on real and synthetic 2DGE images with some statistical criteria. Protein spots in real image dataset have been matched by the proposed method with angular error of 0.05 and end point error of 1.46 and in synthetic image dataset with angular error of 0.13 and end point error of 0.90. These results reveal higher precision and effectiveness and lower matching error of the proposed method.

Keywords: *Two Dimensional Gel Electrophoresis, Spot Segmentation, Matching, Variational Bayes*

*Corresponding Author

Address: Bioelectric Department, Biomedical Engineering Faculty, Sahand University of Technology, P. O. Box: 5331841133, Tabriz, Iran

Tel: +98-41-33458455

Fax: +98-41-33444360

E-mail: shamekhi@sut.ac.ir

تطبیق لکه‌های پروتئینی در تصاویر الکتروفورز ژل دوبعدی بر پایه‌ی مدل‌های احتمالاتی

سینا شامخی^{۱*}، محمدحسین میران‌بیگی^۲، علی گویا^۳

^۱ استادیار، دانشکده‌ی مهندسی پزشکی، دانشگاه صنعتی سهند، تبریز، ایران
^۲ دانشیار، دانشکده‌ی مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، ایران
^۳ دانشکده‌ی مهندسی برق و الکترونیک، دانشگاه شفیلد، شفیلد، انگلستان

تاریخ ثبت در سامانه: ۱۳۹۷/۲/۱۴، بازنگری: ۱۳۹۷/۵/۸، پذیرش قطعی: ۱۳۹۷/۹/۱۳

چکیده

تطبیق لکه‌های پروتئینی در تصاویر الکتروفورز ژل دوبعدی (2DGE)، یکی از فرایندهای اصلی در تحلیل این تصاویر است. به دلیل چالش‌های موجود در تصاویر 2DGE، هم‌چون حضور نویز و آرتیفکت، تطبیق لکه‌های پروتئینی، تحت نظارت نیروی انسانی انجام گرفته که این نظارت، احتمال وقوع خطاهای انسانی را به دنبال خواهد داشت. از این‌رو، در این تحقیق، یک الگوریتم جدید و خودکار برای تطبیق لکه‌های پروتئینی، مبتنی بر مدل‌های احتمالاتی، ارائه شده است. به دلیل پیچیدگی مدل احتمالاتی پیشنهادی، برای حل روابط موجود، از روش بیز-وردشی استفاده شده است. عمل کرد الگوریتم پیشنهادی، به کمک معیارهایی آماری روی مجموعه‌ی تصاویر واقعی و مصنوعی تهیه شده در این تحقیق، مورد ارزیابی قرار گرفته است. بر این اساس، لکه‌های پروتئینی توسط الگوریتم پیشنهادی، در تصاویر واقعی با خطای زاویه‌ای ۰/۰۵ رادیان و خطای نقطه‌ی پایانی ۱/۴۶ پیکسل و در تصاویر مصنوعی با خطای زاویه‌ای ۰/۱۳ رادیان و خطای نقطه‌ی پایانی ۰/۹۰ پیکسل، تطبیق داده شده‌اند. این نتایج، نشان دهنده‌ی دقت، کارایی بالاتر و خطای تطبیق پایین‌تر الگوریتم پیشنهادی در مقایسه با سایر روش‌های مورد بررسی می‌باشد.

کلیدواژه‌ها: الکتروفورز ژل دوبعدی، ناحیه‌بندی لکه، تطبیق، بیز-وردشی

*نویسنده مسئول

نشانی: گروه بیوالکترونیک، دانشکده‌ی مهندسی پزشکی، دانشگاه صنعتی سهند، تبریز، ایران، کد پستی: ۵۳۳۱۸۴۱۱۳۳

تلفن: ۳۳۴۵۸۴۵۵ (۴۱) ۹۸+

دورنگار: ۳۳۴۴۳۶۰ (۴۱) ۹۸+

پست الکترونیکی: shamekhi@sut.ac.ir

۱- مقدمه

پروتئومیک، علمی است که به مطالعه در زمینه‌ی پروتئین‌های تشکیل‌دهنده‌ی یک نمونه از بافت یا ارگانیسم بدن می‌پردازد. در این علم، روش اصلی جداسازی پروتئین‌های یک ترکیب، الکتروفورز ژل دوبعدی (2DGE) است که در اوایل دهه‌ی ۱۹۷۰ معرفی شد [۲،۱]. این روش، از دو مرحله‌ی اصلی تشکیل شده است. در مرحله‌ی اول، پروتئین‌ها در یک میدان الکتریکی، بر اساس بار الکتریکی خود، از یکدیگر جدا شده و بعد اول از ژل دوبعدی شکل می‌گیرد. در مرحله‌ی دوم، پروتئین‌ها بر اساس جرم مولکولی نسبی خود از یکدیگر تفکیک شده و بعد دوم ایجاد می‌شود [۲]. پس از رنگ‌آمیزی ژل‌های دوبعدی، پروتئین‌های جدا شده به صورت تعداد زیادی لکه‌ی^۲ تیره روی ورقه‌ی ژل نمایان می‌شوند. در انتها، تصویر نهایی 2DGE، که حاصل از اسکن ژل دوبعدی است، شکل می‌گیرد. لکه‌های پروتئینی موجود در این تصاویر، پس از تشخیص، تطبیق داده شده و تغییرات ایجاد شده در تصاویر حاصل از یک نمونه، در حالت‌های کنترل، بیمار و تحت درمان مورد مقایسه و بررسی قرار می‌گیرند. از اواخر دهه‌ی ۱۹۹۰، تحقیقات ویژه‌ای برای تطبیق تصاویر 2DGE و تطابق لکه‌های پروتئینی، صورت گرفته است. در این میان، می‌توان به استفاده از گراف‌های ساختاری [۳]، به کارگیری هم‌زمان نشانه‌های نقطه‌ای و شدت روشنایی هر پیکسل [۴] و هم‌چنین روش نزدیک‌ترین نقطه‌ی تکراری [۵] اشاره کرد. روگرز و هم‌کارانش در سال ۲۰۰۷، از یک روش مبتنی بر تطبیق نقطه، برای تطابق لکه‌ها در تصاویر 2DGE استفاده کرد [۶]. در سال‌های ۲۰۰۸ و ۲۰۱۰، سان و هم‌کارانش، روش‌هایی مبتنی بر تخمین جریان نوری برای محاسبه‌ی بردارهای جابه‌جایی در تصاویر را ارائه کردند [۸،۷]. در سال‌های اخیر، آن‌ها روشی مبتنی بر تخمین نوری، با عنوان "کلاسیک+عبارت غیرمحملی" (CNLT^۳) را برای تخمین بردارهای جابه‌جایی در تصویر، ارائه کردند [۹]. رودریگز و هم‌کارانش نیز به کمک روش‌های تطبیق بلوکی و مدل‌های دگرشکلی، روشی را برای محاسبه‌ی جابه‌جایی کش‌سان لکه‌های پروتئینی در تصاویر 2DGE پیشنهاد کردند [۱۰]. در سال ۲۰۱۶، ژین و هم‌کارانش، روشی مبتنی بر اطلاعات فاصله، شدت و الگو را برای تطبیق تصاویر، ارائه کردند [۱۱]. علی‌زاده و هم‌کارانش نیز در سال ۲۰۱۷، با استفاده از روش‌های شکلی و بررسی ویژگی پیکسل‌های موجود در همسایگی، روشی را

برای تطبیق لکه‌های پروتئینی، ارائه دادند [۱۲]. با وجود گستردگی این تحقیقات، نقش نیروی انسانی در تنظیم پارامترها و تصحیح نتایج در روش‌های مذکور، موجب ایجاد خطا در فرایند پردازش و تحلیل تصاویر 2DGE می‌گردد. در این تحقیق، با الهام از روش‌های ناحیه‌بندی یک‌پارچه [۱۳-۱۵]، مدل احتمالاتی نوینی برای کاهش دخالت کاربران و حل خودکار مساله‌ی تطبیق لکه‌های پروتئینی در تصاویر 2DGE، ارائه شده و مورد استفاده قرار گرفته است. در مدل‌های احتمالاتی پیچیده، مانند مدل طراحی شده در این تحقیق، به دلیل ابعاد و پارامترهای متعدد مدل و وجود جمع‌های پیچیده روی متغیرها، ارزیابی توزیع‌های پسین و محاسبه‌ی امیدهای ریاضی، یا غیرممکن بوده و یا مستلزم صرف هزینه‌های بسیار زیادی می‌باشد. در این شرایط، از روش‌های تخمین، هم‌چون روش بیز-وردشی (VB^۴)، استفاده می‌شود [۱۴]. محاسبات مبتنی بر بیز-وردشی، در حوزه‌هایی مانند المان محدود و بیشینه‌ی آن‌تروپی، به طور گسترده مورد استفاده قرار گرفته [۱۴،۱۶] و در تحقیقات متعددی گسترش یافته و بهینه‌سازی شده است [۱۷-۱۹].

مطالب این مقاله، در چهار بخش اصلی ارائه شده است. در بخش دوم، الگوریتم تطبیق خودکار پیشنهادی ارائه می‌گردد. در بخش سوم، نتایج حاصل از این تحقیق، ارائه شده و مورد بررسی و مقایسه قرار گرفته و در نهایت، در بخش چهارم، به جمع‌بندی مطالب و نتیجه‌گیری پرداخته شده است.

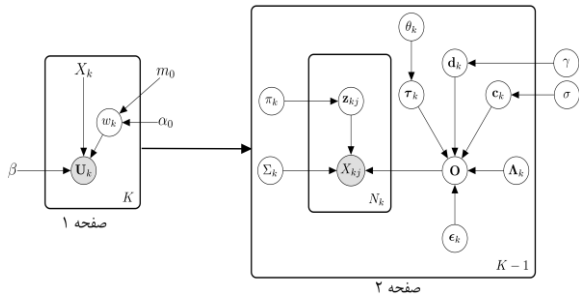
۲- روش پیشنهادی برای تطبیق لکه‌ها

چنانچه پیش‌تر نیز اشاره شد، در این تحقیق، مساله‌ی ناحیه‌بندی و تطبیق خودکار در تصاویر 2DGE، به صورت یک‌پارچه و توسط روش‌های احتمالاتی، مدل‌سازی شده، حل شده و مورد بررسی قرار گرفته است. در این بخش، در راستای طراحی مدل احتمالاتی مذکور، ابتدا تعاریف اولیه و متغیرهای اصلی مورد نیاز، ارائه گشته و در ادامه، جهت بیان بهتر ساختار مدل احتمالاتی و تسهیل در تنظیم روابط، یک مدل گرافیکی از مجموعه‌ی متغیرها و پارامترهای موجود، ترسیم شده و جزئیات مربوط به آن مورد بررسی قرار گرفته است.

۲-۱- تعاریف اولیه و مدل گرافیکی

در فرایند ثبت تصاویر 2DGE، تعدادی نمونه‌ی پروتئینی برای بررسی وجود دارد، که از هر یک از این نمونه‌ها معمولاً ۲ یا ۳

^۳ Classic+Non Local Term^۴ Variational Bayes^۱ Two-Dimensional Gel Electrophoresis^۲ Spot



شکل (۱) - مدل گرافیکی پیشنهادی برای ناحیه‌بندی و تطبیق یک پارچه

۲-۱-۱- توابع احتمال پیشین

در این مقاله، احتمالات پیشین صفحات ۱ و ۲، به صورت آنچه در ادامه آمده، در نظر گرفته شده است. در صفحه‌ی ۱، احتمال مشاهدات به شرط متغیرهای پنهان و پارامترهای جزئی، با فرض نرمال بودن شدت پیکسل در تصاویر، به صورت زیر نوشته شده است:

$$p(\mathbf{U} | \mathbf{W}) = \prod_{k=1}^K \prod_{n=1}^N N(u_{kn} | \mathbf{w}_k^T \phi_{\mathbf{X}_k}(x_n), \beta^{-1}) \quad (۴)$$

با این فرض که بردار وزن‌ها، متغیرهایی نرمال هستند، توزیع \mathbf{W} نیز به صورت رابطه‌ی (۵) تعریف می‌گردد.

$$p(\mathbf{W}) = \prod_{k=1}^K N(\mathbf{w}_k | \mathbf{m}_{0k}, \alpha_{0k}^{-1} \mathbf{I}_{N_k}) \quad (۵)$$

توزیع پیشنهادی برای مکان لکه‌ها در تصویر مرجع، در تابع احتمال مشترک صفحه‌ی ۲، به صورت زیر است:

$$p(\mathbf{X}_R | \mathbf{O}, \square) = \prod_{k=1}^K \prod_{j=1}^{N_R} \prod_{m=1}^{M_k} N(X_{Rkj} | \mathbf{o}_{km}, \Sigma_{km}^{-1})^{z_{kjm}} \quad (۶)$$

احتمال شرطی متغیرهای پنهان \square ، نسبت به ضرایب ترکیب π نیز به صورت رابطه‌ی (۷) تعریف شده‌اند:

$$p(\square | \pi) = \prod_{k=1}^K \prod_{j=1}^{N_R} \prod_{m=1}^{M_k} \pi_{km}^{z_{kjm}} \quad (۷)$$

توزیع احتمال پیشین متغیر π_k نیز در قالب یک توزیع دریکله، به صورت رابطه‌ی (۸) تعریف شده است:

$$p(\pi_k) = C(\kappa_0) \prod_{m=1}^{M_k} \pi_{km}^{\kappa_0 - 1} \quad (۸)$$

عدد ژل دوبعدی تهیه می‌گردد. بنابراین، مشاهدات مدل احتمالاتی، مجموعه‌ای K -تایی از تصاویر 2DGE است که با \mathbf{U}_k نشان داده می‌شود. این مجموعه‌ی K -تایی از تصاویر، در قالب $\mathbf{U} = \{\mathbf{U}_1 \dots \mathbf{U}_k \dots \mathbf{U}_K\}$ قرار داده شده است. هر یک از تصاویر ورودی، دارای N پیکسل است که شدت روشنایی پیکسل n -ام از تصویر k -ام، به صورت u_{kn} نشان داده می‌شود. موقعیت لکه‌های موجود در تصاویر ورودی، توسط مجموعه‌ی \mathbf{O}_k قالب بردار ذخیره می‌گردد. در این بردار، O_{km} موقعیت لکه‌ی m -ام روی مدل بوده و M_k تعداد لکه‌های موجود در مدل‌های احتمالاتی مربوط به تصویر k -ام می‌باشد. مجموعه‌ی متغیرهای پنهان تصاویر نیز توسط $\square = \{Z_1 \dots Z_k \dots Z_K\}$ بیان می‌شود، به طوری که در بردار Z_{kj} تنها یکی از اعضا (Z_{kjm}) برابر با ۱ می‌باشد.

جابه‌جایی لکه‌های پروتئینی متناظر در تصاویر 2DGE، غیر یک‌نواخت می‌باشد. در واقع، لکه‌های پروتئینی متناظر در تصاویر مختلف به دست آمده از یک نمونه‌ی پروتئینی، به صورت مشابه جابه‌جا نمی‌شوند. از این‌رو، در این تحقیق، الگوی جابه‌جایی بین تصاویر غیرصلب^۱ مد نظر قرار داده شده و از تبدیل TPS^۲ (رابطه‌ی ۱)، برای بیان این جابه‌جایی استفاده شده است [۲۱، ۲۰].

$$\mathbf{T}_k = \mathbf{d}_k \tilde{\mathbf{X}}_k + \mathbf{C}_k \Phi(\tilde{\mathbf{X}}_k) \quad (۱)$$

در این رابطه، \mathbf{d}_k بخش نسبی تابع جابه‌جایی بوده و \mathbf{C}_k ماتریس وزن‌ها و بایاس تطبیق غیرصلب می‌باشد. $\tilde{\mathbf{X}}_k$ نیز برابر $[1; \mathbf{X}_k]$ بوده و $\Phi(\tilde{\mathbf{X}}_k)$ ماتریسی است که بر اساس کرنل TPS به دست می‌آید.

بر پایه‌ی متغیرهای موجود در مساله، توابع احتمال مشترک، توسط دو تابع مجزا به صورت روابط (۲) و (۳) بیان می‌شوند.

$$p(\mathbf{U}, \mathbf{W}) = p(\mathbf{U} | \mathbf{W}) p(\mathbf{W}) \quad (۲)$$

$$p(\mathbf{X}, \mathbf{O}, \Theta) = p(\mathbf{X} | \mathbf{O}, \Theta) p(\mathbf{O}) p(\Theta) \quad (۳)$$

همان‌طور که در شکل (۱) نشان داده شده است، مدل گرافیکی مجموعه‌ی مورد بررسی، در دو صفحه‌ی مجزا از هم قابل ترسیم می‌باشد.

^۲ Thin-Plate Spline

^۱ Non-Rigid

در این روابط، D تعداد بعد تصاویر و $\mathbf{d}_{kq} = \mathbf{d}_k(q,:)^T$ و $\mathbf{C}_{kq} = \mathbf{C}_k(q,:)^T$ بردارهای ستونی هستند. هم‌چنین، توزیع احتمال دقت متغیر X_R ، در قالب یک توزیع وایشارت، به صورت رابطه‌ی (۱۶) فرض شده است:

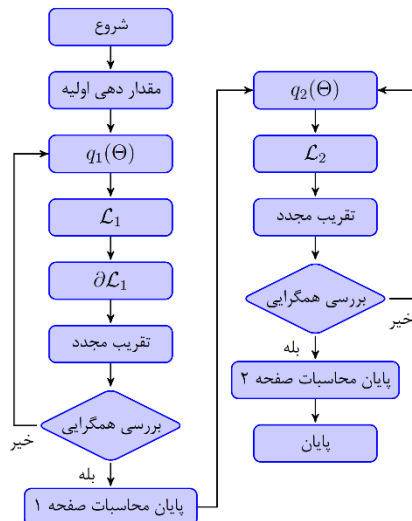
$$p(\Sigma) = \prod_{k=1}^K \prod_{m=1}^{M_k} W(\Sigma_{km} | Y_0, \xi_0) \quad (16)$$

۲-۲- توزیع‌ها و محاسبات بیز-وردشی (VB)

در روش VB، با فرض این‌که توزیع مشترک پارامترها و مشاهدات به صورت $p(U, \Theta)$ در نظر گرفته شده باشد، پارامترهای مدل (Θ) ، به کمک تخمین توزیع پسین پارامترها نسبت به مشاهدات $(p(\Theta | U))$ تخمین زده می‌شوند. لگاریتم احتمال کناری مدل را می‌توان به صورت رابطه‌ی (۱۷) تجزیه کرد:

$$\ln p(U) = L + KL(q \| p) \quad (17)$$

در این رابطه، L باند پایین^۴ و $KL(q \| p)$ هم‌گرایی کولبک-لیبلر^۵ هستند [۱۴]. روند تخمین پارامترهای مدل با بیشینه نمودن باند پایین، توسط بهینه‌سازی تخمین $p(\Theta | U)$ و نیز فاکتوریزه کردن تابع $q(\Theta)$ انجام می‌شود [۱۴، ۱۷، ۱۸]. روند انجام محاسبات وردشی مدل احتمالاتی پیشنهادی، در شکل (۲) نشان داده شده است.



شکل (۲) - روند انجام محاسبات در مدل گرافیکی پیشنهادی

در محاسبات مربوط به صفحه‌ی ۱، با تعیین مقادیر اولیه‌ی پارامترها، محاسبات آغاز می‌گردد. جهت محاسبه‌ی مقدار

از سوی دیگر، توزیع مکان قرارگیری لکه‌ها در مدل، $p(O)$ ، به صورت رابطه‌ی (۹) تعریف شده است:

$$p(O | \mathbf{d}, \mathbf{C}, \Lambda, \delta, \tau) = \prod_{k=1}^K \prod_{m=1}^{M_k} \prod_{t=1}^{T_n} N(\mathbf{o}_{km} | \mathbf{d}_k \tilde{X}_{km} + \mathbf{C}_k \phi(\tilde{X}_{km}) + \delta_{kt}, \Lambda_k^{-1})^{\tau_{km}} \quad (9)$$

در این رابطه، δ_{kt} یک نویز گوسی با میانگین m_0 بوده و احتمال توام متغیرهای Λ و δ ، در قالب یک توزیع گوسین-وایشارت در نظر گرفته شده است. توزیع احتمال متغیر پنهان τ به صورت رابطه‌ی (۱۰) و توزیع احتمال پیشین θ_k نیز به صورت رابطه‌ی (۱۱)، در قالب یک توزیع دریکله فرض شده است:

$$p(\tau | \theta) = \prod_{k=1}^K \prod_{m=1}^{M_k} \prod_{t=1}^{T_n} \theta_{kt}^{\tau_{km}} \quad (10)$$

$$p(\theta_k) = C(\omega_0) \prod_{t=1}^{T_n} \theta_{kt}^{\omega_0 - 1} \quad (11)$$

در طراحی توابع توزیع پیشین مربوط به متغیرهای موجود در تابع تبدیل جابه‌جایی (رابطه‌ی ۱) و تخمین توابع توزیع پسین مربوطه، توجه به موضوع تنگی^۱ ضرایب از اهمیت ویژه‌ای برخوردار است. از این‌رو، برای تخمین تنگ احتمال پسین و افزایش سرعت محاسبات، روش تعیین خودکار رابطه (ARD^۲) به کار گرفته شده است [۱۴، ۱۶]. بیل^۳، در سال ۲۰۰۳، جهت اعمال ARD در توزیع مربوط به ماتریس‌ها، هر یک از سطرهای موجود در ماتریس را در قالب یک بردار ستونی تنظیم نمود و ARD را بر هر یک از این بردارهای ستونی اعمال کرد [۱۶]. در تحقیق حاضر نیز با استفاده از روش ارائه شده توسط بیل، توزیع‌ها و پارامترهای مربوط به متغیرهای d و C از رابطه‌ی جابه‌جایی (رابطه‌ی ۱)، به صورت روابط (۱۲) تا (۱۵) تعریف شده‌اند:

$$p(\mathbf{d} | \gamma) = \prod_{k=1}^K \prod_{q=1}^D N(\mathbf{d}_{kq} | 0, \text{diag}(\gamma^{-1})) \quad (12)$$

$$p(\gamma) = \text{Gamma}(\gamma | a_0, b_0) \quad (13)$$

$$p(\mathbf{C} | \sigma) = \prod_{k=1}^K \prod_{q=1}^D N(\mathbf{C}_{kq} | 0, \text{diag}(\sigma^{-1})) \quad (14)$$

$$p(\sigma) = \text{Gamma}(\sigma | e_0, f_0) \quad (15)$$

^۴ Lower Bound

^۵ Kullback-Leibler Divergence

^۱ Sparsity

^۲ Automatic Relevance Determination

^۳ Beal

حاصل ضرب مجموعه‌ای از توابع نرمال، به صورت رابطه‌ی (۲۵) بازنویسی شده است، ارائه می‌گردد:

$$q(\mathbf{O}) = \prod_{k=1}^K \prod_{m=1}^{M_k} N(\mathbf{o}_{km} | \boldsymbol{\mu}_{km}, \mathbf{L}_{km}^{-1}) \quad (25)$$

در این رابطه، $\boldsymbol{\mu}_{km}$ و \mathbf{L}_{km}^{-1} به ترتیب میانگین و واریانس مدل میانگین بوده و به صورت روابط (۲۶) و (۲۷) تعریف می‌شوند:

$$\mathbf{L}_{km} = \sum_{t=1}^{T_n} g_{kmt} < \boldsymbol{\Lambda}_{kt} > + N_{km} < \boldsymbol{\Sigma}_{km} > \quad (26)$$

$$\boldsymbol{\mu}_{km} = \mathbf{L}_{km}^{-1} \left(\sum_{t=1}^{T_n} g_{kmt} < \boldsymbol{\Lambda}_{kt} > (\mathbf{d}\mathbf{m}_k \tilde{\mathbf{X}}_k + \mathbf{C}\mathbf{m}_k \Phi(\tilde{\mathbf{X}}_k) + \mathbf{m}_{kt}) + \sum_{j=1}^{N_k} r_{kjm} < \boldsymbol{\Sigma}_{km} > X_{kj} \right), \quad N_{km} = \sum_{j=1}^{N_k} r_{kjm} \quad (27)$$

در این روابط، $< \cdot >$ نماد امید ریاضی بوده و \mathbf{g}_{kmt} ، \mathbf{m}_{kt} به ترتیب امید ریاضی Z_{kjm} ، τ_{kmt} و δ_{kt} می‌باشند. پس از انجام محاسبات و ساده‌سازی‌های مورد نیاز، توزیع پسین متغیرهای پنهان ($q(\square)$) نیز به صورت زیر به دست می‌آید:

$$q(\square) = \prod_{k=1}^K \prod_{j=1}^{N_k} \prod_{m=1}^{M_k} r_{kjm}^{z_{kjm}} \quad (28)$$

$$E[z_{kjm}] = r_{kjm} = \frac{\rho_{kjm}}{\sum_{m=1}^{M_k} \rho_{kjm}}, \quad N_m = \sum_{k=1}^K \sum_{j=1}^{N_k} r_{kjm}. \quad (29)$$

توزیع پسین متغیر π ، مانند توزیع پیشین آن، در قالب یک توزیع دریکله، به صورت روابط (۳۰) و (۳۱) به دست می‌آید:

$$q(\boldsymbol{\pi}) = \prod_{k=1}^K C(\kappa_k) \prod_{m=1}^{M_k} \pi_{km}^{\kappa_{km}-1} \quad (30)$$

$$\kappa_{km} = \kappa_0 + N_{km}, \quad N_{km} = \sum_{j=1}^{N_k} r_{kjm}, \quad \kappa = \kappa_0 + N_k. \quad (31)$$

همچنین، توزیع $q(\delta, \boldsymbol{\Lambda})$ در قالب یک توزیع گوسین-وایشارت به دست می‌آید. در انتها و بر پایه‌ی اصول بیز-وردشی، رابطه‌ی کلی L_2 به صورت زیر قابل بیان است:

$$\begin{aligned} L_2 = & E[\ln p(\mathbf{X}_R | \mathbf{O}, \square, \boldsymbol{\Sigma})] + E[p(\mathbf{O} | \mathbf{d}, \mathbf{C}, \boldsymbol{\Lambda}, \delta, \boldsymbol{\tau})] \\ & + E[\ln p(\square | \boldsymbol{\pi})] + E[\ln p(\boldsymbol{\pi})] + E[\ln p(\boldsymbol{\Sigma})] \\ & + E[\ln p(\mathbf{d} | \boldsymbol{\gamma})] + E[\ln p(\boldsymbol{\gamma})] + E[\ln p(\mathbf{C} | \boldsymbol{\sigma})] \\ & + E[\ln p(\boldsymbol{\sigma})] + E[\ln p(\boldsymbol{\tau} | \boldsymbol{\theta})] + E[\ln p(\boldsymbol{\theta})] \\ & + E[\ln p(\delta, \boldsymbol{\Lambda})] - E[\ln q(\mathbf{O})] - E[\ln q(\square)] \\ & - E[\ln q(\boldsymbol{\pi})] - E[\ln q(\boldsymbol{\Sigma})] - E[\ln q(\mathbf{d})] \\ & - E[\ln q(\boldsymbol{\gamma})] - E[\ln q(\mathbf{C})] - E[\ln q(\boldsymbol{\sigma})] \\ & - E[\ln q(\boldsymbol{\tau})] - E[\ln q(\boldsymbol{\theta})] - E[\ln q(\delta, \boldsymbol{\Lambda})] \end{aligned} \quad (32)$$

بهینه‌ی هر یک از این پارامترها، در هر گام از محاسبات، پس از تخمین تابع توزیع پسین $q(\mathbf{W})$ ، رابطه‌ی پارامتری باند پایین برای صفحه‌ی ۱ تشکیل می‌شود. سپس برای محاسبه‌ی مقادیر جدید پارامترها، مشتقات نسبی باند پایین نسبت به هر یک از پارامترها محاسبه می‌شود. محاسبه‌ی این مقادیر جدید در یک الگوریتم تکرارشونده و تا حصول هم‌گرایی، ادامه می‌یابد. جهت انجام محاسبات مربوط به صفحه‌ی ۲، روندی مشابه صفحه‌ی ۱ انجام می‌گردد. در ادامه، روابط نهایی محاسبه شده برای هر یک از دو صفحه‌ی مدل، ارائه شده است.

پس از انجام محاسبات مربوط به صفحه‌ی ۱، تابع توزیع احتمال پسین بردار وزن‌ها، به صورت روابط زیر به دست می‌آید:

$$q(\mathbf{W}) = \prod_{k=1}^K N(\mathbf{w}_k | \mathbf{m}_k, \mathbf{A}_k^{-1}) \quad (18)$$

$$\mathbf{A}_k = \alpha_{0k} \mathbf{I} + \beta \Phi_{X_k}^T \Phi_{X_k} \quad (19)$$

$$\mathbf{m}_k = \mathbf{A}_k^{-1} (\alpha_{0k} \mathbf{m}_{0k} + \beta \Phi_{X_k}^T \mathbf{U}_k) \quad (20)$$

که در این روابط، Φ_{X_k} ماتریس طرح^۱ مدل است. گام بعدی، محاسبه‌ی باند پایین صفحه‌ی ۱ (L_1) بوده که برای بررسی هم‌گرایی و همچنین تخمین مقدار پارامترها، ضروری است. برای باند پایین صفحه‌ی ۱ داریم:

$$L_1 = E[\ln p(\mathbf{U} | \mathbf{W})] + E[\ln p(\mathbf{W})] - E[\ln q(\mathbf{W})] \quad (21)$$

جهت محاسبه‌ی مقدار L_1 ، هر یک از عبارات سمت راست تساوی فوق، که نتایج نهایی آن‌ها در روابط (۲۲) تا (۲۴) درج شده است، مورد محاسبه قرار می‌گیرند:

$$\begin{aligned} E[\ln p(\mathbf{U} | \mathbf{W})] = & \sum_{k=1}^K \sum_{n=1}^N \left(\frac{1}{2} \ln \frac{\beta}{2\pi} - \frac{\beta}{2} (u_{kn}^2) \right. \\ & \left. - 2E[\mathbf{w}_k^T \phi_{X_k}(x_n) u_{kn}] + E[\mathbf{w}_k^T \phi_{X_k}(x_n) \phi_{X_k}^T(x_n) \mathbf{w}_k] \right) \end{aligned} \quad (22)$$

$$\begin{aligned} E[\ln p(\mathbf{W})] = & \frac{1}{2} \sum_{k=1}^K N_k \ln \frac{\alpha_{0k}}{2\pi} - \alpha_{0k} E[\mathbf{w}_k^T \mathbf{w}_k] \\ & + \alpha_{0k} E[\mathbf{w}_k^T] \mathbf{m}_{0k} - \alpha_{0k} \mathbf{m}_{0k}^T \mathbf{m}_{0k} \end{aligned} \quad (23)$$

$$E[\ln q(\mathbf{W})] = \sum_{k=1}^K E[\ln N(\mathbf{w}_k | \mathbf{m}_k, \mathbf{A}_k^{-1})] \quad (24)$$

به دلیل تعداد زیاد متغیرهای موجود در صفحه‌ی ۲، روابط به دست آمده در این تحقیق برای صفحه‌ی ۲، به طور خلاصه ذکر شده است. ابتدا، نتیجه‌ی محاسبه‌ی $q(\mathbf{O})$ ، که در قالب

^۱ Design Matrix

بهره گرفته شده است. روش اول، روش "کلاسیک + عبارت غیرمحملی"^۳ است که در تحقیقات سان و همکارانش ارائه شده است [۲۴،۲۳،۹]. هم‌چنین، از روش ارائه شده توسط بروکس^۴ و همکارانش (مبتنی بر یک روش جریان نوری وردشی) [۲۶،۲۵]، روش هورن-شانک^۵ و روش بلک-آناندون^۶ نیز جهت ارزیابی روش پیشنهادی، استفاده شده است [۲۴،۷]. جهت مقایسه‌ی روش پیشنهادی با روش‌های تطبیق نقطه، از الگوریتم تطبیق نقطه‌ی CPD، که توسط میرورنکو ارائه شده، استفاده شده است [۲۷]. لازم به ذکر است که روش CPD، در تحقیقات سال‌های اخیر نیز جهت تطبیق لکه‌های پروتئینی مورد استفاده قرار گرفته است [۲۸]. کلیه‌ی روش‌های مقایسه‌ای مذکور، در این تحقیق پیاده‌سازی شده و نتایج هر یک با اعمال روی تصاویر پایگاه داده، به دست آمده است. لازم به ذکر است که جهت بررسی و مقایسه‌ی عمل‌کرد روش‌ها در تطبیق لکه‌ها، از معیارهای خطای نقطه‌ی پایانی (EPE)، خطای زاویه‌ای (AE) و نرخ تطبیق (MR)، که به ترتیب در روابط (۳۳) تا (۳۵) ارائه شده‌اند، استفاده شده است [۱۰].

$$EPE = \sqrt{(u_M - u_T)^2 + (v_M - v_T)^2} \quad (33)$$

$$AE = \cos^{-1} \left(\frac{1 + u_M \times u_T + v_M \times v_T}{\sqrt{1 + u_M^2 + v_M^2} \sqrt{1 + u_T^2 + v_T^2}} \right) \quad (34)$$

$$MR = \frac{\text{Number of correct matched spots}}{\text{Total spot number}} \quad (35)$$

در این روابط، (u_M, v_M) بردار جابه‌جایی اندازه‌گیری شده و (u_T, v_T) بردار جابه‌جایی صحیح در نقاط مورد بررسی می‌باشند.

۳-۲-۱- نتایج اعمال الگوریتم تطبیق پیشنهادی

نمونه‌ای از خروجی‌های مراحل الگوریتم یک‌پارچه‌ی تطبیق پیشنهادی روی تصاویر واقعی و مصنوعی، به ترتیب در شکل‌های (۳) و (۴) نشان داده شده است. در شکل (۳)، تصاویر ردیف (الف)، تصاویر اصلی و ورودی، تصاویر ردیف (ب)، شرایط اولیه‌ی صفحه‌ی ۱، تصاویر ردیف (ج)، نتیجه‌ی تخمین پس از پایان محاسبات صفحه‌ی ۱، تصاویر ردیف (د)، شرایط اولیه‌ی صفحه‌ی ۲ و تصاویر ردیف (ه)، نتیجه‌ی انجام محاسبات صفحه‌ی ۲ را نشان می‌دهند. در شکل (۴)، تصاویر ردیف (الف)، تصاویر اصلی مصنوعی، تصاویر ردیف (ب)، نتیجه‌ی تخمین پس

که پس از محاسبه‌ی خودکار کلیه‌ی توزیع‌های پسین و امیدهای ریاضی هر یک، L_2 محاسبه می‌گردد. لازم به ذکر است که با توجه به حجم روابط، از درج جزئیات در رابطه‌ی (۳۲) صرف نظر شده است.

کلیه‌ی روابط اشاره شده در بالا، در هر گام محاسبه شده، با جای‌گذاری در رابطه‌ی باند پایین صفحه‌ی ۲ (رابطه‌ی ۳۲)، هم‌گرایی باند پایین محاسبه گشته و ارزیابی می‌شود. این محاسبات تکرار شونده، با وقوع هم‌گرایی در باند پایین، متوقف شده، آخرین مقادیر تخمین زده شده، به‌عنوان تخمین‌های نهایی ذخیره شده و مورد استفاده قرار می‌گیرند. لازم به ذکر است که کلیه‌ی این محاسبات، به صورت خودکار و بدون دخالت کاربر صورت می‌گیرد. در ادامه، نتایج حاصل از به کارگیری مدل ارائه شده در این تحقیق، مورد بررسی قرار گرفته است.

۳- نتایج

۳-۱- پایگاه داده و ابزار پردازش

در این تحقیق، جهت ارزیابی عمل‌کرد روش پیشنهادی، مجموعه‌ای از تصاویر واقعی و مصنوعی 2DGE، که در این تحقیق تهیه شده‌اند، مورد استفاده قرار گرفته است. پایگاه داده‌ی تصاویر واقعی، که در این تحقیق و در آزمایشگاه پروتئین-شیمی انستیتو پاستور ایران تهیه شده است، شامل ۱۸ تصویر از یک نمونه‌ی پایه‌ای اخذ شده از بافت پستان خوک‌چه‌ی هندی می‌باشد. تصاویر مصنوعی این تحقیق، از یک مجموعه‌ی ۳۶-تایی از تصاویر مصنوعی با تقریباً ۳۰۰۰ لکه‌ی مصنوعی تشکیل شده است. این تصاویر بر اساس روش ارائه شده در تحقیق شامخی و همکارانش تولید شده است [۲۲]. الگوریتم‌های ارائه شده در این تحقیق، در سیستم عامل اوبونتو ۱۲/۰۴، در محیط برنامه‌نویسی مکس^۱ نرم‌افزار متلب و به کمک زبان برنامه‌نویسی ++C و دستوره‌ای کتابخانه‌ی اینسایت تولکیت^۲، پیاده‌سازی شده است. هم‌چنین، از یک سیستم کامپیوتری با پردازنده‌ی اینتل ۳/۲ گیگاهرتز چهار هسته‌ای و حافظه‌ی رم ۴ گیگابایت، برای محاسبه‌ی زمان اجرای برنامه‌ها استفاده شده است.

۳-۲- نتایج الگوریتم پیشنهادی

در این تحقیق، جهت مقایسه و ارزیابی روش یک‌پارچه‌ای که برای تطبیق تصاویر 2DGE پیشنهاد شده، از ۵ روش مقایسه‌ای

^۴ Brox

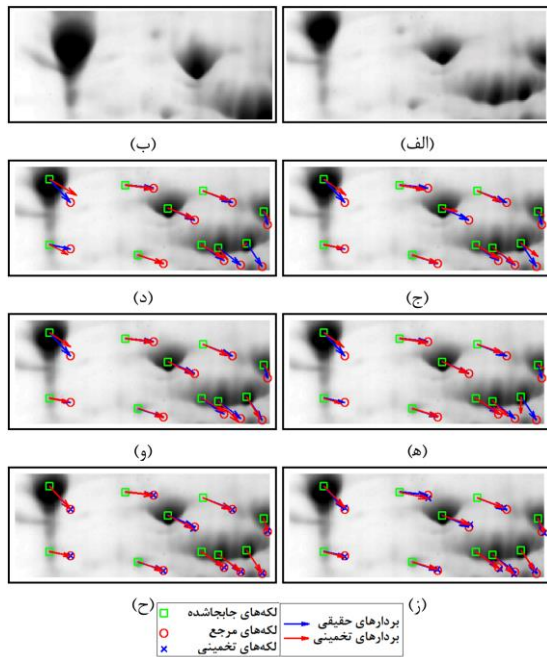
^۵ HS

^۶ Classic Black and Anandon (CBA)

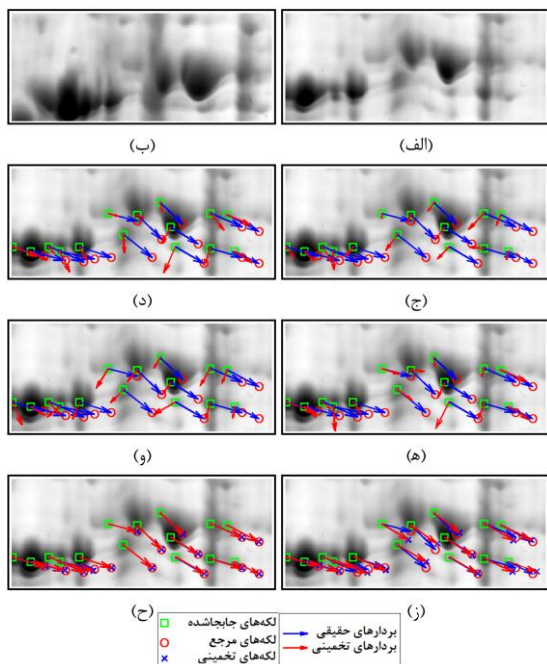
^۱ MEX

^۲ Insight Toolkit 4.0.0

^۳ CNLT



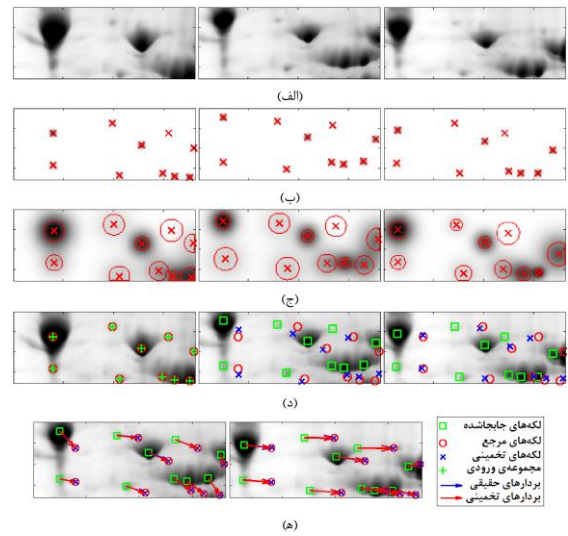
شکل (۵) - نتایج تطبیق، (الف) تصویر انتقال‌یافته، (ب) تصویر مرجع، (ج) روش بروکس، (د) روش HS، (ه) روش CBA، (و) روش CNLT، (ز) روش CPD، (ح) روش پیشنهادی



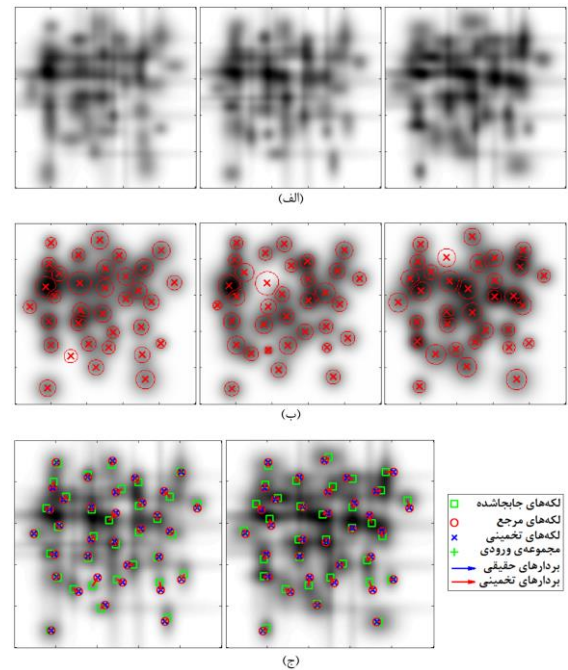
شکل (۶) - نتایج تطبیق، (الف) تصویر انتقال‌یافته، (ب) تصویر مرجع، (ج) روش بروکس، (د) روش HS، (ه) روش CBA، (و) روش CNLT، (ز) روش CPD، (ح) روش پیشنهادی

هم‌چنین، تصاویر شکل (۷)، نتیجه‌ی تطبیق لکه‌های تصویر سمت راست از ردیف (الف) شکل (۴) بر تصویر سمت چپ از همین ردیف، توسط روش پیشنهادی و روش‌های مقایسه‌ای می‌باشند. همان‌طور که مشاهده می‌شود، بردارهای تخمینی در نتایج به دست آمده از روش پیشنهادی، در مقایسه با روش‌های

از پایان محاسبات صفحه‌ی ۱ و تصاویر ردیف (ج)، نتیجه‌ی انجام محاسبات صفحه‌ی ۲ را نشان می‌دهند.



شکل (۳) - نتایج تطبیق تصاویر واقعی توسط روش پیشنهادی، (الف) مجموعه‌ای ۳-تایی از تصاویر واقعی، (ب) تصاویر تولید شده در مراحل الگوریتم، (ه) نتایج نهایی بردارهای جابه‌جایی (الف)



شکل (۴) - نتایج تطبیق تصاویر مصنوعی توسط روش پیشنهادی، (الف) مجموعه‌ای ۳-تایی از تصاویر مصنوعی، (ب) تصاویر تولید شده در مرحله‌ی ۱، (ج) نتایج نهایی بردارهای جابه‌جایی (الف)

نمونه‌هایی از نتایج حاصل از تطبیق تصاویر واقعی توسط روش‌های پیشنهادی و مقایسه‌ای، در شکل‌های (۵) و (۶) و نمونه‌ای از نتایج حاصل از تطبیق تصاویر مصنوعی، در شکل (۷) ارائه شده است. در هر یک از شکل‌های (۵) و (۶)، تصویر (الف)، تصویر انتقال‌یافته و تصویر (ب)، تصویر مرجع می‌باشد.

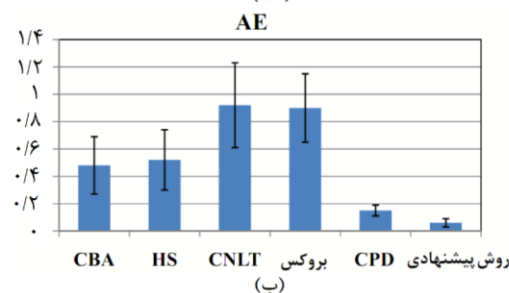
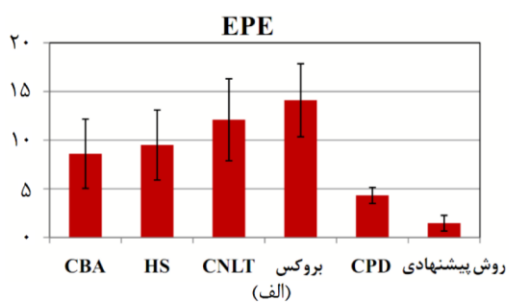
پیشنهادی در این تحقیق، که مبتنی بر مدل‌های احتمالاتی است، در مقایسه با روش‌های دیگر، بسیار کمتر می‌باشد. هم‌چنین، مقایسه‌ی انحراف معیار روش‌ها، نشان‌دهنده‌ی تغییرات زیاد در مقدار خطای بردار جابه‌جایی محاسبه شده در روش‌هایی است که مبتنی بر تطبیق نقطه نمی‌باشند. علاوه بر این، روش پیشنهادی در این تحقیق در مقایسه با سایر روش‌های مورد بررسی، دارای بیش‌ترین نرخ تطبیق صحیح در هر دو مجموعه‌ی تصاویر مصنوعی و واقعی است.

جدول (۱) - مقایسه‌ی کمی روش‌ها، در تطبیق لکه‌های واقعی

روش	EPE±Std	AE±Std	نرخ تطبیق
بلک-آناندون (CBA) (۲۰۰۸)	۸/۶۳±۳/۵۶	۰/۴۸±۰/۲۱	۶۳/۰
هورن-شانک (HS) (۲۰۰۸)	۹/۵۱±۳/۶	۰/۵۲±۰/۲۲	۶۳/۴
CNLT (۲۰۱۴)	۱۲/۳۸±۴/۲۱	۰/۹۲±۰/۳۱	۵۴/۱
بروکس (۲۰۰۴)	۱۴/۱۱±۳/۷۵	۰/۹۰±۰/۲۵	۵۰/۴
CPD (۲۰۱۰ و ۲۰۱۷)	۴/۳۲±۸/۲۳	۰/۰۸±۰/۰۴	۸۵/۸
روش پیشنهادی	۱/۴۶±۰/۸۱	۰/۰۵±۰/۰۲	۹۳/۵

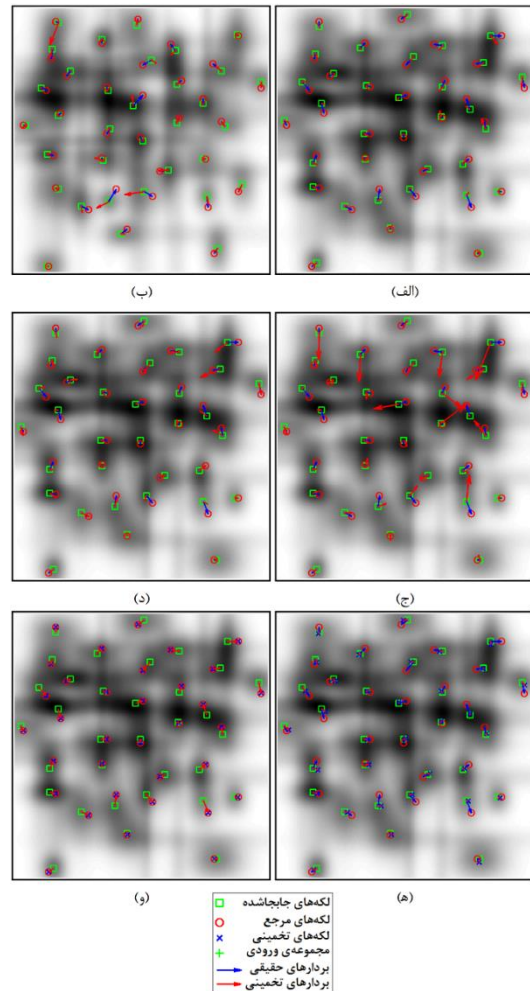
جدول (۲) - مقایسه‌ی کمی روش‌ها، در تطبیق لکه‌های مصنوعی

روش	EPE±Std	AE±Std	نرخ تطبیق
بلک-آناندون (CBA) (۲۰۰۸)	۵/۲۵±۲/۰	۰/۶۱±۰/۳۸	۶۷/۰
هورن-شانک (HS) (۲۰۰۸)	۵/۶±۱/۸۲	۰/۷۲±۰/۳۷	۷۰/۲
CNLT (۲۰۱۴)	۶/۱۳±۱/۲۲	۰/۷۴±۰/۳۴	۶۵/۱
بروکس (۲۰۰۴)	۶/۴۸±۰/۹۴	۰/۸۳±۰/۳۲	۶۴/۷
CPD (۲۰۱۰ و ۲۰۱۷)	۳/۵۷±۰/۸۰	۰/۵۱±۰/۳۱	۸۱/۷
روش پیشنهادی	۰/۹۰±۰/۳۱	۰/۱۳±۰/۰۶	۹۵/۹



شکل (۸) - (الف) خطای نقطه‌ی پایانی، (ب) خطای زاویه‌ای در تطبیق لکه‌های واقعی

مقایسه‌ای، مطابقت بهتری با بردارهای حقیقی دارند. هم‌چنین، بردارهای جابه‌جایی تخمین زده شده توسط روش‌های مبتنی بر تطبیق نقطه (روش پیشنهادی و CPD)، در مقایسه با چهار روش دیگر، مطابقت بیش‌تری با بردارهای حقیقی دارند.



شکل (۷) - نتایج تطبیق لکه‌های تصویر سمت راست از ردیف (الف) شکل (۴) بر تصویر سمت چپ از همین ردیف، توسط روش‌های: (الف) بروکس، (ب) HS، (ج) CBA، (د) CNLT، (ه) CPD، (و) روش پیشنهادی

۲-۲-۳- نتایج کمی (عددی) تطبیق تصاویر

در این تحقیق، جهت مقایسه و ارزیابی کمی عمل کرد روش‌های تطبیق، از معیارهای EPE، AE و MR، که پیش‌تر توضیح داده شدند، استفاده شده است. مقادیر این معیارها، که با اعمال کلیه‌ی روش‌ها در تطبیق مجموعه‌ی تصاویر واقعی و مصنوعی به دست آمده است، به ترتیب در جدول‌های (۱) و (۲) ارائه شده است. هم‌چنین، میانگین این معیارها و انحراف معیار آن‌ها، در نمودار شکل‌های (۸) و (۹) نشان داده شده است. با توجه به نتایج مذکور، مشاهده می‌شود که خطای میانگین روش

جدول (۳) - زمان مورد نیاز برای اجرای روش‌ها (s)

تصاویر واقعی	تصاویر مصنوعی	روش
۱۵۶۲/۹	۴۵/۲	بلک-آناندون (CBA) (۲۰۰۸)
۵۵۱/۳	۱۴/۵	هورن-شانک (HS) (۲۰۰۸)
۳۶۶۲/۹	۹۲/۱	CNLT (۲۰۱۴)
۲۶۴/۵	۷/۵	بروکس (۲۰۰۴)
۱۷۶/۸	۳/۲	CPD (۲۰۱۰ و ۲۰۱۷)
۳۸۰/۳	۷/۹	روش پیشنهادی

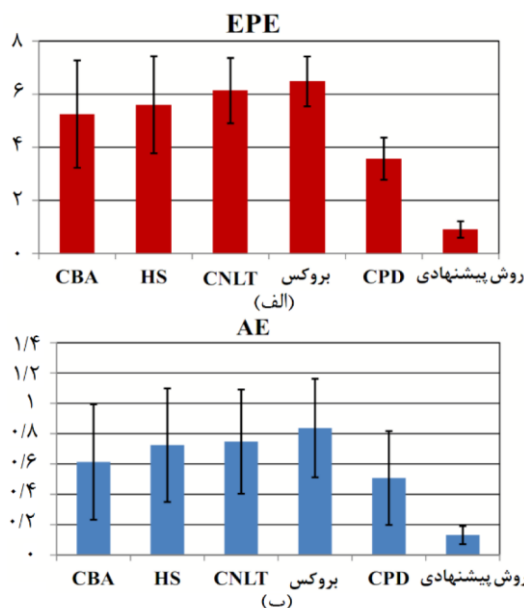
۴- نتیجه‌گیری و جمع‌بندی

در این تحقیق، جهت کاهش دخالت نیروی انسانی در تحلیل تصاویر 2DGE، مسالهی تطبیق لکه‌های پروتئینی، به صورت خودکار و در قالب یک مدل احتمالاتی، بیان شده است. در این مدل، ویژگی‌های هندسی تصاویر، مانند موقعیت مکانی لکه‌های پروتئینی، محاسبه شده، تخمین زده و تطبیق داده شده است. به دلیل گستردگی مسالهی تحقیق و در نتیجه، پیچیدگی مدل احتمالاتی بیان‌کننده‌ی آن، حل روابط مدل و تخمین پارامترهای آن به کمک روش بیز-وردشی انجام گرفت. جهت ارزیابی روش پیشنهادی تطبیق خودکار، نتایج حاصل از این روش، با مجموعه‌ای از روش‌های تطبیق تصاویر، مورد مقایسه قرار گرفته است. مقایسه‌ی معیارهای ارزیابی، هم‌چون خطای نقطه‌ی پایانی، خطای زاویه‌ای تطبیق و نرخ تطبیق، در تصاویر واقعی و مصنوعی، نشان‌گر برتری روش پیشنهادی در مقایسه با سایر روش‌های مورد بررسی است. هم‌چنین، مدت زمان اجرای الگوریتم خودکار تطبیق پیشنهادی، مورد مطالعه و مقایسه قرار گرفت. در حقیقت، با وجود این‌که مدت زمان اجرای الگوریتم خودکار پیشنهادی، در مقایسه با زمان مورد نیاز در سایر الگوریتم‌های موجود، به طور قابل ملاحظه‌ای کم‌تر است، برای فراهم ساختن شرایط، به منظور به کارگیری وسیع روش 2DGE در حوزه‌ی کلینیکی، زمان اجرای روش پیشنهادی باید کاهش داده شود.

جهت بهبود الگوریتم تطبیق ارائه شده در این تحقیق، پیشنهاد می‌گردد تا مطالعات بیش‌تری جهت ساده‌سازی مدل احتمالاتی ارائه شده، به ویژه در صفحه‌ی ۲، انجام شود تا محاسبات در زمانی کوتاه‌تر صورت گیرد.

۶- مراجع

- [1] U. K. Laemmli, "Cleavage of structural proteins during the assembly of the head of bacteriophage T4," *Nature*, vol. 227, no. 5259, pp. 680-685, 1970.



شکل (۹) - (الف) خطای نقطه‌ی پایانی، (ب) خطای زاویه‌ای در تطبیق لکه‌های مصنوعی

۳-۲-۳- بررسی مدت زمان اجرای الگوریتم‌ها

مدت زمان مورد نیاز برای پردازش تصاویر 2DGE، عاملی است که می‌تواند بر توسعه‌ی هر چه بیش‌تر کاربردهای کلینیکی روش 2DGE، تاثیرگذار باشد. زمان مورد نیاز برای اجرای الگوریتم‌های فعلی، به دلیل نظارت و دخالت نیروی انسانی در فرایندهای پردازش، بسیار طولانی است. به طوری که در الگوریتم‌های موجود، آماده‌سازی، ناحیه‌بندی و تطبیق یک مجموعه‌ی تصویر مربوط به یک نمونه‌ی پروتئینی، تا چند روز به طول می‌انجامد. از این‌رو در این بخش، الگوریتم خودکار پیشنهاد شده در این تحقیق، از دیدگاه زمان مورد نیاز برای اجرا، مورد بررسی و مقایسه قرار گرفته است. مدت زمان اجرای روش پیشنهادی و روش‌های مقایسه‌ای روی تصاویر مصنوعی و واقعی، به صورت میانگین برای هر تصویر، در جدول (۳) درج شده است. لازم به ذکر است که اعداد مربوط به روش‌های مقایسه‌ای، بدون در نظر گرفتن دخالت و نظارت کاربر، محاسبه شده و صرفاً زمان اجرای یک مرحله‌ای تطبیق را بیان می‌کنند. مشاهده می‌شود که مدت زمان تطبیق بدون نظارت، توسط روش پیشنهادی، در مقایسه با برخی از روش‌ها، کم‌تر و در مقایسه با برخی دیگر از روش‌ها، بیش‌تر است، که نشان‌دهنده‌ی ضرورت توجه بیش‌تر به موضوع زمان در روش‌های پیشنهادی می‌باشد. البته، با احتساب دخالت و نظارت نیروی انسانی، که در روش‌های مقایسه‌ای امری ضروری است، در مجموع، روش پیشنهادی از نظر زمان اجرا، عمل‌کرد مناسبی دارد.

- [16] M. J. Beal, "Variational algorithms for approximate Bayesian inference," University of London United Kingdom, 2003.
- [17] T. Jaakkola, "10 Tutorial on Variational Approximation Methods," *Advanced mean field methods: theory and practice*, p. 129, 2001.
- [18] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, "An introduction to variational methods for graphical models," *Machine learning*, vol. 37, no. 2, pp. 183-233, 1999.
- [19] A. Corduneanu and C. M. Bishop, "Variational Bayesian model selection for mixture distributions," in *Artificial intelligence and Statistics*, vol. 2001, pp. 27-34: Morgan Kaufmann Waltham, MA.
- [20] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 11, no. 6, pp. 567-585, 1989.
- [21] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Computer Vision and Image Understanding*, vol. 89, no. 2, pp. 114-141, 2003.
- [۲۲] س. شامخی، م. ح. میران بیگی، ب. آذریان، و ع. گویا، «مدل شکلی ناهمسانگرد جدید برای لکه های پروتئینی در تصاویر الکتروفورز ژل دو بعدی»، «مجموعه مقالات بیست و دومین کنفرانس مهندسی زیست پزشکی ایران، ۱۳۹۴»
- [23] C. Sun and X.-m. Wang, "Spot segmentation and verification based on improve marker controlled watershed transform," in *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on, 2010*, vol. 8, pp. 63-66.
- [24] Deqing Sun research page, School of Engineering and Applied Sciences, Harvard University. Available: <http://people.seas.harvard.edu/~dqsun/>
- [25] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," *Computer Vision-ECCV 2004*, pp. 25-36, 2004.
- [26] Binaries/Code, Department of Computer Science, Faculty of Engineering, Lehrstuhl für Mustererkennung und Bildverarbeitung. Available: <http://lmb.informatik.uni-freiburg.de/resources/binaries>
- [27] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 12, pp. 2262-2275, 2010.
- [28] S. Sudhakara, N. Patel, and V. Gadre, "Non-subsampled contourlet transform & coherent point drift based registration of 2D gel electrophoresis images," in *Convergence in Technology (I2CT), 2017 2nd International Conference for, 2017*, pp. 410-415: IEEE.
- [2] P. H. O'Farrell, "High resolution two-dimensional electrophoresis of proteins," *Journal of biological chemistry*, vol. 250, no. 10, pp. 4007-4021, 1975.
- [3] Automated detection and matching of spots in autoradiogram images of two-dimensional electrophoresis for high-speed genome scanning, 1997.
- [4] K. Rohr, P. Cathier, and S. Wörz, "Elastic registration of electrophoresis images using intensity information and point landmarks," *Pattern recognition*, vol. 37, no. 5, pp. 1035-1048, 2004.
- [5] G. Shi, T. Jiang, W. Zhu, B. Liu, and H. Zhao, "Alignment of two-dimensional electrophoresis gels," *Biochemical and biophysical research communications*, vol. 357, no. 2, pp. 427-432, 2007.
- [6] M. Rogers and J. Graham, "Robust and accurate registration of 2-D electrophoresis gels using point-matching," *IEEE transactions on image processing*, vol. 16, no. 3, pp. 624-635, 2007.
- [7] D. Sun, S. Roth, J. P. Lewis, and M. J. Black, "Learning Optical Flow," in *Computer Vision – ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part III*, D. Forsyth, P. Torr, and A. Zisserman, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 83-97.
- [8] D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, 2010*, pp. 2432-2439: IEEE.
- [9] D. Sun, S. Roth, and M. J. Black, "A quantitative analysis of current practices in optical flow estimation and the principles behind them," *International Journal of Computer Vision*, vol. 106, no. 2, pp. 115-137, 2014.
- [10] A. Rodriguez, C. Fernandez-Lozano, J. Dorado, and J. R. Rabunal, "Two-dimensional gel electrophoresis image registration using block-matching techniques and deformation models," *Analytical biochemistry*, vol. 454, pp. 53-59, 2014.
- [11] H.-M. Xin and Y. Zhu, "Spot Matching of 2-DE Images Using Distance, Intensity, and Pattern Information," in *2-D PAGE Map Analysis*: Springer, pp. 109-117, 2016.
- [12] B. Alizadeh Savareh, A. Bashiri, and M. Mostafavi, "Neighborhood matrix: A new idea in matching of two dimensional gel images," *Progress in Biological Sciences*, vol. 6, no. 2, pp. 129-137, 2017.
- [13] J. Ashburner and K. J. Friston, "Unified segmentation," *Neuroimage*, vol. 26, no. 3, pp. 839-851, 2005.
- [14] C. M. Bishop, "Pattern recognition and machine learning," Springer New York, 2006.
- [15] A. Gooya, K. M. Pohl, M. Bilello, L. Cirillo, G. Biros, E. R. Melhem and C. Davatzikos, "GLISTR: glioma image segmentation and registration," *IEEE transactions on medical imaging*, vol. 31, no. 10, pp. 1941-1954, 2012.