

تعیین ضایعه تومور مغزی با استفاده از الگوریتم k -mean و انترویی میانگین در تصاویر MRI

دکتر رحمان فرنوش^۱، فاطمه آقاگلی^{۲*}

۱- استادیار گروه آمار، دانشگاه علم و صنعت ایران f.farnoosh@iust.ac.ir

۲- کارشناسی ارشد آمار ریاضی، دانشکده ریاضی، دانشگاه علم و صنعت ایران f.aghagoly71@gmail.com

چکیده

در حال حاضر روش تصویربرداری MRI به دلیل وضوح و کیفیت بالا از مهم‌ترین ابزار تشخیص و ارزیابی تومورهای غیرقابل لمس مغز است. تجزیه و تحلیل تومورهای موجود در تصاویر MRI توسط متخصصین پزشکی و بر مبنای نواحی استخراج شده توسط الگوریتم‌های ناحیه بندی انجام می‌شود. در این مقاله هدف، استفاده از دو روش الگوریتم k -mean و انترویی میانگین جهت تشخیص تومور مغزی است. با بررسی نتایج به دست آمده، الگوریتم k -mean در مقایسه با انترویی میانگین می‌تواند تمام لبه‌های مربوط به بافت توده را در تصویر مشخص کند و همچنین آن را با اجزای بیشتری جداسازی و نمایش دهد در صورتی که انترویی میانگین لبه‌های بافت مورد نظر را مشخص نکرده؛ ولی اجزای اصلی مورد نظر را در تصویر حفظ کرده و تصویر را خلاصه تر نشان می‌دهد. در نهایت نتیجه می‌شود که انترویی میانگین قابلیت و انعطاف پذیری بیشتری دارد، اما دقت الگوریتم k -mean بیشتر است که بسته به نوع کارایی می‌توان از هر دو روش استفاده کرد.

کلمات کلیدی: تصاویر MRI، تقطیع تصاویر پزشکی، خوشه بندی k -mean، انترویی میانگین، تشخیص تومور مغزی.

۱. مقدمه

پردازش تصویر به منظور شناسایی موارد غیرطبیعی و ناهنجاری‌ها، تصاویر پزشکی را تحلیل و اطلاعات مفید را از آن‌ها استخراج می‌کند. از آنجایی که تشخیص معتبر و صحیح، آن‌هم در کوتاه‌ترین زمان ممکن دارای اهمیت است. مکانیسم‌های پردازش تصویر، به عنوان روش‌هایی ساده و غیرتهاجمی برای شناسایی سلول‌های سرطانی تشخیص اولیه را سرعت بخشیده و در نهایت شانس زنده ماندن بیماران سرطانی را افزایش می‌دهند [۱]. پردازش تصویر شامل چندین مرحله است که مهم‌ترین آن‌ها تقطیع است. تقطیع یک عملیات پایه برای آنالیز و تحلیل تصاویر است که پیکسل‌های تصویر را به نواحی غیر یکسان تقسیم می‌کند به طوری که این نواحی با یکدیگر اشتراکی نداشته باشند [۲].

برای تقطیع تصاویر روش‌های گوناگونی پیشنهاد شده است: روش‌های پیشنهادی را می‌توان به دودسته روش‌های مبتنی بر هیستوگرام و روش‌های مبتنی بر خوشه بندی تصاویر تقسیم بندی کرد. در روش‌های مبتنی بر هیستوگرام، تقسیم بندی تصاویر بر اساس توزیع پیکسل‌ها صورت می‌گیرد و هدف یافتن سطح آستانه‌ای مناسب برای اعمال به تصویر است. در روش‌های مبتنی بر خوشه بندی از شباهت‌ها و روابط موجود بین داده‌ها استفاده کرده و آن‌ها را به نحوی گروه بندی می‌کند که داده‌های داخل یک گروه یا خوشه شبیه به هم بوده و دارای بیشترین تفاوت با داده‌های خوشه‌های دیگر باشند [۳-۵]. در بخش دوم این مقاله روش انترویی میانگین مبتنی بر هیستوگرام و در بخش سوم الگوریتم

خوشه‌بندی k-mean برای تشخیص تومور موجود در یک تصویر MRI مغزی مورداستفاده قرار می‌گیرد و در پایان نیز به نتیجه‌گیری پرداخته می‌شود.

۲. روش انتروپی میانگین

مفهوم انتروپی به‌عنوان شاخه‌ای از پردازش تصویر با برآورد مقدار اطلاعات یک تصویر معرفی می‌شود. در این صورت می‌تواند از توزیع سطوح خاکستری تصویر محاسبه شود و یک تقطیع مناسب برای تصویر به دست آید. به دلیل اینکه مقادیر انتروپی برای تصویر با محاسبه مقدار احتمالات سطوح خاکستری به دست می‌آید. توابع پایه در این روش ماتریس احتمالات و سطوح خاکستری است. اگر احتمال هر سطح خاکستری r_k در تصویر p_k باشد. در این صورت می‌توان احتمال رخداد یک پیکسل از تصویر را به‌صورت زیر تعریف نمود:

$$p_k = \frac{n_k}{N \times M}, k = 0, \dots, L - 1 \quad (1)$$

که n_k تعداد پیکسل‌ها در سطح خاکستری k و L تعداد سطوح خاکستری را نشان می‌دهند. N و M نیز به ترتیب نشان‌دهنده‌ی تعداد سطر و ستون در تصویر هستند [۶].

انتروپی برای اولین بار توسط شانون* [۷] در سال ۱۹۴۸ و برای متغیرهای تصادفی گسسته تعریف شد که به انتروپی شانون معروف است. Pun [۸] برای اولین بار مفهوم شانون را برای تعریف انتروپی از یک تصویر به کاربرد با این فرض که یک تصویر به‌طور کامل تنها توسط هیستوگرام سطح خاکستری آن ارائه شده باشد. در سال‌های اخیر برای تعمیم انتروپی شانون برای متغیرهای تصادفی پیوسته چندین مدل پیشنهاد شده است [۹-۱۱]. انتروپی میانگین یک نوع استانداردسازی اعمال شده بر روی انتروپی شانون برای مطابقت دادن حالت‌های گسسته با پیوسته است، درحالی‌که آن بسیاری از ویژگی‌های اساسی انتروپی شانون را حفظ می‌کند. به همین دلیل انتروپی میانگین کلی‌تر از انتروپی شانون است در مفهوم اینکه تعریف آن شامل هر دو متغیر تصادفی گسسته و پیوسته می‌شود. کیتنه[†] و همکاران [۱۲] در سال ۲۰۱۶ با توجه به استدلال‌های جینز[‡] (۱۹۶۳، ۱۹۶۸) [۱۳، ۱۴]، عواد و امین[§] (۱۹۸۷) [۱۵] برای تعریف این انتروپی، انتروپی میانگین امروزی را به‌صورت زیر تعریف کردند:

$$A = \sum_{k=0}^{L-1} -p_k \frac{\log p_k}{p_k^2} = \sum_{k=0}^{L-1} \frac{-\log p_k}{p_k}, k = 0, \dots, L - 1 \quad (2)$$

و انتروپی میانگین برای تصویر به‌صورت زیر استدلال می‌شود:

$$A(X) = -E(\log \frac{f(X)}{a_X}), a_X = E(f(X)) = \sum_X f^2(X) \quad (3)$$

انتروپی میانگین تصویر، فاصله بین توزیع احتمال تصویر ورودی و تصویر خروجی را اندازه‌گیری می‌کند Li, Lee و Tam آستانه را به‌عنوان مینیمم فاصله در این انتروپی در نظر گرفتند [۱۲]؛ پس آستانه بهینه به نحوی انتخاب می‌شود که این فاصله حداقل شود. با توجه به اینکه ناحیه‌های مربوط به تومور، عمدتاً در ظاهر سفید هستند انتخاب صحیح و

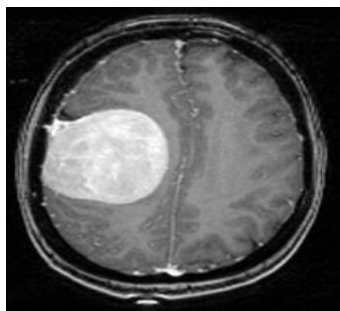
* Shannon

† Kittaneh

‡ Jaynes

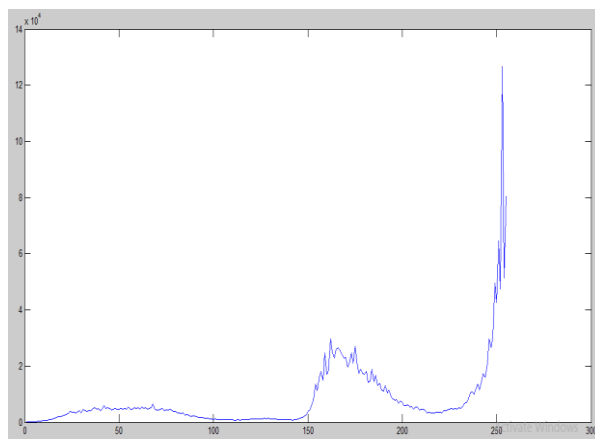
§ Awad and Ameen

منطقی برای تعیین مقدار آستانه به این صورت است که با استفاده از کمترین مقدار انترویی که به سطح خاکستری سفید نزدیک‌تر است اقدام به تعیین آستانه بهینه شود [۱۶-۱۸].
تصویر ورودی انتخاب‌شده یک تصویر MRI مغزی فردی بیمار با تعداد سطر ۳۳۹ و ستون ۳۶۹ است (شکل ۱).



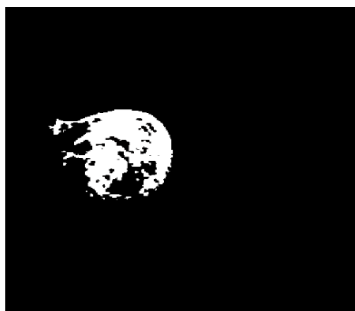
شکل ۱- تصویر MRI مغزی یک فرد بیمار

مقدار مینیمم انترویی میانگین با توجه به نمودار انترویی (شکل ۲) در سطح خاکستری ۲۱۴ رخ می‌دهد که برابر با مقدار ۰,۲۳۹۴ است.



شکل ۲- نمودار انترویی میانگین تصویر MRI مغزی فرد بیمار

با اعمال این مقدار آستانه بر روی تصویر ورودی تصویر خروجی به این صورت حاصل می‌شود (شکل ۳).



شکل ۳- تقطیع تصویر MRI مغزی فرد بیمار بر اساس مینیمم انترویی میانگین

۳. روش خوشه‌بندی k-mean

خوشه‌بندی یا آنالیز خوشه (Clustering) در آمار و یادگیری ماشین، یکی از شاخه‌های یادگیری غیر نظارتی است. این شاخه هنگامی استفاده می‌شود که ما به دنبال یافتن گروه‌هایی از داده‌های مشابه هستیم. بدون این‌که از قبل پیش‌بینی در مورد شباهت‌های موجود داشته باشیم.

پس خوشه مجموعه‌ای از اشیاء است که در آن اشیاء با یکدیگر مشابه بوده و با اشیاء موجود در خوشه‌های دیگر مشابه نیستند. الگوریتم k-mean یکی از پرکاربردترین الگوریتم‌های خوشه‌بندی است و برای اولین بار توسط جی. بی مک کوئین در سال ۱۹۶۷ معرفی شد. [۱۹]

الگوریتم زیر الگوریتم پایه برای این روش محسوب می‌شود:

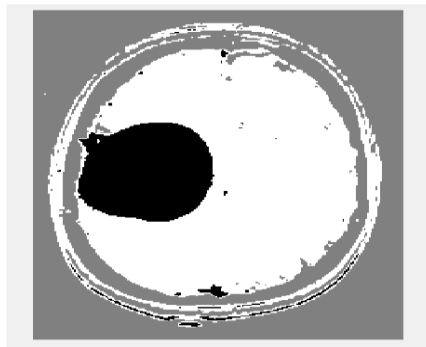
۱. در ابتدا K نقطه (k تعداد خوشه) به صورت تصادفی از میان n نقطه به‌عنوان مراکز خوشه‌ها انتخاب می‌شوند.

۲. هر نمونه داده ($n-k$ نقطه باقیمانده) به خوشه‌ای که مرکز آن خوشه کمترین فاصله تا آن داده را دارد، نسبت داده می‌شود.

۳. پس از تعلق تمام داده‌ها به یکی از خوشه‌ها برای هر خوشه یک نقطه جدید (میانگین نقاط متعلق به هر خوشه) به‌عنوان مرکز محاسبه می‌شود.

۴. مراحل ۲ و ۳ تکرار می‌شوند تا زمانی که دیگر هیچ تغییری در مراکز خوشه‌ها حاصل نشود [۲۰].

با اجرای الگوریتم خوشه‌بندی k-mean روی تصویر MRI مفروض داریم:



شکل ۳- تقطیع تصویر MRI مغزی فرد بیمار بر اساس روش خوشه‌بندی k-mean

از آنجاکه در خوشه‌بندی، به برچسب خوشه موردنظر دسترسی نداریم؛ لذا مجبوریم خوشه موردنظر خود را خودمان تعیین کنیم. هدف ما جدا کردن تومور در تصویر است؛ بنابراین خوشه مربوط به این ناحیه را پیدا می‌کنیم. همان‌طور که از تصویر مشخص است ناحیه سیاه‌رنگ، خوشه موردنظر ما است. باید توجه شود که رنگ‌بندی خوشه‌ها تصادفی است. ممکن است در تکرار بعدی الگوریتم k-mean رنگ خوشه‌ها باهم عوض شود. چون که مقادری اولیه خوشه‌ها در این الگوریتم در هر تکرار متفاوت است. همان‌طور که از تصویر مشخص است، خوشه مربوط به تومور در مقایسه با دو خوشه دیگر کمترین سطح را دارد. لذا تعداد عناصر هر خوشه را می‌شماریم، هر کدام تعداد عنصر کمتری داشته باشد خوشه موردنظر ما است. [۲۱].

۴. نتیجه‌گیری

با بررسی نتایج به‌دست‌آمده، الگوریتم k -mean در مقایسه با انتروپی میانگین توانسته تمام لبه‌های مربوط به بافت توده را در تصویر مشخص کند و همچنین آن را با اجزای بیشتری جداسازی و نمایش دهد در صورتی که انتروپی میانگین لبه‌های بافت موردنظر را مشخص نکرده ولی اجزای اصلی موردنظر را حفظ کرده و تصویر را خلاصه‌تر نشان می‌دهد. در نهایت می‌توان نتیجه گرفت که انتروپی میانگین قابلیت و انعطاف‌پذیری بیشتری دارد اما دقت الگوریتم k -mean بیشتر است که بسته به نوع کارایی می‌توان از هر دو روش استفاده کرد.

۵. مراجع

- [1] Chandra P, Noh H-B, Shim Y-B. (2013), "Cancer cell detection based on the interaction between an anticancer drug and cell membrane components". *Chemical Communications*. Vol. 49(19), pp. 1900-2.
- [2] Hadziavdic.V. (1999), "A comparative study of active contour models for boundary detection in brain images". Diploma Project of Faculty for Mathematical and Natural Sciences University of Tromso (URL: http://www.uib.no/med/avd/miapr/arvid/vedad_diploma.Pdf)
- [3] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. (2011), "Contour detection and hierarchical image segmentation". *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, pp. 898-916.
- [4] M. Gong, Y. Liang, J. Shi, W. Ma, and J. Ma. (2013), "Fuzzy c-means clustering with local information and kernel metric for image segmentation". *Image Processing, IEEE Transactions on*, vol.23, pp. 573-584.
- [5] A. Dirami, K. Hammouche, M. Diaf, and P. Siarry. (2013), "Fast multilevel thresholding for image segmentation through a multiphase level set method". *Signal Processing*, vol.93, pp. 139-153
- [6] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Prentice Hall, 2002, PP 220-235.
- [7] Shannon, C. E. (1948), "A Mathematical Theory of Communications," *Bell System Technical Journal*, 27, 379-423
- [8] T. Pun, "Entropic thresholding, a new approach," *Computer Graphics and Image Processing*, vol. 16, pp. 210-239, 1981.
- [9] A. Renyi, on measures of entropy and information, in: *Proceedings of the Fourth Berkeley Symposium on Math. Statist. Prob.* Vol. 1, 1960, University of California Press, Berkeley, 1961, pp. 547-561.
- [10] Tsallis, C. 1988. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* 52, 480-487.
- [11] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, A new method for gray-level picture thresholding using the entropy of the histogram, *Comput. Vision Graphics Image Process.* 29, 1985, 273-285.



- [12] Omar A. Kittaneh, Mohammad A. U. Khan, Muhammed Akbar & Husam A. Bayoud (2016) Average Entropy: A New Uncertainty Measure with Application to Image Segmentation, pp 18-24.
- [13] Jaynes, E. T. (1963), "Information Theory and Statistical Mechanics," in Statistical Physics, ed. K. Ford, New York: W. A. Benjamin, Inc. pp. 181–217.
- [14] Jaynes, E. T. (1968), "Prior Probabilities," IEEE Transactions on Systems Science and Cybernetics, SSC-4, 227–241.
- [15] Awad, M. and Ameen, J. (1987), "Application of Entropy to a Life-Time Model," IMA Journal of Mathematical Control & Information, 4, 143–147
- [16] Pinsker, M. S. (1964), Information and Information Stability of Random Variables and Processes, San Francisco, CA: Holden-Day (Originally published in Russian in 1960).
- [17] Li, C.H. Lee, C.K. 993. Minimum cross entropy thresholding. Pattern Recognition 26, 617-625.
- [18] Ferraro, M. Boccignone, G. and Caeili T. (1999), "On the Representation of Image Structures via Scale Space Entropy Conditions," IEEE Transaction on Pattern Analysis and Machine Intelligence, 21, 1199–1203.
- [19] Macqueen, J. (1967), "Some methods for classification and analysis of multivariate observations". In Proceedings of the 5th Berkeley symposium on mathematical statistics and probability, vol.1, pp. 281–297. Berkeley, CA: University of California Press.
- [20] Hartigan, J. (1975). "Clustering Algorithms". Toronto: JohnWiley & Sons.
- [21] C.A. Glasbey, and G.W. Horgan. (1995), "Image Analysis for the Biological Sciences", John Willey & Sons Pub.Comp.