

تشخیص ملانوما با یک مدل یادگیری عمیق

چکیده

دریافت: ۱۳۹۸/۱۱/۱۸ ویرایش: ۱۳۹۸/۱۱/۲۵ پذیرش: ۱۳۹۹/۰۳/۲۴ آنلاین: ۱۳۹۹/۰۳/۳۱

علی عامری*

گروه مهندسی پزشکی، دانشکده پزشکی،
دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران.

زمینه و هدف: سرطان پوست یکی از شایعترین سرطان‌ها و ملانوما (Melanoma) کشنده‌ترین نوع سرطان پوست می‌باشد. خال ملانوسیتیک (Melanocytic nevi) و ملانوما هر دو از ملانوسایت‌ها (سلول‌های تولیدکننده رنگدانه) به وجود می‌آیند، اما خال ملانوسیتیک خوش خیم و ملانوما بدخیم هستند. این مقاله یک مدل یادگیری عمیق (Deep learning) برای طبقه‌بندی (Classification) این دو ضایعه پوستی ارائه می‌کند.

روش بررسی: در این مطالعه تحلیلی که در بهمن ۱۳۹۸ در دانشگاه علوم پزشکی شهید بهشتی انجام شد، از مجموعه داده عکس‌های درماسکوبی (HAM10000, Human against machine with 10000 training images)، هزار تصویر خال ملانوسیتیک و هزار تصویر ملانوما استخراج گردید. از هر مورد، ۹۰۰ تصویر به شکل تصادفی برای آموزش سیستم انتخاب شدند و ۱۰۰ تصویر باقیمانده برای تست اختصاص داده شد. یک مدل یادگیری عمیق شبکه عصبی کانولوشنال (Convolutional neural network)، با استفاده از AlexNet (Krizhevsky et al., 2012) به‌عنوان مدل از پیش آموزش دیده شده (Pretrained)، به‌کار گرفته شد. در ابتدا این شبکه با ۱۸۰۰ تصویر آموزش داده شد و سپس عملکرد آن بر روی ۲۰۰ تصویر ارزیابی گردید.

یافته‌ها: مدل پیشنهادی به دقت ۹۳٪ (Accuracy) در طبقه‌بندی تصاویر به دو کلاس خوش‌خیم و بدخیم دست یافت. همچنین مساحت زیر منحنی (Receiver operating characteristic (ROC)، حساسیت ۹۴٪ (Sensitivity) و اختصاصیت ۹۲٪ (Specificity) به دست آمد. همچنین با تنظیم پارامتر آستانه طبقه‌بندی مدل، امکان افزایش حساسیت، به قیمت کاهش اختصاصیت وجود دارد و بالعکس.

نتیجه‌گیری: با توجه به دشواری تشخیص ملانوما حتی برای متخصصین با تجربه، یافته‌های این مطالعه، توانایی بالای یادگیری عمیق را در تشخیص سرطان پوست نشان می‌دهد.

کلمات کلیدی: یادگیری عمیق، ملانوما، خال رنگی، نئوپلاسم پوستی.

*نویسنده مسئول: تهران، ولنجک، بلوار دانشجو،
دانشگاه علوم پزشکی شهید بهشتی، دانشکده پزشکی.
تلفن: ۰۲۱-۲۲۴۳۹۹۱
E-mail: aliameri86@gmail.com

مقدمه

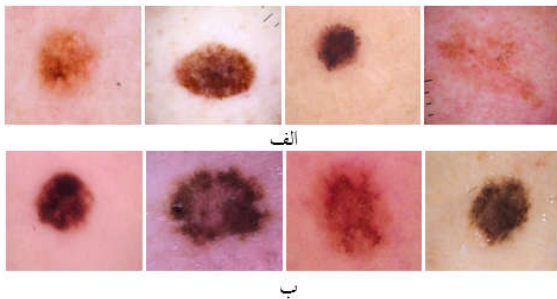
می‌گردد. هر سال در این کشور، ۱۰ هزار نفر بر اثر ابتلا به ملانوما جان خود را از دست می‌دهند. تشخیص زودهنگام برای این بیماری بسیار حیاتی است، به طوری که اگر در مراحل ابتدایی تشخیص داده شود احتمال مرگ و میر ۱٪ است ولی اگر دیر تشخیص داده شود احتمال مرگ به ۸۶٪ افزایش می‌یابد.^۲ به همین دلیل، سیستمی با

هر سال چهار، پنج میلیون نفر در آمریکا به سرطان پوست مبتلا می‌شوند.^۱ اگرچه ملانوما ۵٪ موارد سرطان پوست در این کشور را شامل می‌شود، ولی منجر به ۷۵٪ موارد مرگ در اثر سرطان پوست

دانشگاه کوئینزلند استرالیا گردآوری شده است.^{۱۰} این تصاویر شامل انواع مختلفی از ضایعات پوستی هستند و تشخیص بیشتر این تصاویر به وسیله پاتولوژی انجام شده و بقیه تصاویر با توافق نظر متخصصین، یا میکروسکوپ کانفوکال تشخیص داده شده‌اند. در مطالعه حاضر، با استفاده از تصاویر ملانوما و خال ملانوسیتیک در این مجموعه داده، یک شبکه یادگیری عمیق ارایه شده است.

روش بررسی

در این مطالعه تحلیلی که در بهمن ۱۳۹۸ در دانشگاه علوم پزشکی شهید بهشتی انجام شد، از مجموعه تصاویر درماسکوپی (Human against machine with 10000 training images, HAM10000) هزار تصویر ملانوما و هزار تصویر خال ملانوسیتیک مورد استفاده قرار گرفتند. در شکل ۱، چهار تصویر به عنوان نمونه از تصاویر ملانوما و خال ملانوسیتیک نشان داده شده‌اند. در ابتدا از هر مورد، به طور تصادفی ۹۰۰ تصویر به عنوان داده آموزشی انتخاب شدند و ۱۰۰ تصویر باقیمانده به عنوان داده تست در نظر گرفته شدند. یک مدل یادگیری عمیق شبکه عصبی کانولوشنال (Convolutional neural network) برای طبقه‌بندی تصاویر به کار گرفته شد. به منظور بهبود عملکرد شبکه، از روش انتقال یادگیری (Transfer learning) استفاده شد. در این روش، از یک شبکه که برای یک کاربرد مرتبط دیگر آموزش داده شده به عنوان شبکه پیش آموزش داده شده استفاده می‌شود، یعنی مقادیر اولیه وزن‌های فیلترهای لایه‌های کانولوشنال،

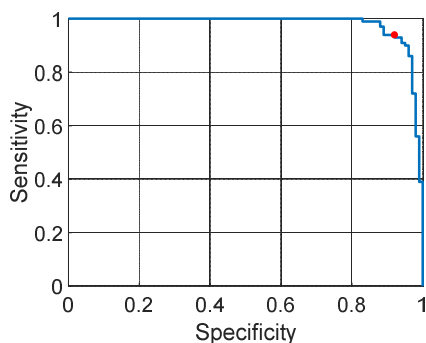


شکل ۱: چهار تصویر به عنوان نمونه از تصاویر (الف) ملانوما و (ب) خال ملانوسیتیک نشان داده شده‌اند.

توانایی تشخیص ملانوما از روی تصویر بسیار مفید خواهد بود و می‌تواند از آمار مرگ‌ومیر بکاهد. ملانوما در سلول‌های ملانوسیت ایجاد می‌شود. این سلول‌ها ساخت ملانین که رنگدانه پوستی است را بر عهده دارند. خال ملانوسیتی که یک ضایعه خوش خیم است هم در این سلول‌ها شکل می‌گیرد.^۳ با توجه به خطر بالای ملانوما، تشخیص این دو ضایعه از هم اهمیت بسزایی دارد.

تا سال ۲۰۱۶، مطالعاتی که سیستم‌های مبتنی بر کامپیوتر برای تشخیص سرطان پوست ارایه دادند، از روش‌های سنتی یادگیری ماشین (Machine learning) استفاده می‌کردند، به طوری که ابتدا تصویر ضایعه مورد نظر را با تکنیک‌هایی از پوست پیرامون جداسازی کرده و سپس ویژگی (Feature) های خاصی (مانند شکل، بافت و رنگ ضایعه) را تعریف و از آن استخراج می‌کردند.^{۴-۶} در پایان هم، این ویژگی‌ها را به عنوان ورودی به یک طبقه‌بند (Classifier) برای تشخیص نوع ضایعه می‌دادند. اما از سال ۲۰۱۶، شبکه‌های یادگیری عمیق (Deep learning) به تدریج وارد گستره‌ی تشخیص سرطان پوست شد. این شبکه‌ها از چند سال پیش از آن در کاربردهایی مانند تشخیص اشیاء در تصویر، تحول بزرگی ایجاد کرده بودند.

امروزه، یادگیری عمیق در حوزه‌های مختلف از جمله در مهندسی پزشکی کاربرد گسترده‌ای یافته است.^{۷-۹} مزیت مهم روش‌های یادگیری عمیق این است که احتیاجی به جداسازی ضایعه و استخراج ویژگی از آن را ندارند و می‌توانند از روی تصویر خام، نوع ضایعه را تشخیص دهند. یکی از مشکلاتی که برای آموزش این سیستم‌ها وجود دارد، نبود مجموعه داده‌های با حجم زیاد می‌باشد. در سال ۲۰۱۷، پژوهشگرانی از دانشگاه استنفورد، یک مطالعه بر روی ۱۳۰ هزار تصویر ضایعات پوستی انجام دادند که در آن یک شبکه یادگیری عمیق بر روی بخشی از تصاویر آموزش داده شد و سپس عملکرد مدل بر روی تصاویر دیگر ارزیابی گردید.^۲ همچنین از ۲۱ متخصص پوست هم خواسته شد که در مورد بخشی از تصاویر تست، نظر تشخیصی ارایه دهند. یافته‌ها نشان داد که دقت تشخیص روش یادگیری عمیق در سطح دقت تشخیص متخصصین بود. متأسفانه داده‌های این مطالعه، در دسترس عموم قرار ندارد. در سال ۲۰۱۸، یک مجموعه داده با ۱۰ هزار تصویر درماسکوپی ارایه شد که طی مدت ۲۰ سال در دانشگاه پزشکی وین اتریش و همچنین



شکل ۲: منحنی ROC عملکرد مدل پیشنهادی بر روی داده تست. نقطه قرمز، آستانه اطمینان طبقه‌بندی ۰/۵ را نشان می‌دهد.

است. نتایج ارزیابی عملکرد مدل برای طبقه‌بندی ۲۰۰ تصویر تست (۱۰۰ مورد از هر کلاس) نشان داد که سیستم با دقت بالایی قادر به تمایز این دو نوع ضایعه پوستی از یکدیگر می‌باشد. در یک مطالعه که از ۱۵۷ متخصص پوست در آلمان (۳۰٪ آن‌ها متخصصین با تجربه بالا بودند)، برای طبقه‌بندی ملانوما و خال ملانوسیتیک بر روی ۱۰۰ تصویر درماتوسکوپی نظرخواهی شده بود، به‌طور میانگین توانستند طبقه‌بندی با $AUC=0.67$ انجام دهند.^{۱۱} هرچند دقت تشخیص، به‌شدت به مجموعه داده ارتباط دارد، اما این نتیجه گویای دشواری تشخیص ملانوما برای متخصصین می‌باشد. بنابراین، دقت بالای به‌دست آمده ($AUC=0.98$) در روش یادگیری عمیق در مطالعه حاضر، بیانگر اهمیت این روش هم برای تشخیص در زمانی‌که پزشک در دسترس نیست و هم برای تشخیص کمکی برای متخصص می‌باشد.

همانطور که در منحنی ROC در شکل ۲ مشخص است، با تغییر آستانه اطمینان طبقه‌بندی می‌توان حساسیت و اختصاصیت طبقه‌بندی را تغییر داد. در واقع این امر، بیانگر یک Trade-off می‌باشد که با افزایش یکی، دیگری کاهش می‌یابد. بنابراین، کاربر می‌تواند با توجه به این‌که کدامیک به چه میزان اهمیت بیشتری دارند، آستانه بهینه را انتخاب کند. برای مثال اگر بخواهد تا حد امکان، ملانوما را به اشتباه خال ملانوسیتیک طبقه‌بندی نکند، باید اهمیت بیشتری به حساسیت داده، به قیمت اینکه اختصاصیت پایین‌تری داشته باشد. برای این

به‌جای مقادیر تصادفی، مقادیر وزن‌های این شبکه قرار داده می‌شوند. در این‌جا از شبکه AlexNet (Krizhevsky et al., 2012) برای انتقال یادگیری استفاده شد، چرا که این شبکه برای طبقه‌بندی تصویر اشیاء آموزش دیده که مرتبط با کاربرد طبقه‌بندی ضایعات پوستی می‌باشد.^{۱۱} سه لایه آخر AlexNet یعنی آخرین لایه Fully-connected، لایه‌های Softmax و Classification جایگزین شدند تا شبکه برای طبقه‌بندی با دو کلاس مناسب شود. برای آموزش شبکه از الگوریتم Stochastic gradient descent with momentum (SGDM) و تعداد ۴۰ اپیک استفاده شد. همچنین نرخ یادگیری اولیه 0.0001 و اندازه Minibatch ۳۰ در نظر گرفته شد. شبکه با استفاده از یک کامپیوتر با پردازنده گرافیکی GeForce™ GTX 1050 (NVIDIA Corporation, Santa Clara, CA, USA) و نرم‌افزار MATLAB®, version 2018Ra (Mathworks Inc., Natick, MA, USA) مورد نیاز برای AlexNet یعنی 227×227 تغییر داده شد. برای به‌دست آوردن کیفیت طبقه‌بندی، چهار معیار محاسبه شدند: دقت طبقه‌بندی (Classification accuracy)، حساسیت (Sensitivity)، اختصاصیت (Specificity) و مساحت زیر منحنی ROC (Area under the curve, AUC). با فرض اینکه منظور از کلاس مثبت، ملانوما و منظور از کلاس منفی، خال ملانوسیتیک است.

یافته‌ها

منحنی ROC عملکرد مدل پیشنهادی بر روی داده تست در شکل ۲ نشان داده شده است که در آن $AUC=0.98$ به‌دست آمد. اگر آستانه اطمینان طبقه‌بندی را ۰/۵ بگیریم (نقطه قرمز بر روی منحنی)، شبکه با دقت ۹۳٪ ملانوما و خال ملانوسیت را طبقه‌بندی می‌کند. همچنین حساسیت ۹۴٪ و اختصاصیت ۹۲٪ به‌دست می‌آید.

بحث

در این مطالعه، برای طبقه‌بندی تصاویر ضایعه‌های پوستی به دو کلاس ملانوما و خال ملانوسیتیک، یک مدل یادگیری عمیق پیشنهاد شد که با ۹۰۰ تصویر درماتوسکوپی از هر کلاس آموزش دیده

هوشمند، امکان تشخیص زودهنگام سرطان پوست را میسر خواهند کرد که می‌تواند نقش مهمی در کاهش مرگ‌ومیر بر اثر سرطان داشته باشد.

از یافته‌های این مطالعه می‌توان نتیجه گرفت که مدل‌های یادگیری عمیق، توانایی بالایی در تشخیص ملانوما از روی عکس‌های ضایعات پوستی دارند. مزیت اصلی یادگیری عمیق نسبت به روش‌های پیشین در این است که بدون هیچ پیش‌پردازی می‌تواند از روی تصویر خام، نوع ضایعه را تشخیص دهد. طبیعی است با افزایش حجم داده‌های آموزشی، قدرت تشخیص این مدل‌ها نیز افزایش می‌یابد.

سپاسگزاری: این مطالعه با حمایت دانشگاه علوم پزشکی شهید بهشتی انجام شده است.

هدف باید آستانه اطمینان را کاهش داد (حرکت نقطه قرمز به سمت چپ روی منحنی ROC). شکل ۲ نشان می‌دهد که برای اینکه مدل، حتی یک مورد ملانوما را هم به اشتباه خال ملانوسیتیک تشخیص ندهد (حساسیت ۱۰۰٪)، مدل همچنان می‌تواند عملکرد قابل‌قبولی با اختصاصیت ۸۳٪ داشته باشد، یعنی فقط ۱۷٪ موارد خال ملانوسیتیک را به اشتباه ملانوما تشخیص می‌دهد.

در این مطالعه، مدل یادگیری عمیق با ۹۰۰ تصویر در هر کلاس آموزش دید. مدل‌های یادگیری عمیق هم مانند انسان، با تعداد و نوع بیشتر داده‌های آموزشی، عملکرد بهتری خواهند داشت. در سال‌های اخیر، با افزایش کیفیت دوربین‌های گوشی‌های همراه، امکان ارتقاء تصویربرداری از ضایعات پوستی فراهم شده است. بنابراین در آینده، مدل‌های یادگیری عمیق، در قالب نرم‌افزار بر روی گوشی‌های

References

1. Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the U.S. population, 2012. *JAMA Dermatol* 2015;151(10):1081-6.
2. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542(7639):115-8.
3. Xu YG, Aylward JL, Swanson AM, Spiegelman VS, Vanness ER, Teng JM, et al. Nonmelanoma skin cancers: basal cell and squamous cell carcinomas. In: Niederhuber JE, Armitage JO, Doroshow JH, Kastan MB, Tepper JE, et al, editors. *Abeloff's Clinical Oncology: 6th ed.* Philadelphia, PA: Elsevier; 2020. P. 1052-73. e8.
4. Masood A, Al-Jumaily AA. Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms. *Int J Biomed Imaging* 2013;2013:323268.
5. Rosado B, Menzies S, Harbauer A, Pehamberger H, Wolff K, Binder M, Kittler H. Accuracy of computer diagnosis of melanoma: a quantitative meta-analysis. *Arch Dermatol* 2003;139(3):361-7; discussion 366.
6. Burroni M, Corona R, Dell'Eva G, Sera F, Bono R, Puddu P, Perotti R, Nobile F, Andreassi L, Rubegni P. Melanoma computer-aided diagnosis: reliability and feasibility study. *Clin Cancer Res* 2004;10(6):1881-6.
7. Ameri A. EMG-based wrist gesture recognition using a convolutional neural network. *Tehran Univ Med J* 2019;77(7):434-9.
8. Ameri A, Akhaee MA, Scheme E, Englehart K. Real-time, simultaneous myoelectric control using a convolutional neural network. *PLoS One* 2018;13(9):e0203835.
9. Ameri A, Akhaee MA, Scheme E, Englehart K. Regression convolutional neural network for improved simultaneous EMG control. *J Neural Eng* 2019;16(3):036015.
10. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* 2018;5:180161.
11. Krizhevsky, A, Sutskever, I, Hinton, GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inform Process Syst* 2012;25(2):1097-105.
12. Brinker TJ, Hekler A, Hauschild A, Berking C, Schilling B, Enk AH, et al. Comparing artificial intelligence algorithms to 157 German dermatologists: the melanoma classification benchmark. *Eur J Cancer* 2019;111:30-7.

Melanoma detection with a deep learning model

Ali Ameri Ph.D. *

Department of Biomedical
Engineering, School of Medicine,
Shahid Beheshti University of
Medical Sciences, Tehran, Iran.

* Corresponding author: School of
Medicine, Shahid Beheshti University of
Medical Sciences, Daneshjoo Blvd.,
Velenjak Ave., Tehran, Iran.
Tel: +98-21-22439941
E-mail: aliameri86@gmail.com

Abstract

Received: 07 Feb. 2020 Revised: 14 Feb. 2020 Accepted: 13 Jun. 2020 Available online: 20 Jun. 2020

Background: Skin cancer is one of the most common forms of cancer in the world and melanoma is the deadliest type of skin cancer. Both melanoma and melanocytic nevi begin in melanocytes (cells that produce melanin). However, melanocytic nevi are benign whereas melanoma is malignant. This work proposes a deep learning model for classification of these two lesions.

Methods: In this analytic study, the database of HAM10000 (human against machine with 10000 training images) dermoscopy images, 1000 melanocytic nevi and 1000 melanoma images were employed, where in each category 900 images were selected randomly and were designated as the training set. The remaining 100 images in each category were considered as the test set. A deep learning convolutional neural network (CNN) was deployed with AlexNet (Krizhevsky et al., 2012) as a pretrained model. The network was trained with 1800 dermoscope images and subsequently was validated with 200 test images. The proposed method removes the need for cumbersome tasks of lesion segmentation and feature extraction. Instead, the CNN can automatically learn and extract useful features from the raw images. Therefore, no image preprocessing is required. Study was conducted at Shahid Beheshti University of Medical Sciences, Tehran, Iran from January to February, 2020.

Results: The proposed model achieved an area under the receiver operating characteristic (ROC) curve of 0.98. Using a confidence score threshold of 0.5, a classification accuracy of 93%, sensitivity of 94%, and specificity of 92% was attained. The user can adjust the threshold to change the model performance according to preference. For example, if sensitivity is the main concern; i.e. false negative is to be avoided, then the threshold must be reduced to improve sensitivity at the cost of specificity. The ROC curve shows that to achieve sensitivity of 100%, specificity is decreased to 83%.

Conclusion: The results show the strength of convolutional neural networks in melanoma detection in dermoscopy images. The proposed method can be deployed to help dermatologists in identifying melanoma. It can also be implemented for self diagnosis of photographs taken from skin lesions. This may facilitate early detection of melanoma, and hence substantially reduce the mortality chance of this dangerous malignancy.

Keywords: deep learning, melanoma, pigmented nevus, skin neoplasms.