Viewpoint

LesionMap: A Method and Tool for the Semantic Annotation of Dermatological Lesions for Documentation and Machine Learning

Bell Raj Eapen[1], MD, MSc; Norm Archer[1], PhD; Kamran Sartipi[2], PhD

[1]Information Systems, McMaster University, Hamilton, ON, Canada

[2]Department of Computer Science, East Carolina University, Greenville, NC, United States

Corresponding Author:
Bell Raj Eapen, MD, MSc
Information Systems
McMaster University
DSB - 211
1280 Main Street West
Hamilton, ON, L8S 4L8
Canada
Phone: 1 905 525 9140 ext 26392
Email: eapenbp@mcmaster.ca

Abstract

Diagnosis and follow-up of patients in dermatology rely on visual cues. Documentation of skin lesions in dermatology is time-consuming and inaccurate. Digital photography is resource-intensive, difficult to standardize, and has privacy concerns. We propose a simple method—LesionMap—and an electronic health software tool—LesionMapper—for semantically annotating dermatological lesions on a body wireframe. We discuss how the type, distribution, and progression of lesions can be represented in a standardized way. The tool is an open-source JavaScript package that can be integrated into web-based electronic medical records. We believe that LesionMapper will facilitate documentation in dermatology that can be used for machine learning in a privacy-preserving manner.

Introduction

Documenting the origin, distribution, and nature of dermatological lesions in a textual form is inefficient and imprecise. Dermatologists often document the images of the patient or draw the lesions on a body wireframe for later reference. Digital photography for clinical documentation is time-consuming and resource intensive to capture, organize, and maintain [1]. Additionally, there is a growing privacy-related concern over the use of these images [2].

Capturing a detailed account of dermatological lesions in a privacy-preserving way is becoming increasingly important in the era of machine learning and artificial intelligence (AI). Documentation in electronic medical records (EMRs) requires a simple and efficient tool that fits into the clinical workflow. There is a growing need for a standardized methodology and an annotation schema to facilitate the capture of rich data related to dermatological conditions for machine learning. For this, we propose a simple method—LesionMap (LM)—and an electronic health (eHealth) software tool—LesionMapper (LMR)—that fits into the clinical workflow.

The sharing of clinical images between dermatologists for learning purposes is common, and most images are published with the consent of the patient [2]. However, increasingly, social media platforms are used for the easy sharing of such resources, with the associated implications on privacy [3]. Machine learning and AI applications need access to a large volume of data to build machine learning models for clinical decision support. Emerging techniques in machine learning and AI such as convolutional neural networks (CNN) and transfer learning [4] have several applications in dermatology [5]. Interestingly, some computer-vision methods can be applied to machine-generated images in addition to digital images [6].

In this paper, we describe common skin lesions, the semantic annotation methodology (LM), and a software tool (LMR) that can be used for semantic annotations. The tool is designed as an extensible software library (JavaScript) that can be

incorporated into web-based EMRs. We briefly describe two such integrations with open-source EMRs—OpenMRS and OSCAR EMR.

Classification of Skin Lesions

Dermatologists use numerous descriptive terms to identify and describe skin lesions [7]. Flat skin lesions that are small are called macules, and when they exceed 1 cm in size, they are called patches. An elevated dome-shaped lesion is called a nodule, whereas a flat elevated lesion is called a plaque. Small fluid-filled lesions are called vesicles, and if they exceed 1 cm, they are called bullae. If vesicles are filled with pus instead of clear fluid, they are called pustules.

Scales refer to a thickened outer layer of skin while the crust is a liquid debris. An ulcer is an irregularly shaped, deep loss of skin, and if it is superficial it is called an erosion. Atrophy is a thinning of skin, and a fissure is a linear cleft. Necrosis is dead skin tissue, and the scar is the replacement of lost skin by connective tissue. Localized hemorrhage into the skin is called purpura, and petechiae, when the hemorrhagic lesions are small.

The color of the lesion can provide diagnostic cues, along with the shape, arrangement, and distribution. Discoid and annular are terms used to describe the shape. The distribution can be grouped, discrete, linear, serpiginous, reticular, generalized, symmetrical, or photodistributed. The size, location, and severity are also important. Although this is not an exhaustive list of dermatological descriptions, the most common conditions are included here. Discrepancies in the terminology of dermatological lesions exist in the literature [8]. LM does not attempt to formalize the ontology, but proposes a pragmatic standard using the iconographic method.

Iconographic Representations of Skin Lesions

Most descriptive terms used in dermatology can be represented by iconographic images representative of the lesion or feature. The use of iconography in clinical documentation has been demonstrated in the context of pain [9]. The type of lesion can be easily represented by icons due to their visual similarity. The list of icons can be supplemented with custom icons for representing descriptive characteristics, such as the site of onset. LMR provides a set of icons for representing visual and nonvisual characteristics of common skin lesions and additional icons for descriptive characteristics (see Figure 1A and B).

Figure 1. The LesionMapper interface: (A,B) icons for descriptive characteristics; (C) large nodule; (D) ulcer on lateral side; (E) multiple discrete plaques; (F) Christmas tree pattern in pityriasis rosea; (G) vitiligo patch showing variation in opacity.
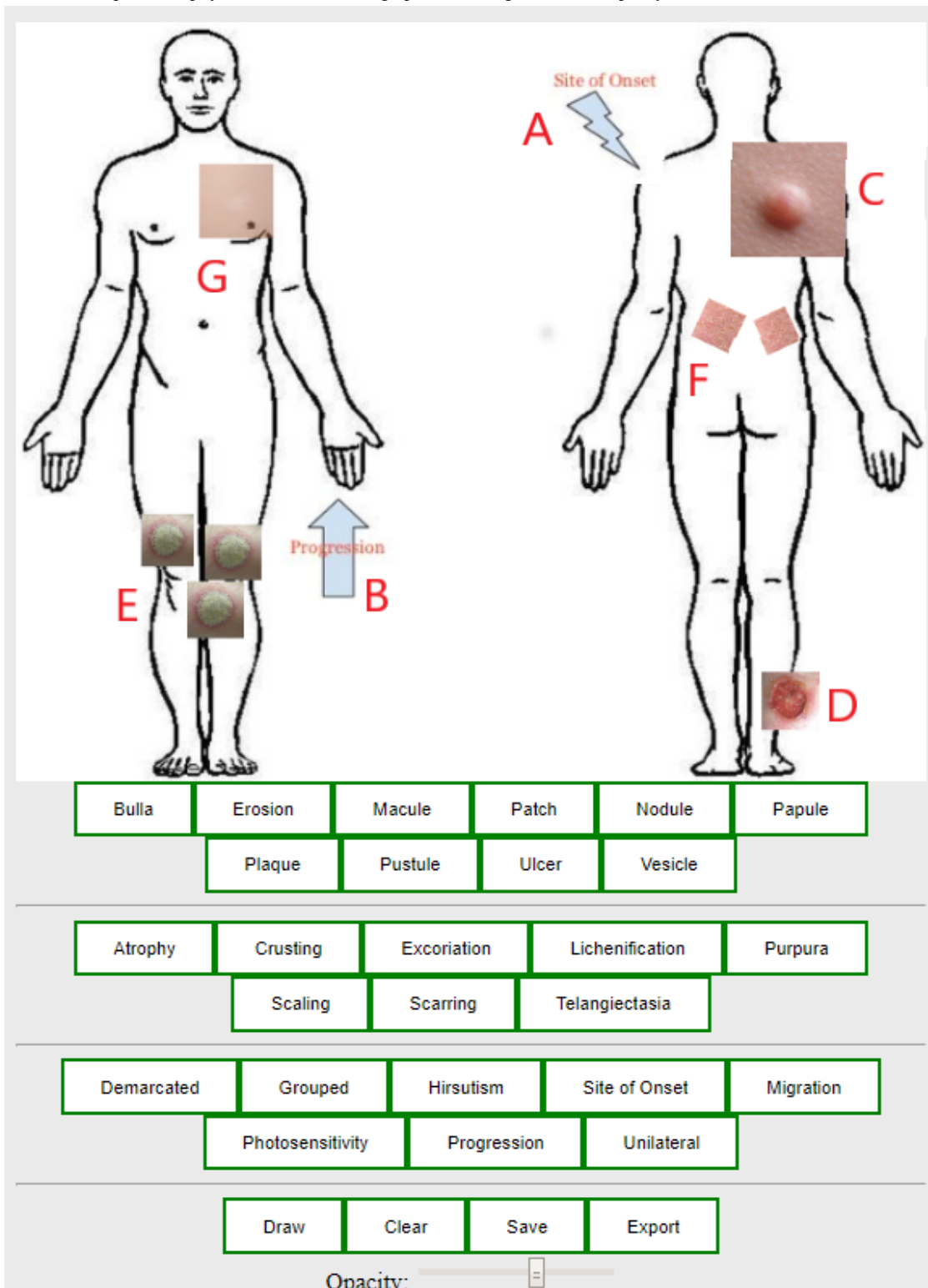
In addition to the type of lesions, there are five other characteristics of each icon that can be changed: size, position, number, orientation, and opacity. Additional information pertaining to the lesion can be encoded using the following characteristics:

- The size of the icon can be used to indicate the average size of the lesion in conditions where lesion size points toward a diagnosis or a particular subtype of the primary diagnosis.

For example, the size of the plaques can be a differentiating feature for small-plaque and large-plaque parapsoriasis. The original size of the icon, when placed on the LM, can be used for comparison (see the large nodule in Figure 1C).

- The position of icon placement indicates the distribution of the lesions. The front and back of the body are depicted in the LM. The lateral view is not included to simplify the interface. To represent lateral distribution, the icons can be

placed in the corresponding edge of the wireframe with an overlap of 50% (see the ulcer on the legs in Figure 1D).

- Multiple icons of the same type can be used to represent discrete lesions, and a single large icon can be used to represent confluent distribution (see discrete plaques in Figure 1E).
- The orientation can be used to indicate a pathognomonic distribution, such as the Christmas tree pattern in pityriasis rosea (see Figure 1F).
- The opacity of the lesion can be used to indicate the severity of the presentation. For example, it can be used to represent the degree of depigmentation in a vitiligo patch or the severity of contact dermatitis (see Figure 1G).

Mapping lesions consistently and accurately requires a tool that supports the various functions described above. In addition, from a design perspective, the tool should have the capability to integrate with other health information systems and EMRs.

LesionMapper

LMR is a prototype implementation of the LM method described above. We adopted the design science principles of Hevner et al [10] for information systems to design LMR. We searched the literature for similar approaches and available tools to address the problem of lesional documentation. Based on the success of similar approaches (Pain-QuILT for annotating pain [9]), we chose iconography as the method and standardized it based on our domain expertise in dermatology. Thereafter, we distinguished some of the easily identifiable characteristics of icons that can be programmatically controlled, such as size, orientation, and transparency. Subsequently, we converged on a popular framework (VueJS JavaScript framework [11]) for implementation. We designed the artifact adopting a modular pattern—as a JavaScript package shared as open source (see the GitHub repository [12])—that can be incorporated into web-based EMRs. LMR provides buttons to add various icons to the canvas. These icons can be independently moved and resized. The opacity and orientation can also be independently modified. The LM can be exported as an image or as a JavaScript Object Notation (JSON) string. LMR supports freehand drawing in the canvas to represent features that are not represented by icons though machine interpretation of the freehand drawing is challenging.

Next, we describe the integration of LMR into two open-source EMRs—OpenMRS and OSCAR.

Integration With Electronic Medical Records

The modular design helps in the integration of LMR into web-based EMRs. The prototype is created using the VueJS JavaScript framework following the Universal Module Definition (UMD) pattern [13] that can be imported by different module loaders into other browser-based applications. The icons are converted into Base64 strings and included in the JavaScript files.

Open Medical Records System (OpenMRS) is an open-source, Java-based EMR for developing countries with a modular and extensible architecture [14]. OpenMRS supports the Open Web Apps (OWA) specifications that make it possible to design external applications that extend the core functions. The OWA communicates with OpenMRS using REST APIs (representational state transfer application programming interfaces), a software architectural style used for creating Web services, and is embedded in the same server instance. OpenMRS has a custom concept dictionary that helps map data points to a uniform terminology. Nontextual data such as images are stored as "complex concepts" outside the relational database. LMR can easily fit into an OWA design pattern, and the exported LM images can be stored as complex observations in the patient record. We have a prototype integration that can be used as an example [15].

OSCAR EMR is a web-based EMR system initially developed for primary care practitioners in Canada. OSCAR EMR has a complex data model, and additional data points are supported by an electronic form (eForm) module that stores data as key-value pairs [16]. eForms do not support images or other nontextual data. The ability of LMR to save LMs as a JSON string makes the integration of the LMR module into eForms possible.

Machine Learning Applications

Dermatological diseases have diverse presentations, with skin type and skin color adding to this variation. Some of these diseases involve hair, nails, and mucous membranes in addition to the skin. Traditional computer-vision algorithms such as convolutional neural networks (CNNs) and other variants of neural networks have limited application when there are many decision alternatives [17]. Hence, AI algorithms have had limited application in dermatology except in problems associated with classification (eg, the presence or absence of cancer) [18]. Such algorithms can classify only a given lesion rather than the patient as a whole (ie, a lesion is cancerous vs patient has cancer). Although few CNN-based image search algorithms have proven to be useful, AI algorithms for diagnostic decision making in clinical dermatology lag far behind areas such as radiology [17].

Text analytics and natural language processing (NLP) can be more useful than image analytics when the decision alternatives are numerous, as in dermatology. Multimodal approaches where an image is combined with metadata have shown promise [19]. Machine learning models built using LMs—especially the models created using the JSON representation—resemble text more than an image. Some relevant metadata such as the position and distribution of the lesions, which are difficult to be captured in text and hard to precisely decipher with NLP, are implicitly captured in LMs. The icons represent ontological concepts from dermatology and can map to any standard terminology system [7]. We posit that LMs are semantically rich enough to be used for machine learning applications. Machine learning models from LMs are likely to be more "explainable" than traditional black box image search algorithms [20]. The implicit metadata

captured by LMs can supplement regular digital images, leading to better machine learning models.

Advantages and Limitations

LMs may save time for busy practitioners while capturing the type, distribution, and characteristics of the lesion; these data can be used to assess clinical progress. The LM exported as JSON resembles a markup language amenable to data mining and machine learning methods [21]. LMs are portable and can be easily and safely exchanged without privacy concerns.

LMR can export LMs as images. These images can be used as a proxy for patient images in computer vision–based applications. Computer vision has been successfully applied to identify metabolic defects from gene expression maps [22]. MNIST (Modified National Institute of Standards and Technology database), a dataset widely used in machine learning, consists of images of handwritten digits [23].

It is widely accepted that machine learning can reinforce some health care disparities in dermatology [24]. Skin color is a significant background noise that needs to be accounted for in any machine learning model. It is possible that some of the existing models are biased toward particular skin types that predominate in the training data set. Such models tend to be less sensitive in making predictions on different skin types [25]. The LM, being of a neutral skin type, is not affected by such bias.

LMs, however, do not capture all the features, both explainable and unexplainable, captured by a digital image. Hence, LMs are not useful in scenarios where accurate and sensitive extraction of features from an image is important for prediction. For example, LMs are not appropriate for skin cancer classification [5] and mole mapping [26]. LMs do not support annotating dermatopathology images [27]; they are also not applicable for dermatoscopic images that rely on pixel-level analysis [28]. Examination findings such as fluctuation, consistency, and tenderness are not represented by icons at present to keep the interface simple. More icons can be added if the user community requires them.

LMR and the LM method have not been clinically tested. The integration of LMR into existing EMRs may be difficult. Despite its anticipated ease of use, the actual impact of LMR on physician workflow, if any, needs to be investigated further.

Discussion

The skin is the largest organ in the human body, and as such, skin conditions are commonly encountered in any health care practice. Although dermatology is a specialization within clinical medicine, 50% of skin conditions are assessed and documented by nondermatologists [29].

There is no universal standard for pictographic documentation of the type, distribution, progression, and severity of lesions in dermatology, as in dentistry [30] and ophthalmology [31]. The LM standardizes visual representation using icons that can be extended to accommodate different use cases in clinical and cosmetic dermatology. The simplicity of the mapping rules facilitates use by nondermatologists in the skincare industry; LMs are also semantically rich enough to capture most relevant information about a skin condition with minimal effort.

Image analytics in dermatology is not as popular, as it is in visually oriented medical specialties such as radiology and pathology; the exception is the field of skin cancer diagnostics. This is because of the privacy concerns associated with dermatological images and the difficulties in standardizing image capture. The LM is not a replacement for a digital image of the lesion. However, some of the diagnostic aspects that are difficult to be captured in images, such as distribution and progression, can be useful for machine learning applications, especially when combined with the textual representation of a patient's history. Such multimodal approaches mimic the clinical workflow more so than CNN-based search algorithms [32]. New computer-vision algorithms are proving to be capable of learning from computer-generated images [22]. We believe that LMs can be similarly used with computer-vision methods. Finally, we urge the open-source community to help us improve LMR and potential users to report issues on the repository [12] so that we can fix them. We will work on a 3D wireframe for better accuracy, and we welcome other feature requests from the user community.

Conflicts of Interest

None declared.

References

1. Aspres N, Egerton IB, Lim AC, Shumack SP. Imaging the skin. Australas J Dermatol 2003 Feb;44(1):19-27. [doi: 10.1046/j.1440-0960.2003.00632.x]
2. Kunde L, McMeniman E, Parker M. Clinical photography in dermatology: Ethical and medico-legal considerations in the age of digital and smartphone technology. Australasian Journal of Dermatology 2013 May 29;54(3):192-197. [doi: 10.1111/ajd.12063]
3. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P & T 2014 Jul;39(7):491-520 [FREE Full text] [Medline: 25083128]
4. Shin H, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans. Med. Imaging 2016 May;35(5):1285-1298. [doi: 10.1109/tmi.2016.2528162]

5. 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 Jan 25;542(7639):115-118. [doi: 10.1038/nature21056]

6. Levine MD, Nazif AM. Dynamic Measurement of Computer Generated Image Segmentations. IEEE Trans Pattern Anal Mach Intell 1985 Mar;PAMI-7(2):155-164. [doi: 10.1109/tpami.1985.4767640]

7. Eapen BR. ONTODerm--a domain ontology for dermatology. Dermatol Online J 2008 Jun 15;14(6):16. [Medline: 18713597]

8. Cardili RN, Roselino AM. Elementary lesions in dermatological semiology: literature review. An Bras Dermatol 2016 Oct;91(5):629-633. [doi: 10.1590/abd1806-4841.20164931]

9. Lalloo C, Kumbhare D, Stinson JN, Henry JL. Pain-QuILT: clinical feasibility of a web-based visual pain assessment tool in adults with chronic pain. J Med Internet Res 2014 May 12;16(5):e127 [FREE Full text] [doi: 10.2196/jmir.3292] [Medline: 24819478]

10. Hevner AR, March ST, Park J, Ram S. Design Science in Information Systems Research. MIS Quarterly 2004;28(1):75. [doi: 10.2307/25148625]

11. Street M, Passaglia A, Halliday P. Complete Vue. In: Complete Vue.js 2 Web Development: Practical Guide To Building End-to-end Web Development Solutions With Vue.js 2. Birmingham: Packt Publishing; 2018.

12. Eapen BR. GitHub. 2020. LesionMapper - Vue component for standardized mapping of lesions in dermatology URL: https://github.com/dermatologist/lesion-mapper [accessed 2020-01-16]

13. Rich H. Dismantling the barriers to entry. Queue 2015;13(5):37 [FREE Full text] [doi: 10.1145/2742580.2742813]

14. Mamlin BW, Biondich PG, Wolfe BA, Fraser H, Jazayeri D, Allen C, et al. Cooking up an open source emr for developing countries: OpenMRS, a recipe for successful collaboration. In: American Medical Informatics Association. 2006 Presented at: AMIA Annual Symposium Proceedings; 2006; Washington p. 529.

15. Eapen BR. OpenMRS. 2017. Dermatology LesionMapper URL: https://addons.openmrs.org/show/org.openmrs.module.dermatology-lesionmap [accessed 2020-01-18]

16. Safadi H, Chan D, Dawes M, Roper M, Faraj S. Open-source health information technology: A case study of electronic medical records. Health Policy and Technology 2015 Mar;4(1):14-28. [doi: 10.1016/j.hlpt.2014.10.011]

17. Li C, Shen C, Xue K, Shen X, Jing Y, Wang Z, et al. Artificial intelligence in dermatology. Chinese Medical Journal 2019;132(17):2017-2020. [doi: 10.1097/cm9.0000000000000372]

18. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. BMC Med Res Methodol 2019 Mar 19;19(1). [doi: 10.1186/s12874-019-0681-4]

19. Wang H, Wang Y, Liang C, Li Y. Assessment of Deep Learning Using Nonimaging Information and Sequential Medical Records to Develop a Prediction Model for Nonmelanoma Skin Cancer. JAMA Dermatol 2019 Nov 01;155(11):1277. [doi: 10.1001/jamadermatol.2019.2335]

20. Andreas H, Chris B, Constantinos SP, Douglas BK. What do we need to build explainable AI systems for the medical domain? arXiv.org 2017 [FREE Full text]

21. Guazzelli A, Lin W, Jena T. PMML in action: unleashing the power of open standards for data mining and predictive analytics. California: Createspace Independent Publishing Platform; 2020.

22. Zhang J, Naik HS, Assefa T, Sarkar S, Reddy RVC, Singh A, et al. Computer vision and machine learning for robust phenotyping in genome-wide studies. Sci Rep 2017 Mar 8;7(1). [doi: 10.1038/srep44048]

23. Li Deng. The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web]. IEEE Signal Process. Mag 2012 Nov;29(6):141-142. [doi: 10.1109/msp.2012.2211477]

24. Adamson AS, Smith A. Machine Learning and Health Care Disparities in Dermatology. JAMA Dermatol 2018 Nov 01;154(11):1247. [doi: 10.1001/jamadermatol.2018.2348]

25. Kamulegeya LH, Okello M, Bwanika JM, Musinguzi D, Lubega W, Rusoke D, et al. Using artificial intelligence on dermatology conditions in uganda: A case for diversity in training data sets for machine learning. bioRxiv.org 2019 [FREE Full text] [doi: 10.1101/826057]

26. Berk-Krauss J, Polsky D, Stein JA. Mole Mapping for Management of Pigmented Skin Lesions. Dermatologic Clinics 2017 Oct;35(4):439-445. [doi: 10.1016/j.det.2017.06.004]

27. Toro P, Corredor G, Romero E, Arias V. A visualization, navigation, and annotation system for dermatopathology training. In: International Society for Optics and Photonics. 2018 Presented at: 14th International Symposium on Medical Information Processing and Analysis, volume 10975; 2018; Mazatlán, Mexico. [doi: 10.1117/12.2511638]

28. Ferreira B, Barata C, Marques JS. What is the role of annotations in the detection of dermoscopic structures? In: Iberian Conference on Pattern Recognition and Image Analysis. 2019 Presented at: 9th Iberian Conference, IbPRIA 2019; 2019; Madrid, Spain p. 3-11. [doi: 10.1007/978-3-030-31321-0_1]

29. Ahiarah A, Fox C, Servoss T. Brief intervention to improve diagnosis and treatment knowledge of skin disorders by family medicine residents. Fam Med 2007;39(10):720-723 [FREE Full text] [Medline: 17987414]

30. Wu M, Koenig L, Lynch J, Wirtz T. Spatially-oriented EMR for Dental Surgery. In: American Medical Informatics Association. 2006 Presented at: AMIA Annual Symposium Proceedings; 2006; Washington p. 1147.

31. Rajkomar A, Kannan A, Chen K, Vardoulakis L, Chou K, Cui C, et al. Automatically Charting Symptoms From Patient-Physician Conversations Using Machine Learning. JAMA Intern Med 2019 Jun 01;179(6):836. [doi: 10.1001/jamainternmed.2018.8558]

32.   Baltrusaitis T, Ahuja C, Morency L. Multimodal Machine Learning: A Survey and Taxonomy. IEEE Trans Pattern Anal
       Mach Intell 2019 Feb 1;41(2):423-443. [doi: 10.1109/tpami.2018.2798607]

Abbreviations

AI: artificial intelligence
API: application programming interface
CNN: convolutional neural network
eForm: electronic form
eHealth: electronic health
EMR: electronic medical record
JSON: JavaScript Object Notation
LM: LesionMap
LMR: LesionMapper
MNIST: Modified National Institute of Standards and Technology database
NLP: natural language processing
OWA: Open Web Apps
REST: representational state transfer
UMD: Universal Module Definition

XSL•FO
RenderX