

DETEKSI DINI KANKER KULIT MENGGUNAKAN K-NN DAN CONVOLUTIONAL NEURAL NETWORK

Teresia R. Savera¹, Winsya H. Suryawan², Agung W. Setiawan^{*3}

^{1,2}Program Studi Sarjana Teknik Elektro, Sekolah Teknik Elektro dan Informatika ITB

³Teknik Biomedika, Sekolah Teknik Elektro dan Informatika ITB

Email: ¹teresiasavera@gmail.com, ²winsyahesap@gmail.com, ³awsetiawan@stei.itb.ac.id

*Penulis Korespondensi

(Naskah masuk: 24 Oktober 2019, diterima untuk diterbitkan: 10 Februari 2020)

Abstrak

Kanker kulit adalah salah satu jenis kanker yang dapat menyebabkan kematian sehingga diperlukan sebuah aplikasi perangkat lunak yang dapat digunakan untuk membantu melakukan deteksi dini kanker kulit dengan mudah. Sehingga diharapkan deteksi dini kanker kulit dapat terdeteksi lebih cepat. Pada penelitian ini terdapat dua metode yang digunakan untuk melakukan deteksi dini kanker kulit yaitu deteksi dengan klasifikasi secara regresi dan *artificial neural network* dengan arsitektur *convolutional neural network*. Akurasi yang diperoleh dengan menggunakan klasifikasi secara regresi adalah sebesar 75%. Sementara, akurasi deteksi yang didapatkan dengan menggunakan *convolutional neural network* adalah sebesar 76%. Hasil yang diperoleh dari kedua metoda ini masih dapat ditingkatkan pada penelitian lanjutan, yaitu dengan cara melakukan prapengolahan pada set data citra yang digunakan. Sehingga set data yang digunakan memiliki tingkat pencahayaan, sudut (pengambilan), serta ukuran citra yang sama. Apabila tersedia sumber daya komputasi yang besar, akan dilakukan penambahan jumlah citra yang digunakan, baik itu sebagai set data latihan maupun uji.

Kata kunci: *benign, convolutional neural networks, kanker kulit, melanoma, regresi*

EARLY DETECTION OF SKIN CANCER USING K-NN AND CONVOLUTIONAL NEURAL NETWORK

Abstract

Skin cancer is one type of cancer that can cause death for many people. Because of this, an application is needed to easily detect skin cancer early that the cancer can be handled with more quickly. In this study there were two methods used to detect skin cancer, namely detection by regression classification and detection by classifying using artificial neural networks with network convolutional architecture. Detection with regression classification gives an accuracy of 75%. While detection using convolutional neural networks gives an accuracy of 76%. These proposed early detection systems can be improved to increase the accuracy. For further development, several image processing techniques will be applied, such as contrast enhancement and color equalization. For future works, if there is more computational resource, more images can be used as dataset and implement the deep learning algorithm to improve the accuracy.

Keywords: *benign, convolutional neural networks, melanoma, regression, skin cancer*

1. PENDAHULUAN

Penyakit kanker adalah salah satu penyakit utama yang dapat menyebabkan kematian di dunia. Menurut *World Health Organization*, sebanyak 160.000 orang mengidap kanker kulit setiap tahun di dunia. Australia merupakan salah satu negara yang memiliki jumlah penderita kanker kulit tertinggi, empat kali lebih tinggi dibandingkan Amerika Serikat. Berdasarkan data dari *Australian Bureau of Statistics*, dari semua orang Australia yang mengidap kanker, sebanyak 32,6% adalah penderita kanker

kulit. Pada tahun 2012, terdapat 971.279 kasus kanker kulit dan 2.162 diantaranya menyebabkan kematian. Sementara, menurut data dari *Centers for Disease Control and Prevention*, pada tahun 2015, di Amerika Serikat terdapat 80.442 kasus melanoma kulit, dan 8.885 orang meninggal. Di Indonesia, penderita kanker kulit terbilang relatif sedikit apabila dibandingkan dengan ketiga negara tersebut. Namun demikian, kanker kulit ini harus mendapat perhatian sedini mungkin, karena tidak hanya dapat menyebabkan kecacatan yang dapat merusak

penampilan, tetapi juga dapat mengakibatkan kematian apabila sudah masuk ke stadium lanjut.

Dikutip dari Pusat Penelitian Sumber Daya dan Pelayanan Kesehatan Kemenkes RI, pada tahun 2012, diperkirakan terdapat 14 juta kasus baru kanker dan 8,2 juta kematian akibat kanker di dunia. Dan di Indonesia, menurut data Badan Registrasi Kanker, Ikatan Ahli Patologi Indonesia, dari 1.530 kasus kanker kulit, kasus terbanyak adalah karsinoma sel basal, yaitu sebesar 39,93%. Oleh karena itu, diperlukan suatu aplikasi perangkat lunak yang dapat digunakan untuk membantu melakukan deteksi dini kanker kulit dengan mudah. Diharapkan, dengan adanya aplikasi perangkat lunak ini dapat membantu masyarakat untuk mendeteksi kelainan kulit, apakah masuk ke dalam kategori kanker atau tidak. Sehingga, masyarakat tidak terlambat mengetahui jenis kelainan kulit, yang kemudian dapat berkonsultasi lebih lanjut ke dokter, terutama dokter spesialis kulit dan kelamin.

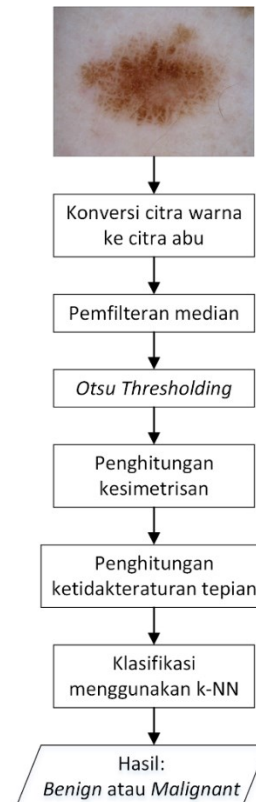
Deteksi dini kanker kulit berbasis pengolahan citra (Zortea, dkk., 2014.) dapat dilakukan dengan menggunakan ekstraksi ciri (Mehta & Shah, 2016), (Kaya, dkk., 2016), (Navarro, Escudero-Viñolo, & Bescós, 2018), (Pathan, Prabhu, & Siddalingaswamy, 2018) maupun *machine learning* (Arroyo & Zapirain, 2014), (Marchetti, dkk., 2018). Metode ekstraksi ciri digunakan untuk mendapatkan nilai dari parameter *Assymetry*, *Border Irregularity*, *Color*, dan *Diameter* dari suatu citra kanker kulit. Keempat parameter ini dikenal dengan parameter ABCD yang umum digunakan untuk mendeteksi kanker kulit menggunakan citra (Jain & Pise, 2015), (Rastgoo, dkk., 2015), (Kasmi & Mokrani, 2016).

Pada penelitian ini, terdapat dua usulan metode klasifikasi yang digunakan untuk melakukan deteksi kanker kulit berbasis pengolahan citra digital, yaitu klasifikasi secara regresi dan *artificial neural network* (Dalila, dkk., 2017) menggunakan *Convolutional Neural Network* (CNN) (Nasr-Esfahani, dkk., 2016). Implementasi klasifikasi secara regresi dilakukan dengan menggunakan perangkat lunak MATLAB. Terdapat dua *toolbox* yang digunakan, yaitu *image processing* dan *Statistics and Machine Learning*. Sedangkan metode CNN diimplementasikan dengan menggunakan bahasa pemrograman *Python*.

2. METODE

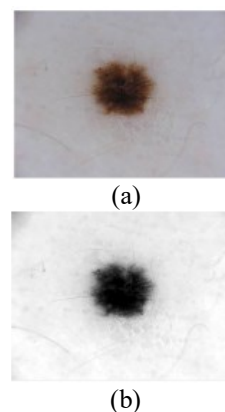
Meskipun ABCD merupakan metode yang umumnya digunakan, tetapi pada deteksi kanker kulit berbasis pengolahan citra digital, dari keempat fitur ini hanya dua yang memiliki fitur yang paling penting, yaitu ketidaksimetrisan dan fitur bentuk (Ruela, dkk., 2017). Oleh karena itu, kedua fitur ini yang akan digunakan dalam penelitian ini. Ketidaksimetrisan, terjadi apabila area "kelainan kulit" bersifat simetris maka tidak bersifat kanker (*benign*). Apabila area "kelainan kulit" bersifat asimetri, maka luka dianggap bersifat kanker (*malignant*). Ketidakteraturan tepian (*border irregularity*), beberapa area "kelainan kulit" yang

bersifat kanker memiliki tepian yang kabur, berlekuk, dan kasar.



Gambar 1. Diagram blok deteksi dini kanker menggunakan ekstraksi fitur dan k-NN

Diagram blok yang digunakan untuk ekstraksi fitur serta klasifikasi dengan metode regresi ditunjukkan pada Gambar 1. Dalam penelitian ini, *preprocessing* dilakukan dalam dua tahap, yaitu mengubah citra warna RGB (24 bit) menjadi citra abu (8 bit). Tahap kedua adalah proses penghilangan derau yang dilakukan dengan menggunakan filter median. Tipe filter ini adalah filter lolos-rendah, yang bekerja dengan mengganti nilai suatu piksel pada citra asal dengan nilai median dari piksel tersebut dan tetangganya. Citra sebelum dan sesudah dilakukan prapengolahan dapat dilihat pada Gambar 2.



Gambar 2. Citra (a) sebelum dan (b) sesudah prapengolahan



Gambar 3. Citra setelah proses binerisasi

Terdapat beberapa metode segmentasi yang dapat digunakan pada citra kanker kulit, yaitu *Kapur's multithresholding* maupun *level set-based segmentation*. *Otsu thresholding* merupakan metode binerisasi citra dengan memilih nilai batas ambang (*threshold*) secara otomatis dari tingkat keabuan histogram melalui analisis diskriminan. Metode ini dipilih karena merupakan metode yang paling umum digunakan dalam pengolahan citra kanker kulit (Rajinikanth, dkk., 2017). Pada penelitian ini, nilai ambang batas yang dipilih adalah 140, sehingga piksel yang bernilai di atas 140 akan diubah menjadi putih (1), sementara di bawah 140 diubah menjadi hitam (0). Citra hasil binerisasi dapat dilihat pada Gambar 3.

Proses binerisasi ini dilakukan untuk mempermudah proses selanjutnya. Area pada daerah penting (*Region of Interest / RoI*) adalah area yang berwarna putih atau bernilai 1, oleh karena itu dilakukan operasi pembalikan (*invert*). Untuk melakukan analisis terhadap tepian citra, dilakukan deteksi objek terlebih dahulu. Algoritma deteksi objek yang dilakukan pada penelitian ini adalah *Canny edge detection*. Ilustrasi penghitungan ketidakteraturan tepian citra kanker kulit ini dapat dilihat pada Gambar 4.

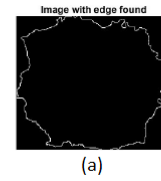
Secara umum tingkat asimetri suatu objek dapat dihitung dengan membagi objek menjadi dua bagian dan menghitung bagian yang berbeda satu sama lain. Sehingga tingkat asimetri suatu objek dapat dihitung dengan persamaan berikut:

$$AS = \frac{\Delta T}{T} \times 100\% \tag{1}$$

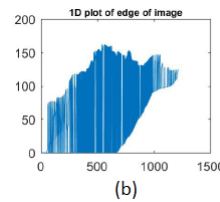
dimana,

ΔT adalah jumlah area yang berbeda, dan T adalah area keseluruhan dari objek.

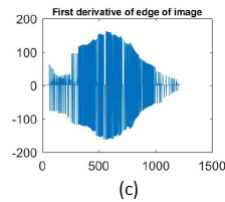
Pada penelitian ini tingkat asimetri dihitung dengan cara mencerminkan objek pada sumbu tertentu lalu menghitung area yang tumpang tindih sama lain. Jumlah area yang saling tumpang tindih kemudian dibandingkan dengan area keseluruhan. Apabila jumlah area yang saling tumpang tindih sama dengan area keseluruhan, hal itu berarti objek memiliki bentuk yang simetris, namun apabila jumlah area yang saling tumpang tindih lebih kecil dari jumlah area keseluruhan, hal itu berarti objek tidak simetris. Semakin kecil jumlah area yang tumpang tindih, maka semakin besar tingkat ketidaksimetrisannya.



(a)



(b)

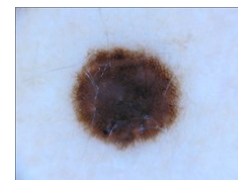


(c)

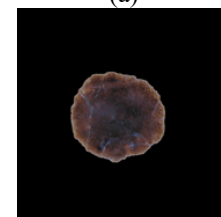
Gambar 4. Penghitungan tingkat ketidakteraturan tepian

Analisis ketidakteraturan tepian cukup kompleks karena sifatnya yang subjektif. Pada penelitian ini analisis ketidakteraturan tepian dilakukan dengan menghitung jarak setiap tepian terhadap pusat massa daerah "kelainan kulit" dan menggambarannya dalam grafik 1 dimensi seperti pada Gambar 4. Penggambaran tepian dalam grafik 1 dimensi ini digunakan untuk memudahkan analisis. Apabila grafik cenderung lurus atau hiperbola tanpa ada banyak nilai maksimum maupun nilai minimum lokal, maka berarti tepian memiliki tepian yang jelas dan normal. Penentuan jumlah nilai maksimum dan minimum lokal dalam 1 grafik tepian dilakukan dengan menghitung turunan dari grafik tersebut lalu melakukan perhitungan jumlah *zero crossing* dalam grafik turunan tersebut. *Zero crossing* dalam suatu grafik turunan menggambarkan adanya nilai maksimum ataupun minimum lokal. Semakin banyak nilai *zero crossing* dalam grafik turunan tepian, maka semakin tidak teratur tepian dari citra tersebut.

Sebelum dilakukan klasifikasi, data dipetakan dalam suatu grafik untuk diidentifikasi penyebarannya. Dari hasil persebaran data, klasifikasi dengan *hyperplane* yang paling optimal adalah dengan menggunakan algoritma *K-Nearest Neighbor*.



(a)



(b)

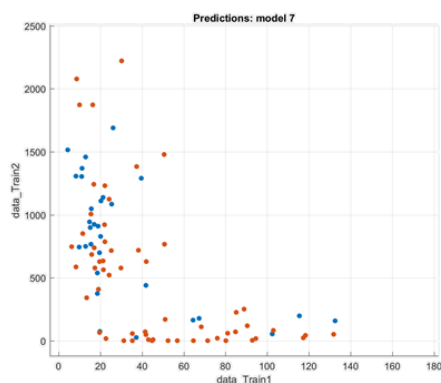
Gambar 5. Citra (a) sebelum dan (b) sesudah *resizing & masking*

CNN ini merupakan algoritma yang menggunakan konvolusi dan *neural network*.

Konvolusi pada algoritma akan menghasilkan fitur-fitur yang paling menonjol dari sebuah citra. Kemudian, hasil konvolusi tersebut akan diberikan komponen non-linieritas. Proses selanjutnya dari CNN ialah mengkonversi dimensi citra menjadi lebih kecil dengan teknik *max pooling*. Lalu vektor matriks citra akan diubah menjadi satu dimensi dengan metode *flattening*. Setelah itu, hasil proses citra tersebut akan diklasifikasikan dengan algoritma *neural network*. Selanjutnya, dilakukan pelatihan agar dapat mendeteksi kanker kulit. Proses yang terakhir adalah menentukan akurasi sistem.

Pada penelitian ini, sebelum dilakukan deteksi dini kanker kulit menggunakan CNN, terlebih dahulu dilakukan prapengolahan dengan menggunakan *library OpenCV*. Tahap pertama adalah melakukan penyeragaman dimensi citra menjadi 512×512 piksel. Kemudian dilakukan proses *masking* untuk memperoleh RoI. Citra sebelum dan sesudah prapengolahan dapat dilihat pada Gambar 5.

Pada penelitian ini digunakan arsitektur CNN *Convolutional Auto Encoder (CAE)* sebanyak 6 *layer*. Terdapat beberapa parameter serta pengaturan yang digunakan. Pertama, konvolusi filter 3×3 dengan 32 tipe sebanyak 3 *layer*. Konvolusi ini disesuaikan dengan ukuran piksel yang diberikan sebagai masukan, yaitu $32 \times 32 \times 3$. Fungsi dari *layer* konvolusi adalah melakukan ekstraksi fitur-fitur tertentu pada suatu citra. Selanjutnya adalah *activation layer*, fungsi yang digunakan adalah aktivasi ReLU (*Rectified Linear Unit*), karena fungsi ini mampu untuk menanggulangi *vanishing gradient problem* pada saat proses latih. Parameter selanjutnya adalah *max pooling* yang merupakan tipe *non-linear down sampling* yang dilakukan terhadap model hasil konvolusi untuk memperkecil dimensi citra. Ukuran *kernel* yang dipakai untuk *Max Pooling* adalah 2×2 , sehingga ukuran dari dimensi modelnya diperkecil menjadi setengahnya. Parameter terakhir adalah *dense* atau *hidden layer* yang terdapat pada *classifier network*. Pada penelitian ini digunakan tiga *hidden layer*. *Hidden layer* pertama mempunyai 64 neuron, *hidden layer* kedua memiliki 32 neuron, dan *hidden layer* ketiga mempunyai 1 neuron.



Gambar 6. Persebaran data

Pada penelitian ini digunakan algoritma ADAM (*adaptive moment estimation*) *optimizers* yang

merupakan pengganti *classical stochastic gradient* untuk memperbarui *network weight iterative* dalam proses pelatihan data. Setelah itu, dilakukan augmentasi data pada citra dengan membuat *classifier model* yang lebih efektif dalam melakukan pengenalan dan klasifikasi citra, meskipun dengan jumlah dataset yang relatif lebih sedikit.

3. HASIL DAN ANALISIS

3.1. Klasifikasi dengan k-NN

Pada penelitian yang menggunakan klasifikasi dengan model regresi, set data yang digunakan berasal dari *International Skin Imaging Collaboration (ISIC)* dengan rincian sebagai berikut: 80 citra latih *benign*, 80 citra latih *malignant*, 20 citra uji *benign*, dan 20 citra uji *malignant*. Sebelum dilakukan klasifikasi, data dipetakan dalam suatu grafik untuk diidentifikasi penyebarannya. Persebaran data dengan fitur ketidakteraturan tepi dan ketidaksimetrisan dapat dilihat pada Gambar 6.

Deteksi kanker kulit dengan ekstraksi fitur dan regresi k-NN menghasilkan akurasi sebesar 75% dengan *true positive rate (precision)* sebesar 85% dan *true negative rate (specificity)* sebesar 72%. Nilai akurasi ini juga dipengaruhi oleh set data yang digunakan dimana dataset yang digunakan tidak memiliki standar pengambilan gambar yang sama sehingga tingkat kecerahan pada setiap gambar menjadi berbeda-beda, sementara binerisasi yang dilakukan untuk melakukan segmentasi gambar memiliki nilai yang tepat. Hal ini dapat menyebabkan adanya penurunan performa sistem dalam melakukan segmentasi dan kalkulasi parameter.

3.2. Klasifikasi dengan CNN

Citra kanker kulit yang digunakan dalam klasifikasi menggunakan CNN berjumlah 800, dengan rincian sebagai berikut: 320 citra latih *benign*, 320 citra latih *malignant*, 80 citra uji *benign*, dan 80 citra uji *malignant*. Tahap selanjutnya adalah mendefinisikan parameter latih. Jumlah *epochs* yang digunakan adalah sebesar 20 kali. Pada penelitian ini, algoritma ADAM *optimizers* digunakan untuk melatih sistem yang merupakan metode optimasi stokastik dengan memanfaatkan prinsip *gradient descent*. Dengan menggunakan metode yang diusulkan ini, didapatkan akurasi sebesar 76,56%.

4. KESIMPULAN

Kedua metode tersebut menghasilkan hasil yang kurang lebih sama, yaitu 75% untuk klasifikasi menggunakan k-NN dan 76,56% dengan CNN. Akurasi yang didapatkan ini dipengaruhi oleh set data yang digunakan, dimana set data yang digunakan terlalu bervariasi dan jumlahnya yang terlalu sedikit sehingga sulit untuk melakukan proses generalisasi.

Pengembangan lanjutan dari penelitian adalah klasifikasi dilakukan dengan menggunakan lebih

banyak data dan dataset yang lebih terstandarisasi sehingga keempat parameter dapat dipertimbangkan dalam proses klasifikasi kanker kulit untuk metode k-NN. Disamping itu, metode yang dikembangkan ini dapat diimplementasikan ke dalam sistem deteksi berbasis *smartphone* yang dapat digunakan sebagai alat deteksi dini kanker kulit karena memiliki beban komputasi yang relatif lebih rendah.

DAFTAR PUSTAKA

- ARROYO, J.L.G. dan ZAPIRAIN, B.G., 2014. Detection of pigment network in dermoscopy images using supervised machine learning and structural analysis. *Computers in biology and medicine*, 44, pp.144-157.
- DALILA, F., ZOHRA, A., REDA, K. dan HOCINE, C., 2017. Segmentation and classification of melanoma and benign skin lesions. *Optik*, 140, pp.749-761.
- HARANGI, B., 2018. Skin lesion classification with ensembles of deep convolutional neural networks. *Journal of biomedical informatics*, 86, pp.25-32.
- JAIN, S. dan PISE, N., 2015. Computer aided melanoma skin cancer detection using image processing. *Procedia Computer Science*, 48, pp.735-740.
- JAFARI, M.H., NASR-ESFAHANI, E., KARIMI, N., SOROUSHMEHR, S.R., SAMAVI, S. dan NAJARIAN, K., 2017. Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma. *International journal of computer assisted radiology and surgery*, 12(6), pp.1021-1030.
- KASMI, R. dan MOKRANI, K., 2016. Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule. *IET Image Processing*, 10(6), pp.448-455.
- KAYA, S., BAYRAKTAR, M., KOCKARA, S., METE, M., HALIC, T., FIELD, H.E. dan WONG, H.K., 2016, October. Abrupt skin lesion border cutoff measurement for malignancy detection in dermoscopy images. In *BMC bioinformatics*, 17(13), pp. 367. BioMed Central.
- MARCHETTI, M.A., CODELLA, N.C., DUSZA, S.W., GUTMAN, D.A., HELBA, B., KALLOO, A., MISHRA, N., CARRERA, C., CELEBI, M.E., DEFAZIO, J.L. dan JAIMES, N., 2018. Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. *Journal of the American Academy of Dermatology*, 78(2), pp.270-277.
- MEHTA, P. dan SHAH, B., 2016. Review on techniques and steps of computer aided skin cancer diagnosis. *Procedia Computer Science*, 85, pp.309-316.
- NASR-ESFAHANI, E., SAMAVI, S., KARIMI, N., SOROUSHMEHR, S.M.R., JAFARI, M.H., WARD, K. dan NAJARIAN, K., 2016, August. Melanoma detection by analysis of clinical images using convolutional neural network. In 2016 38th International Conference of the IEEE Engineering in Medicine and Biology Society, pp.1373-1376. IEEE.
- NAVARRO, F., ESCUDERO-VIÑOLO, M. and Bescós, J., 2018. Accurate segmentation and registration of skin lesion images to evaluate lesion change. *IEEE journal of biomedical and health informatics*, 23(2), pp.501-508.
- PATHAN, S., PRABHU, K.G. dan SIDDALINGASWAMY, P.C., 2018. Techniques and algorithms for computer aided diagnosis of pigmented skin lesions— A review. *Biomedical Signal Processing and Control*, 39, pp.237-262.
- RAJINIKANTH, V., MADHAVARAJA, N., SATAPATHY, S.C. dan FERNANDES, S.L., 2017. Otsu's multi-thresholding and active contour snake model to segment dermoscopy images. *Journal of Medical Imaging and Health Informatics*, 7(8), pp.1837-1840.
- RASTGOO, M., GARCIA, R., MOREL, O. dan MARZANI, F., 2015. Automatic differentiation of melanoma from dysplastic nevi. *Computerized Medical Imaging and Graphics*, 43, pp.44-52.
- RUELA, M., BARATA, C., MARQUES, J.S. dan ROZEIRA, J., 2017. A system for the detection of melanomas in dermoscopy images using shape and symmetry features. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 5(2), pp.127-137.
- ZORTEA, M., SCHOPF, T.R., THON, K., GEILHUF, M., HINDBERG, K., KIRCHESCH, H., MØLLERSEN, K., SCHULZ, J., SKRØVSETH, S.O. dan GODTLIEBSEN, F., 2014. Performance of a dermoscopy-based computer vision system for the diagnosis of pigmented skin lesions compared with visual evaluation by experienced dermatologists. *Artificial intelligence in medicine*, 60(1), pp.13-26.

Halaman ini sengaja dikosongkan