

RINGKASAN

PENGARUH PERBEDAAN RUANG WARNA PADA DATASET HAM10000 SEBAGAI MASUKAN CONVOLUTIONAL NEURAL NETWORK (CNN) UNTUK KLASIFIKASI KANKER KULIT

Muhammad Arif Fahrudin

Salah satu kanker paling berbahaya diantara jenis kanker lain yaitu kanker kulit yang mana disebabkan oleh pembelahan sel yang tidak normal. Kanker kulit diklasifikasikan berdasarkan jenisnya, yaitu melanoma dan non-melanoma. Tergantung pada jenis asalnya, lesi sel melanoma termasuk melanoma dan melanomatik nevus. Demikian pula, lesi non-melanoma termasuk karsinoma sel basal, keratosis teraktivasi, neoplasma dermal, lesi vaskular, dan keratosis jinak. *Convolutional Neural Network* (CNN) lebih baik dalam mengklasifikasi kanker kulit. Dokter kulit yang melakukan pemeriksaan visual pada kanker kulit memiliki akurasi 60% - 80%. Terdapat berbagai ruang warna alternatif yang dapat membantu meningkatkan kinerja CNN dan mempelajari representasi data yang lebih baik. Sehingga dilakukan perbandingan dimensi warna yang paling baik untuk CNN pada kanker kulit.

Model arsitektur klasifikasi dibuat sendiri menggunakan *MobileNetV2*. Dengan bantuan modul *OpenCV* dilakukan konversi dimensi warna dari RGB ke HSV, YCrCb, XYZ, *greyscale*. Konversi akan dilakukan ke *training data*, *validation data*, dan *test data*. Pada penelitian ini akan dibandingkan hasil model yang dihasilkan oleh keempat dimensi warna tersebut. Hasil model yang akan diambil berasal dari *test data*.

Ditinjau dari *test data* didapatkan hasil terbaik pada ruang warna RGB ukuran data 224x224 dengan nilai akurasi sebesar 0.9511 serta didapatkan hasil paling rendah yaitu pada ruang warna *greyscale* ukuran data 32x32 dengan nilai akurasi sebesar 0.6824. Ruang warna RGB, HSV, XYZ, YcbCr dan *greyscale* nilai dari *accuracy*, *precision* dan *recall* memiliki nilai yang sama atau mendekati pada setiap ruang warna. Model dengan ukuran data 224x224 memiliki akurasi yang lebih baik dibandingkan dengan model dengan ukuran data 32x32.

Kata kunci : CNN, *MobileNetV2*, Ruang Warna, Kanker Kulit

SUMMARY

THE EFFECT OF DIFFERENCES IN HAM10000 DATASET COLOR SPACE AS CONVOLUTIONAL NEURAL NETWORK (CNN) INPUT FOR SKIN CANCER CLASSIFICATION

Muhammad Arif Fahrudin

One of the most dangerous cancers among other types of cancer is skin cancer which is caused by abnormal cell division. Skin cancer is classified according to its type, namely melanoma and non-melanoma. Depending on the type of origin, melanoma cell lesions include melanoma and melanoma nevus. Similarly, non-melanoma lesions include basal cell carcinoma, activated keratitis, dermal neoplasms, vascular lesions, and benign keratitis. Convolutional Neural Network (CNN) is better at classifying skin cancer. Dermatologists who perform visual examinations of skin cancer have an accuracy of 60% - 80%. There are various alternative color spaces that can help improve CNN performance and learn better data representation. So that the comparison of the best color dimensions for CNN in skin cancer was carried out.

The classification architecture model was made by myself using MobileNetV2. With the help of the OpenCV module, color space conversion is carried out from RGB to HSV, YCrCb, XYZ, greyscale. Conversion will be carried out to training data, validation data, and test data. In this study, the results of the model generated by the four color space will be compared. The results of the model to be taken are derived from the test data.

Judging from the test data, the best results were obtained in the RGB color space with data size 224x224 with an accuracy value of 0.9511 and the lowest results obtained, namely in the greyscale color space, data size 32x32 with an accuracy value of 0.6824. Color space RGB, HSV, XYZ, YcbCr and greyscale values of accuracy, precision and recall have the same or close values in each color space. The model with a data size of 224x224 has better accuracy than the model with a data size of 32x32.

Keywords : CNN, MobileNetV2, Color Space, Skin Cancer