

Pendeteksian Penyakit Malaria Menggunakan *Medical Images* Analisis Dengan *Deep Learning Python*

Detection of Malaria Using Medical Images Analysis with Deep Learning Python

Afrizal Zein

Program Studi Teknik Informatika, STMIK Eresha
Jl. Raya Puspitek Serpong No. 10 Tangerang Selatan Banten
E-mail : zeinafrizal@gmail.com

Abstrak--- *Tahukah Anda bahwa Penyakit Malaria adalah penyakit menular yang menyebabkan lebih dari 400.000 kematian per tahun. Penyakit Malaria adalah endemik sejati di beberapa wilayah di dunia, yang berarti bahwa penyakit ini secara teratur ditemukan di wilayah tersebut.*

Malaria adalah masalah kesehatan global yang serius, dan diagnosis yang cepat dan akurat diperlukan untuk mengendalikan penyakit. Algoritma pemrosesan gambar untuk mengotomatisasi diagnosis malaria pada apusan darah tipis dikembangkan. Sistem klasifikasi citra dirancang untuk secara positif mengidentifikasi parasit malaria yang ada dalam noda darah tipis, dan membedakan spesies malaria. Dari masalah di atas penulis mengadakan penelitian sistem komputer vision yang secara otomatis dapat mendeteksi Penyakit Malaria menggunakan deep learning menggunakan dataset Rajaraman yang telah dipublikasikan olehnya.

Dari hasil pengujian Medical Images Analisis menunjukkan perolehan tingkat akurasi pendeteksian didapat mencapai sebesar di atas 97% pada data pengujian,

Kata Kunci : *Deteksi Malaria, Medical Images Analisis, Deep learning*

Abstract--- *Did you know that Malaria is an infectious disease that causes more than 400,000 deaths per year. Malaria is a true endemic in several regions of the world, which means that this disease is regularly found in the region. So, what makes some regions of the world more vulnerable to malaria while others are completely free of malaria?*

There are many components that make the area vulnerable to infectious diseases. We will be the main constituents below. When assessing the risk of an infectious disease outbreak we usually check how many people are in the population or at or below the poverty level. The higher the level of poverty, the higher the risk of infectious diseases, although some researchers will say the opposite - that malaria causes poverty.

From the problem above the authors conducted a computer vision system research that can automatically detect Malaria using deep learning using the Rajaraman dataset that was published by him.

From the results of the use of Medical Images Analysis shows that the accuracy of the obtained detection reaches above 97%. on test data,

Keywords : *Detection of Malaria, Medical Images Analisis, Deep learning*

1. PENDAHULUAN

Malaria adalah penyakit infeksi serius yang disebabkan oleh parasit darah perifer dari genus *Plasmodium*. Menurut *World Health Organization* (WHO), itu menyebabkan lebih dari 1 juta kematian yang timbul dari sekitar 300-500 juta infeksi setiap tahun. Meskipun ada teknik yang lebih baru], mikroskopi manual untuk pemeriksaan apusan darah (ditemukan pada akhir abad ke-19), saat ini "standar emas" untuk diagnosis malaria. Diagnosis menggunakan mikroskop membutuhkan pelatihan khusus dan keahlian yang cukup. Telah ditunjukkan dalam beberapa studi lapangan bahwa mikroskop manual bukan metode skrining yang dapat diandalkan bila dilakukan oleh non-ahli karena kurangnya pelatihan terutama di daerah pedesaan di mana malaria endemik Sebuah sistem otomatis bertujuan untuk melakukan tugas ini tanpa campur tangan manusia dan untuk menyedia

kan alat yang obyektif, dapat diandalkan, dan efisien untuk melakukannya.

Sistem diagnosis otomatis dapat dirancang dengan memahami keahlian diagnostik dan merepresentasikannya dengan algoritma pemrosesan gambar, analisis, dan pengenalan pola khusus yang disesuaikan. Meskipun ini bukan topik penelitian yang populer, sejumlah studi visi yang terlihat secara langsung membahas diagnosis otomatis malaria. Meskipun sangat khusus, jika angka kematian dianggap hasil mereka dapat dianggap lebih penting daripada beberapa aplikasi visi komputer populer lainnya.

Studi ini memberikan tinjauan tentang studi visi komputer diagnosis malaria dan bermaksud untuk mengisi celah di daerah ini dengan melakukannya. Ada beberapa interpretasi yang berbeda dari persyaratan dan dengan demikian penerapan solusi yang diusulkan untuk masalah tersebut. Di

sini, perbedaan-perbedaan ini dibahas; kepraktisan, ketangguhan, ketepatan solusi yang diusulkan dan penerapannya untuk melakukan tugas diagnosis yang sebenarnya dipertanyakan. Selain itu, metode evaluasi yang dipilih untuk mengukur dan menilai akurasi dibahas. Selain itu, beberapa karya lain dari literatur yang menyangkut sub-masalah atau sub-komponen yang diperlukan diperiksa dan ditempatkan dalam kerangka kerja pengenalan pola umum untuk aplikasi diagnosis. Tujuan dari makalah ini adalah untuk: 1) survei *state-of-the-art* dari metode tentang masalah; 2) menggambarkan kerangka visi komputer umum untuk melakukan tugas diagnosis; 3) menyelesaikan beberapa *ambiguitas* dari perspektif yang berbeda mengenai masalah, dan 4) menunjukkan beberapa karya masa depan untuk studi penelitian potensial.

Diagnosis mikroskopi dilakukan dengan pemeriksaan visual pemeriksaan noda darah secara manual. Seluruh proses membutuhkan kemampuan untuk membedakan antara komponen / benda bernoda non-parasit (misalnya sel darah merah, sel darah putih, trombosit, dan artefak) dan parasit malaria menggunakan informasi visual. Jika sampel darah didiagnosis positif (ada parasit yang hadir), kemampuan tambahan untuk membedakan spesies dan tahap kehidupan (yaitu identifikasi) diperlukan untuk menentukan infeksi.

Dari sudut pandang visi komputer, diagnosis malaria adalah masalah multi-bagian. Suatu sistem yang lengkap harus dilengkapi dengan fungsi-fungsi yang harus dilakukan: akuisisi citra, pra-pemrosesan, segmentasi (kandidat objek lokalisasi), dan tugas-tugas klasifikasi. Oleh karena itu, sistem diagnosis lengkap juga memerlukan beberapa fungsi seperti penempatan slide mikroskop, fokus otomatis, cepat, dan andal, serta akuisisi gambar. Beberapa studi tentang akuisisi citra diperiksa dalam bagian akuisisi Gambar. Biasanya, gambar yang diperoleh dari mikroskop memiliki beberapa variasi yang dapat mempengaruhi proses. Ini biasanya dibahas oleh fungsi pra-pemrosesan yang dibahas di bagian variasi Gambar. Langkah penting dalam analisis otomatis adalah untuk mendapatkan / menemukan kemungkinan sel yang terinfeksi (yaitu kandidat) yang merupakan obyek bernoda dalam gambar. Deteksi pewarnaan dan lokalisasi objek-objek ini dibahas dalam bagian Segmentasi dan pixel bernoda dan objek.

Untuk melakukan diagnosis pada sampel darah tepi, sistem harus mampu membedakan antara parasit malaria, artefak, dan komponen darah yang sehat. Mayoritas studi analisis citra terkait malaria yang ada (misalnya tidak membahas persyaratan ini. Ini menghasilkan solusi yang terlalu sederhana, yang tidak berlaku untuk diagnosis secara langsung. Di sisi lain, beberapa metode yang membahas diferensiasi misalnya memiliki hasil eksperimen yang terbatas untuk menunjukkan

bahwa solusi yang mereka usulkan sebanding dengan diagnosis mikroskop manual atau dapat menggantikannya. Untuk ini efek persyaratan untuk data eksperimen yang tepat dan set-up dibahas dalam bagian Diskusi. Untuk mengatur adegan, pengenalan singkat tentang parasit malaria, spesiesnya, dan tahapan siklus hidup disediakan di bagian berikutnya, diikuti dengan deskripsi singkat diagnosis mikroskopi.



Gambar 1. Nyamuk dapat membawa penyakit infeksi seperti malaria

2. TINJAUAN PUSTAKA

Menurut penelitian yang dilakukan oleh Eli Gibson, Wenqi Li (2018) yang berjudul "*Computer Methods and Programs in Biomedicine*", Latar belakang dan tujuan: Analisis citra medis dan masalah intervensi berbantuan komputer semakin diatasi dengan solusi berbasis pembelajaran yang mendalam. Platform *deep-learning* yang mapan bersifat fleksibel tetapi tidak menyediakan fungsionalitas khusus untuk analisis citra medis dan mengadaptasinya untuk domain aplikasi ini membutuhkan upaya implementasi yang substansial. Akibatnya, telah ada duplikasi substansial dari upaya dan infrastruktur yang tidak kompatibel dikembangkan di banyak kelompok penelitian. Karya ini menyajikan platform NiftyNet sumber terbuka untuk pembelajaran mendalam dalam pencitraan medis. Ambisi NiftyNet adalah untuk mempercepat dan menyederhanakan pengembangan solusi ini, dan untuk menyediakan mekanisme umum untuk menyebarluaskan hasil penelitian bagi masyarakat untuk digunakan, diadaptasi, dan dikembangkan. Metode: Infrastruktur NiftyNet menyediakan jalur pipa pembelajaran dalam modular untuk berbagai aplikasi pencitraan medis termasuk segmentasi, regresi, pembuatan gambar dan aplikasi pembelajaran representasi. Komponen dari pipa NiftyNet termasuk pemuatan data, augmentasi data, arsitektur jaringan, fungsi kehilangan dan metrik evaluasi dirancang untuk, dan mengambil keuntungan dari, keanehan analisis citra medis dan intervensi yang dibantu komputer. NiftyNet dibangun di atas kerangka kerja TensorFlow dan mendukung fitur-

fitur seperti visualisasi TensorBoard dari gambar 2D dan 3D dan grafik komputasi secara default. Hasil: Kami menyajikan tiga aplikasi analisis gambar medis ilustratif yang dibangun menggunakan infrastruktur NiftyNet : (1) segmentasi beberapa organ perut dari *computed tomography*; (2) regresi gambar untuk memprediksi peta atenuasi tomografi komputer dari citra resonansi magnetik otak; dan (3) generasi gambar USG disimulasikan untuk pose anatomi tertentu. Kesimpulan: Infrastruktur NiftyNet memungkinkan para peneliti untuk dengan cepat mengembangkan dan mendistribusikan solusi pembelajaran mendalam untuk segmentasi, regresi, pembuatan gambar dan aplikasi pembelajaran representasi, atau memperluas platform ke aplikasi baru.

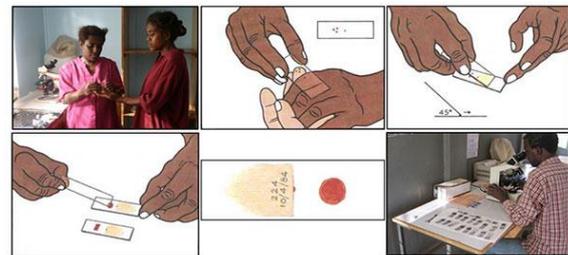
Menurut penelitian yang dilakukan oleh Geert Litjens, Thijs Kooi (2017), yang berjudul “ *A Survey on Deep Learning in Medical Image Analysis* “; Algoritma pembelajaran mendalam, khususnya jaringan konvolusional, telah dengan cepat menjadi metodologi pilihan untuk meng analisis gambar medis. Makalah ini mengulas konsep-konsep pembelajaran mendalam utama yang berkaitan dengan analisis citra medis dan merangkum lebih dari 300 kontribusi ke lapangan, yang sebagian besar muncul di tahun lalu. Kami mensurvei penggunaan pembelajaran mendalam untuk klasifikasi gambar, deteksi objek, segmentasi, registrasi, dan tugas-tugas lainnya. Gambaran singkat disediakan dari studi per area aplikasi: neuro, retina, paru, patologi digital, payudara, jantung, perut, muskuloskeletal. Kami diakhiri dengan ringkasan terkini, diskusi kritis tentang tantangan terbuka dan arahan untuk penelitian di masa depan.

3. METODOLOGI PENELITIAN

Ada banyak paradigma visi komputer yang berbeda, yang dapat digunakan untuk membangun sistem analisis / pengenalan visual otomatis. Pekerjaan yang ada pada malaria umumnya menggunakan morfologi matematis untuk memroses gambar karena cocok dengan analisis objek seperti blob seperti sel darah. Di sisi lain, untuk membedakan antara pola yang diamati pendekatan pembelajaran berbasis statistik sangat populer. Pembaca dapat menemukan dalam makalah ini banyak istilah teknis yang digunakan untuk menjelaskan masalah atau pendekatan yang berbeda. File tambahan 1 memberikan definisi singkat untuk beberapa istilah yang terkait dengan pemrosesan gambar (misalnya pixel, histogram, gradient), operator morfologi matematika (misalnya erosi, pelebaran, pembukaan, granulometri), konsep klasifikasi pola (misalnya fitur, penggolong, dan pelatihan). Informasi lebih rinci dapat ditemukan dalam sumber-sumber berikut: pada morfologi matematika, pada pengenalan pola statistik, dan pada pemrosesan gambar umum].

Akuisisi gambar

Dalam jumlah gambar yang diperlukan untuk menangkap 2 cm² wilayah spesimen pada 20 × pembesaran dihitung menjadi hampir 1.300 gambar menggunakan 1.300 × 1.030 piksel 2/3 inci muatan perangkat gabungan (sensor CCD) kamera. Diagnosis malaria membutuhkan pembesaran objektif 100x (disarankan untuk pemeriksaan manual), sehingga jumlah foto yang diambil akan 25 kali lebih tinggi. Oleh karena itu, itu kira-kira sesuai dengan lebih dari 30.000 gerakan slide, fokus, dan operasi sensor rana CCD yang membutuhkan teknik yang sangat cepat. Untuk mengurangi kebutuhan waktu, Wetzel dkk [34] mengusulkan untuk menangkap gambar sementara slide terus bergerak, yang memperkenalkan masalah gambar kabur. Mereka mengusulkan untuk menggunakan lampu strobo Xenon bukan lampu konvensional untuk memecahkan masalah ini, yang mungkin meningkatkan biaya secara substansial.



Gambar 2. Dua metode pengujian untuk malaria meliputi (1) smear darah, dan (2) pengujian antigen (yaitu tes cepat). Ini adalah dua cara umum pengujian malaria yang paling sering didiskusikan dan digunakan

Harus dicatat bahwa seorang ahli manusia akan membutuhkan lebih banyak waktu untuk pergi melalui slide dan memfokuskan mikroskop untuk mengamati 30.000 bidang. Oleh karena itu, jumlah bidang yang akan diperiksa oleh pakar biasanya lebih kecil. Dalam tutorial mikroskopi malaria WHO, pemeriksaan hanya 100 bidang disarankan sebelum memberikan keputusan negatif. Selain itu, dalam film tebal, jika parasit diamati di lapangan, 100 lebih banyak ladang (atau 200 sel darah putih, 0,025 µl darah) akan cukup untuk menghitung parasitemia. Karena kurang sensitif, pemeriksaan rutin pada film darah tipis tidak dianjurkan untuk jenis diagnosis positif / negatif. Namun, jika parasit ditemukan, pemeriksaan 50 bidang (rata-rata 200 per lapangan menghasilkan 10.000 sel darah merah secara total) akan cukup untuk menghitung parasitemia dalam film tipis. Dengan demikian, persyaratan kecepatan sistem akuisisi citra relatif mudah dicapai. Selain itu, solusi pemfokusan cepat yang baru-baru ini muncul dan mesin pemindaian geser komersial khusus (misalnya. Paten No. 563437 yang diajukan pada tanggal 2000-05-03)

menjanjikan untuk menyelesaikan kendala praktis yang penting ini.

4. HASIL DAN PEMBAHASAN

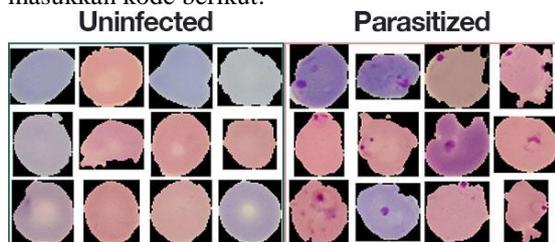
Kumpulan data malaria kami tidak memiliki data pra-pemisahan untuk pelatihan, validasi, dan pengujian, jadi kami harus melakukan pemisahan sendiri. Untuk membuat pemisahan data kami, kami akan menggunakan skrip `build_dataset.py` - skrip ini akan:

Ambil jalan ke semua gambar contoh kami dan acak acak mereka. Pisahkan jalur gambar ke pelatihan, validasi, dan pengujian.

Buat tiga sub-direktori baru di direktori / malaria, yaitu pelatihan /, validasi /, dan pengujian /.

Secara otomatis menyalin gambar ke direktori yang sesuai.

Untuk melihat bagaimana proses pemisahan data dilakukan, buka `build_dataset.py` dan masukkan kode berikut:



Gambar 3. Sebuah subset dari Kumpulan Malaria yang disediakan oleh National Institute of Health (NIH). Kami akan menggunakan dataset ini untuk mengembangkan model klasifikasi pencitraan medis pembelajaran mendalam dengan Python, OpenCV, dan Keras.

```
# import paket yang diperlukan
from pyimagesearch import config
from imutils import paths
import random
import shutil
import os

imagePaths = list(paths.list_images(config.ORIG_INPUT_DATASET))
random.seed(42)
random.shuffle(imagePaths)
```

Sekarang setelah gambar diatur ke dalam perpecahan masing-masing, mari kita mendefinisikan kumpulan data yang akan kita bangun:

```
# Mendefinisikan kumpulan data yang akan kami buat
datasets = [
    ("training", trainPaths,
     config.TRAIN_PATH),
    ("validation", valPaths,
     config.VAL_PATH),
    ("testing", testPaths,
     config.TEST_PATH)
]
```

Selanjutnya kita membuat 3 tupel sebagai berikut :

```
# loop diatas datasets
for (dtype, imagePath, baseOutput) in datasets:
    # tampilkan pemisahan data yang kami buat
    print("[INFO] building '{}'
    split".format(dtype))

    # jika direktori output basis output tidak ada
    ,
    buatlah
    if not os.path.exists(baseOutput):
        print("[INFO] 'creating {}'
        directory".format(baseOutput))
        os.makedirs(baseOutput)

    for inputPath in imagePath:
        # extract the filename of the input image along with its
        filename =
        inputPath.split(os.path.sep)[-1]
        label =
        inputPath.split(os.path.sep)[-2]

        # build the path to the label directory
        labelPath =
        os.path.sep.join([baseOutput, label])

        if not
        os.path.exists(labelPath):
            print("[INFO]
            'creating {}' directory".format(labelPath))

            os.makedirs(labelPath)

        p =
        os.path.sep.join([labelPath, filename])
        shutil.copy2(inputPath, p)
```

Pelatihan model pembelajaran mendalam untuk analisis citra medis

Sekarang setelah kami membuat pemisahan data kami, mari lanjutkan dan latih model pembelajaran mendalam kami untuk analisis gambar medis.

Seperti yang saya sebutkan sebelumnya dalam tutorial ini, tujuan saya adalah untuk menggunakan kembali kode sebanyak mungkin dari bab-bab dalam buku saya, Deep Learning for Computer Vision with Python. Bahkan, lebih dari 75% + kode langsung dari contoh teks dan kode.

Waktu adalah esensi ketika datang ke analisis gambar medis, sehingga semakin kita dapat bersandar pada kode yang stabil dan stabil, semakin baik.

Seperti yang akan kita lihat, kita akan dapat menggunakan kode ini untuk mendapatkan akurasi 97%.

```
# set the matplotlib backend so figures can be
saved in the background
import matplotlib
matplotlib.use("Agg")

# import the necessary packages
from keras.preprocessing.image import
ImageDataGenerator
from keras.callbacks import
LearningRateScheduler
from keras.optimizers import SGD
from py.resnet import ResNet
from py import config
from sklearn.metrics import classification_report
from imutils import paths
import matplotlib.pyplot as plt
import numpy as np
import argparse

# construct the argument parser and parse the
arguments
ap = argparse.ArgumentParser()
ap.add_argument("-p", "--plot", type=str,
default="plot.png",
help="path to output loss/accuracy
plot")
args = vars(ap.parse_args())
```

Kami menggunakan keras untuk melatih model pembelajaran mendalam citra medis kami, sklearn untuk mencetak classification_report, mengambil jalur dari dataset kami, numpy untuk pemrosesan numerik, dan argparse untuk parsing argumen baris perintah.

Yang rumit adalah matplotlib. Karena kami menyimpan alur ke disk (dan dalam kasus saya, pada mesin tanpa kepala), kami perlu menggunakan backend "Agg".

Kami tidak akan membahas arsitektur ResNet dalam tutorial ini, tetapi jika Anda tertarik untuk mempelajari lebih lanjut, pastikan untuk merujuk ke publikasi ResNet resmi serta Pembelajaran Jauh untuk Visi Komputer dengan Python tempat saya meninjau ResNet secara mendetail.

Kami memiliki argumen baris perintah tunggal yang diuraikan pada Baris 18-21, --plot. Secara default, plot kami akan ditempatkan di direktori kerja saat ini dan bernama plot.png. Sebagai alternatif, Anda dapat menyediakan nama file / jalur yang berbeda pada baris perintah ketika Anda pergi untuk menjalankan program.

Hasil medical image analysis.

```
$ python train_model.py
Found 19842 images belonging to 2 classes.
Found 2204 images belonging to 2 classes.
Found 5512 images belonging to 2 classes.
...
Epoch 1/50
620/620
[=====] -
67s - loss: 0.8723 - acc: 0.8459 - val_loss:
0.6020 - val_acc: 0.9508
Epoch 2/50
620/620
[=====] -
66s - loss: 0.6017 - acc: 0.9424 - val_loss:
0.5285 - val_acc: 0.9576
Epoch 3/50
620/620
[=====] -
65s - loss: 0.4834 - acc: 0.9525 - val_loss:
0.4210 - val_acc: 0.9609
...
Epoch 48/50
620/620
[=====] -
65s - loss: 0.1343 - acc: 0.9646 - val_loss:
0.1216 - val_acc: 0.9659
Epoch 49/50
620/620
[=====] -
65s - loss: 0.1344 - acc: 0.9637 - val_loss:
0.1184 - val_acc: 0.9678
Epoch 50/50
620/620
[=====] -
65s - loss: 0.1312 - acc: 0.9650 - val_loss:
0.1162 - val_acc: 0.9678
[INFO] serializing network...
[INFO] evaluating network...
precision recall f1-score support

Parasitized 0.97 0.97 0.97 2786
Uninfected 0.97 0.97 0.97 2726

avg / total 0.97 0.97 0.97 5512
```

5. SIMPULAN

Secara keseluruhan, seluruh proses pelatihan hanya membutuhkan waktu 54 menit (jauh lebih cepat daripada proses pelatihan 24 jam dari metode NIH). Pada kita memperoleh :

Akurasi 96,50% pada data pelatihan

Akurasi 96,78% pada data validasi

Akurasi 97% pada data pengujian

Ada sejumlah manfaat untuk menggunakan model berbasis ResNet yang kami latih di sini hari ini untuk analisis citra medis.

DAFTAR PUSTAKA

- [1] Eli Gibson, Wenqi Li. 2018, *Computer Methods and Programs in Biomedicine, journal* homepage: www.elsevier.com/locate/cmpb
- [2] Geert Litjens, Thijs Kooi. 2017. *A Survey on Deep Learning in Medical Image Analysis*, Radboud University Medical Center Nijmegen, The Netherlands
- [3] I Wayan Suartika E. P, Arya Yudhi Wijaya. 2016. *Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101, Jurnal Teknik ITS Vol. 5, No. 1, (2016) ISSN: 2337-3539 (2301-9271 Print)*
- [4] Afrizal Zein. 2016. *Pendeteksian Multi Wajah dan Recognition Secara Real Time Menggunakan Metoda Principal Component Analisis*, Jurnal ESIT Vol 12 No. 1 (2018)
- [5] Afrizal Zein. 2018. *Menggabungkan Dua Wajah Dengan Metoda Ensemble Regression Trees Menggunakan Pustaka Dlib Dan Opencv Python*, Jurnal ESIT Vol 13 No. 2 (2018)
- [6] Afrizal Zein. 2018. *Pendeteksian Kantuk Secara Real Time Menggunakan Pustaka OPENCV dan DLIB PYTHON*, Jurnal SAINSTECH Vol. 28 No. 2, (2018).