KLASIFIKASI PENYAKIT MALARIA BERDASARKAN CITRA SEL DARAH MENGGUNAKAN RESNET34

KERJA PRAKTEK

Diajukan untuk Memenuhi Persyaratan Akademik dalam

Menyelesaikan Pendidikan pada Program Studi

S1 Teknik Informatika Universitas Kristen Maranatha

Oleh

**Stefanus Hermawan**

**1772023**



**PROGRAM STUDI S1 TEKNIK INFORMATIKA**

**FAKULTAS TEKNOLOGI INFORMASI**

**UNIVERSITAS KRISTEN MARANATHA**

**BANDUNG**

**2020**

# DAFTAR ISI

[DAFTAR ISI 2](#_Toc39770335)

[BAB 1 PENDAHULUAN 5](#_Toc39770336)

[1.1 Latar Belakang 5](#_Toc39770337)

[1.2 Rumusan Masalah 6](#_Toc39770338)

[1.3 Tujuan Pembahasan 6](#_Toc39770339)

[1.4 Ruang Lingkup 6](#_Toc39770340)

[1.5 Sumber Data 6](#_Toc39770341)

[1.6 Sistematika Penyajian 7](#_Toc39770342)

[BAB 2 KAJIAN TEORI 8](#_Toc39770343)

[2.1 *Residual Neural Networks* 8](#_Toc39770344)

[*2.2* *Learning Rate Finder* 8](#_Toc39770345)

[*2.3* *One Cycle Policy* 9](#_Toc39770346)

[*2.4* *FastAi* 9](#_Toc39770347)

[2.5 *Flask* 9](#_Toc39770348)

[2.6 Google Colaboratory 10](#_Toc39770349)

[2.7 Freelancer – Bootstrap Theme 11](#_Toc39770350)

[2.8 Dataset 11](#_Toc39770351)

[BAB 3 ANALISIS DAN RANCANGAN SISTEM 12](#_Toc39770352)

[*3.1* *Data Preprocessing* 12](#_Toc39770353)

[3.2 Rancangan Tampilan Website 12](#_Toc39770354)

[3.2.1 Rancangan Tampilan Utama 12](#_Toc39770355)

[3.2.2 Rancangan Tampilan *Prediction Result* 14](#_Toc39770356)

[BAB 4 IMPLEMENTASI 15](#_Toc39770357)

[4.1 Pelatihan Model 15](#_Toc39770358)

[4.2 Pembuatan *Backend Website* 15](#_Toc39770359)

[4.3 Sub Topik Bab 4 15](#_Toc39770360)

[BAB 5 PENGUJIAN 16](#_Toc39770361)

[5.1 Sub Topik Bab 5 16](#_Toc39770362)

[5.2 Pengujian *Black Box* 16](#_Toc39770363)

[5.3 Pengujian *White* *Box* 16](#_Toc39770364)

[5.4 *User* *Acceptance* *Test* 16](#_Toc39770365)

[5.5 *Benchmarking* 16](#_Toc39770366)

[BAB 6 SIMPULAN DAN SARAN 17](#_Toc39770367)

[6.1 Simpulan 17](#_Toc39770368)

[6.2 Saran 17](#_Toc39770369)

[DAFTAR PUSTAKA 18](#_Toc39770370)

# PENDAHULUAN

## Latar Belakang

Menurut data World Health Organization pada tahun 2017, terdapat 219 juta kasus penyakit malaria yang tersebar di 90 negara, diperkirakan penderita malaria yang meninggal sebanyak 435.000 kasus dan kasus terbanyak berada di wilayah Afrika yang tersebar sebanyak 92% kasus malaria dunia dan 93% kasus penderita yang meninggal [1].

Sedangkan di Indonesia, menurut Kementerian Kesehatan Republik Indonesia pada tahun 2017, terdapat 42% wilayah Indonesia yang belum terbebas dari penyakit malaria, dan jumlah kasus malaria tercatat sebanyak 261.617 kasus malaria secara nasional dengan kasus wilayah Papua, Papua Barat, dan NTT dengan total 10,7 juta penduduk [2].

Parasit malaria dapat diidentifikasi dengan cara memeriksa darah pasien dengan menggunakan mikroskop. Sebelum pemeriksaan spesimen, sel darah akan terlebih dahulu diwarnai agar parasit pada sel darah dapat dibedakan.

Citra sel darah merah diperoleh melalui penangkapan citra menggunakan kamera ponsel cerdas berbasis Android yang dilengkapi dengan optik mikroskop medan terang (*bright-field microscopy*) [3].

Maka dari itu dalam kerja praktek ini akan dibahas bagaimana cara mendeteksi penyakit malaria menggunakan Convolutional Neural Networks dari citra digital sel darah serta mengembangkan *website* sebagai tampilan antarmuka untuk mempermudah penggunaan bagi pengguna.

## Rumusan Masalah

Berdasarkan latar belakang masalah di atas, rumusan masalah yang dapat dirumuskan adalah:

1. Bagaimana cara model *Convolutional Neural Networks* untuk mendeteksi malaria?
2. Bagaimana cara mengembangkan *website* sebagai tampilan antarmuka pengguna untuk model *Convolutional Neural Networks*?
3. Bagaimana cara mengukur kinerja model dalam mendeteksi malaria?

## Tujuan Pembahasan

Adapun tujuan pembahasan dari Kerja Praktek ini adalah :

1. Menggunakan ResNet34 untuk membangun model Convolutional Neural Networks untuk mendeteksi malaria
2. Mengembangkan website sebagai tampilan antarmuka pengguna untuk model ResNet34
3. Melaksanakan penilaian kinerja terhadap model dengan sekumpulan data citra sel darah

## Ruang Lingkup

Ruang lingkup yang akan dibahas dalam laporan Kerja Praktek ini adalah :

1. Menggunakan model ResNet34 yang disediakan FastAi untuk melatih model
2. Mengembangkan website sebagai antarmuka untuk pengguna
3. Data untuk melatih dan mengukur kinerja model adalah citra sel darah yang diambil dari website U.S National Library of Medicine.

## Sumber Data

Dataset yang digunakan untuk melatih dan mengukur kinerja model adalah sekumpulan citra yang telah dikumpulkan dan telah diberi label oleh U.S National Library of Medicine [4].

Model yang digunakan adalah model ResNet34 [5] yang terdapat dalam FastAi library [6].

## Sistematika Penyajian

Sistematika penyajian laporan kerja praktek ini diuraikan sebagai berikut. Bab pertama diisi dengan pendahuluan yang mencakup latar belakang masalah, rumusan masalah, tujuan pembahasan, ruang lingkup, sumber data, serta sistematika penyajian, bab kedua diisi kajian teori, bab ketiga diisi dengan analisis dan perancangan sistem model yang digunakan, bab keempat diisi dengan implementasi model yang telah dilatih, bab kelima diisi dengan pengujian model, dan pada bab keenam diisi dengan kesimpulan dan saran.

# KAJIAN TEORI

## *Residual Neural Networks*

C*onvolutional Neural Networks (CNN)* merupakan model yang cocok diterapkan dalam kasus klasifikasi, tetapi arsitektur CNN dengan lapisan (*layers)* yang semakin dalam menyebabkan penurunan kinerja yang cukup signifikan karena akan menyebabkan hilang atau membesarnya nilai *gradient* (*vanishing gradient*) [5].

Untuk mengatasi penurunan kinerja ini, Kaiming He, Xiangyu Zhang, Shaoqing Ren, dan Jian Sun pada tahun 2015 menemukan solusi untuk mengatasi penurunan kinerja model CNN dengan menerapkan *skip connection/shortcut* yang kemudian dipublikasikan dengan nama *Residual Neural Networks (ResNet)* [5].

ResNet menjadi pemenang dalam ImageNet Challenge 2015 karena mampu melatih jaringan syaraf sebanyak 152 lapisan dengan baik serta lebih mudah untuk dioptimalkan, dan memperoleh akurasi yang meningkat secara signifikan.

Dalam kerja praktek ini akan menggunakan *pre-trained model Residual Neural Networks 34* atau yang juga disebut ResNet34 yang disediakan oleh *library* FastAi.

## *Learning Rate Finder*

*Learning Rate* merupakan salah satu *hyperparameter* yang penting dan perlu ditentukan sebelum dilakukan proses *training* pada model karena hal ini akan mempengaruhi performa model secara signifikan ketika dilakukan proses *training. Learning rate* yang optimal akan menurunkan tingkat *loss function* sehingga *training* dapat mencapai tingkat akurasi yang optimal.

*Learning rate* yang diatur dengan angka yang sangat kecil, akan memperlambat proses *training* karena memakan waktu yang lama serta sumber daya komputasi yang besar, sedangkan jika diatur dengan angka yang terlalu besar akan menyebabkan tingkat akurasi yang rendah.

  
Gambar 1.1 Learning Rate [10]

Dalam kerja praktek ini akan menggunakan *Cyclical Learning Rates (CLR)* yang dikembangkan oleh Leslie Smith [7] yang dalam implementasi akan menggunakan fungsi yang telah tersedia dalam *library* FastAI. Dalam metode ini, *learning rate* tidak akan ditetapkan secara eksak tetapi akan diperbaharui secara berkala dengan variasi nilai yang telah ditentukan batasan berupa nilai minimum dan maksimumnya.

## *FastAi*

FastAi adalah *library deep learning* yang tersedia secara *open source* dan dapat diinstallsecara langsung dengan menggunakan *Package Manager* bahasa pemrograman Python [6]. FastAi dapat digunakan secara langsung untuk melatih model *deep learning* yang telah disediakan (*pre-trained)* serta memiliki dokumentasi yang mudah untuk dipahami.

Dalam kerja praktek ini, FastAi digunakan sebagai library utama untuk mengakses dan melatih pre-trained model ResNet34 untuk klasifikasi penyakit malaria.

## *Flask*

Flask merupakan *micro web* *framework* yang ditulis menggunakan bahasa pemrograman Python. Flask disebut *Micro Framework* karena Flask hanya menyediakan fungsi dasar dalam pengembangan web sehingga aplikasi web lebih mudah dikembangkan secara sederhana tetapi fungsionalitas lain juga dapat dikembangkan dan fungsionalitas tersebut dapat ditambahkan sebagai *extension* sesuai dengan kebutuhan [8].

Flask yang digunakan dalam kerja praktek ini untuk mengembangkan antarmuka pengguna website dan secara bersamaan dengan bahasa pemrograman HTML dan CSS akan dikoneksikan dengan model ResNet34 yang telah dilatih untuk mendeteksi malaria dari citra digital yang diunggah oleh pengguna.Flask merupakan *micro web* *framework* yang ditulis menggunakan bahasa pemrograman Python. Flask disebut *Micro Framework* karena Flask hanya menyediakan fungsi dasar dalam pengembangan web sehingga aplikasi web lebih mudah dikembangkan secara sederhana tetapi fungsionalitas lain juga dapat dikembangkan dan fungsionalitas tersebut dapat ditambahkan sebagai *extension* sesuai dengan kebutuhan [8].

Flask yang digunakan dalam kerja praktek ini untuk mengembangkan antarmuka pengguna website dan secara bersamaan dengan bahasa pemrograman HTML dan CSS akan dikoneksikan dengan model ResNet34 yang telah dilatih untuk mendeteksi malaria dari citra digital yang diunggah oleh pengguna.

## Google Colaboratory

Google Colaboratory merupakan produk dari Google Research yang memberikan akses kepada pengguna untuk menulis dan menjalankan kode berbasis bahasa pemrograman Python melalui *browser* yang cocok untuk menjalankan kode program seperti machine learning, analisis data serta kebutuhan untuk edukasi [8].

Secara teknis Google Colaboratory memberikan layanan aplikasi Jupyter Notebook bagi pengguna dan memberikan akses gratis ke sumber daya komputasi berupa GPU Nvidia K80s dan TPU yang memiliki batasan *timeout* 12 jam.

Dalam kerja praktek ini, Google Colaboratory digunakan untuk menjalankan kode program untuk melatih dan melakukan analisis terhadap model ResNet34 dan dataset sel darah merah.

## Freelancer – Bootstrap Theme

Freelancer adalah template HTML CSS dengan desain *One Page Website* yang memiliki fitur seperti bagian profil, bagian portofolio berbentuk *grid* yang responsive, *window modal* untuk menampilkan detail setiap item portfolio dan formulir kontak berbasis PHP [10].

Template Freelancer dikembangkan oleh David Miller selaku pemilik Blackrock Digital LLC menggunakan *framework* Bootstrap yang kemudian dirilis dibawah lisensi MIT secara *open source*.

Dalam kerja praktek ini, template Freelancer akan digunakan sebagai struktur tampilan utama *website* dan dilakukan penyesuaian sesuai dengan kebutuhan tampilan *website*.

## Dataset

Dataset yang digunakan dalam kerja praktek ini diambil dari website US National Library of Medicine. Citra digital sel darah diambil menggunakan mikroskop optik medan cahaya yang terpasang di perangkat Android dan diberikan label oleh peneliti Lister Hill National Center for Biomedical Communications (LHNCBC) yang merupakan bagian dari US National Library of Medicine (NLM).

Dataset citra sel darah merah yang telah dikumpulkan sebanyak 27.558 citra digital yang telah dibagi menjadi 2 label yaitu parasitized dan uninfected dengan jumlah citra sel darah merah yang berimbang.

Publikasi dataset yang dipakai dalam kerja praktek ini [9], He et al (2016) membandingkan beberapa pre-trained model Convolutional Neural Networks (CNN) seperti AlexNet, VGG-16, ResNet-50, Xception, DenseNet-121 serta model CNN yang dikustomisasi dengan tiga layer konvolusional (convolutional layers) dan dua layer yang terhubung penuh (fully connected layers). Hasil dari penelitian ini menunjukkan bahwa pre-trained model ResNet-50 memiliki tingkat akurasi paling tinggi.

# ANALISIS DAN RANCANGAN SISTEM

## *Data Preprocessing*

*Data Preprocessing* dataset citra sel darah merah dilakukan dengan membagi dataset citra sel darah merah menjadi 3 bagian yaitu 80% dataset digunakan dalam *training set,* 10% untuk *test set* dan 10% untuk *validation set* yang dipisahkan menjadi 3 *folder* secara acak dan berimbang menggunakan pustaka *split\_folders . https://pypi.org/project/split-folders/*

Dataset ditransformasikan dengan melakukan augmentasi data berupa citra sel darah yang diputar secara vertikal dan horizontal serta dilakukan normalisasi dataset berdasarkan statistika tiga channel RGB dari dataset ImageNet.

Semua proses *preprocessing* dilakukan menggunakan *library fastai* yaitu fungsi *ImageDataBunch* [11] dan citra sel darah merah yang telah dikumpulkan oleh U.S National Library of Medicine [4] telah diberikan label berupa citra sel darah yang telah dipisahkan dalam 2 folder yaitu *parasitized* dan *uninfected*.

## Rancangan Tampilan Website

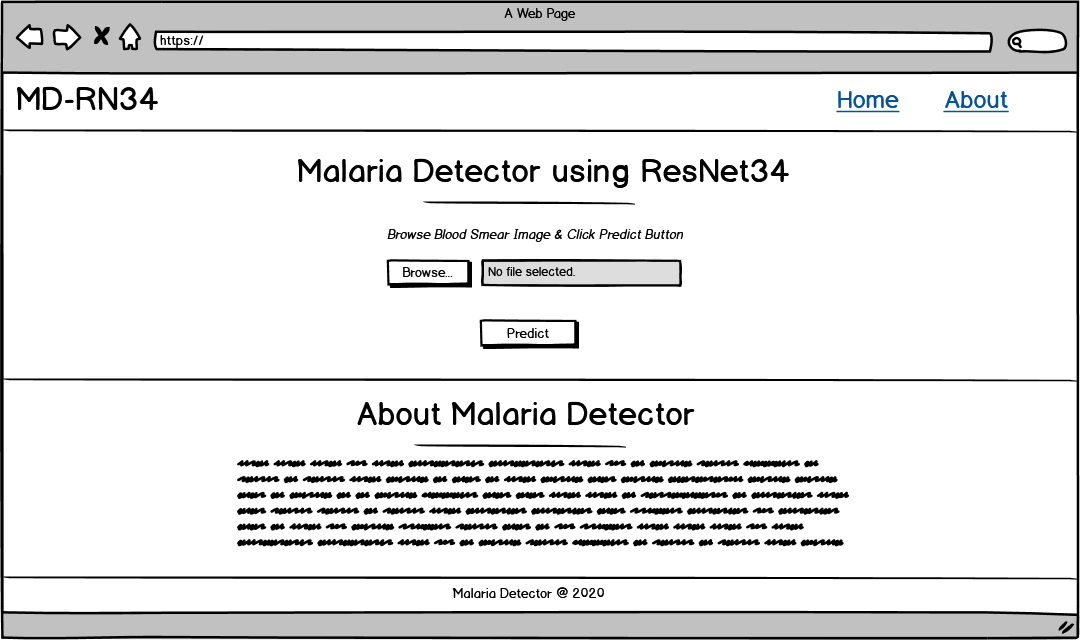
Rancangan tampilan antarmuka pengguna website dibuat sederhana dan menarik untuk mempermudah dalam pengunaan website ini.

## Rancangan Tampilan Utama

Pada Gambar 3.1 menunjukkan rancangan tampilan halaman utama *website* yang terdiri dari bagian untuk fitur utama yaitu klasifikasi citra digital malaria dan bagian *about.*

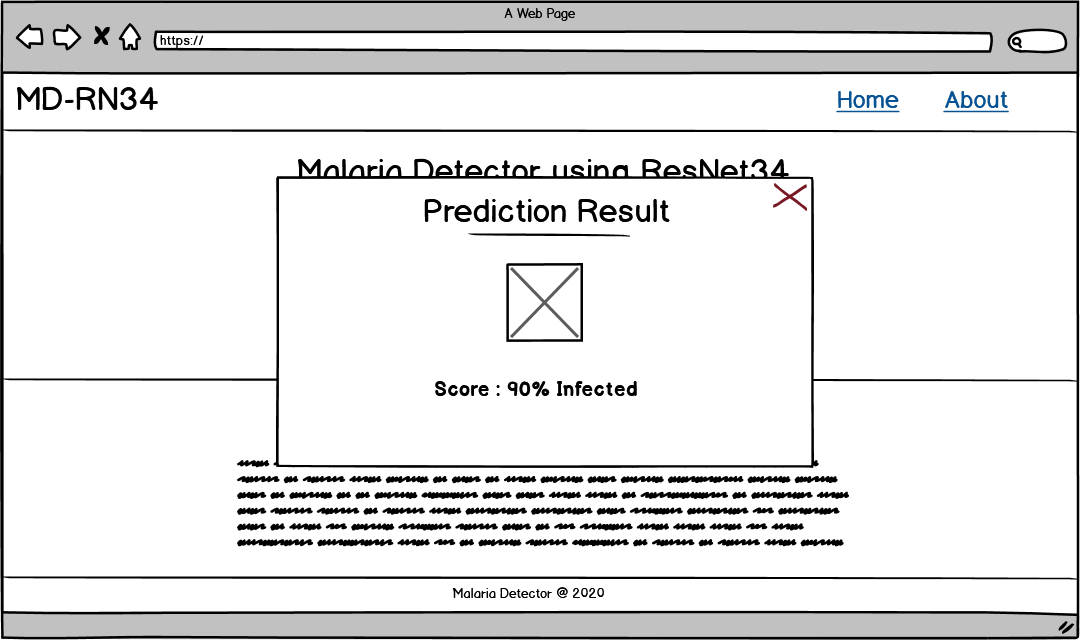
Pada bagian klasifikasi citra digital malaria, pengguna dapat melakukan pengunggahan citra digital sel darah merah kemudian dengan menekan tombol *predict* yang akan memunculkan tampilan *modal window* berupa hasil prediksi yang dilakukan.

Pada bagian *about*, pengguna dapat membaca detail tentang tujuan dari *website* dibuat, penjelasan tentang penyakit malaria, detail tentang algoritma *ResNet34* serta hasil pelatihan model yang digunakan dalam *website*.

  
Gambar 3.1 Rancangan Tampilan Utama

## Rancangan Tampilan *Prediction Result*

Gambar 3.2 menunjukkan hasil dari prediksi model *ResNet34* yang ditampilkan berupa *pop-up window* yang berisi citra digital sel darah yang telah diunggah dan skor hasil prediksi. Jika hasil prediksi menunjukkan kemungkinan terinfeksi lebih besar dari 50% maka *website* akan menampilkan hasil prediksi berupa teks “*XX% infected”* dan jika hasil prediksi menunjukkan kemungkinan tidak terinfeksi lebih besar dari 50% maka website akan menampilkan hasil prediksi berupa teks “*XX% uninfected*”.



Gambar .2 Tampilan Result

# IMPLEMENTASI

## Pelatihan Model

Penjelasan ....

## Implementasi *Backend Website*

Penjelasan ....

## Implementasi Tampilan Website

Penjelasan ....

# PENGUJIAN

## Sub Topik Bab 5

(Style: Report Content). Bab 5 ini berisi pembahasan dan uji coba hasil penelitian yang Anda lakukan. Berupa hasil pengujian aplikasi, jika kerja praktek/ tugas akhir Anda merupakan pengembangan aplikasi. Bisa berupa hanya pembahasan saja jika merupakan audit atau perancangan sistem informasi atau bagian ini bisa dihilangkan apabila sudah dibahas di bab sebelumnya. Sesuaikan sub-bab yang dibutuhkan, dan gunakan sub-heading sesuai dengan yang dibutuhkan.

## Pengujian *Black Box*

## Pengujian *White* *Box*

## *User* *Acceptance* *Test*

## *Benchmarking*

# SIMPULAN DAN SARAN

## Simpulan

Simpulan berisi kesimpulan berdasarkan tujuan penelitian Anda. Jelaskan dalam paragraf pengantar, kemudian langsung berikan dalam butiran-butiran yang menjawab tujuan. Jumlah simpulan diharapkan sama dengan jumlah tujuan penelitian.

## Saran

Saran dapat secara garis besar berisi dua hal yaitu saran untuk pengembangan aplikasi di masa yang akan datang atau saran untuk pelaksanaan penelitian dengan lebih baik di masa yang akan datang.

Melakukan hosting website , melakukan training dengan epoch yg lebih banyak, dan menggunakan arsitektur ResNet dengan layer yang lebih besar atau arsitektur Deep Learning yang lebih baik

# DAFTAR PUSTAKA

|  |  |
| --- | --- |
| [1] | WHO, “Malaria,” [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/malaria. [Diakses 10 Desember 2019]. |
| [2] | Biro Komunikasi dan Pelayanan Masyarakat, Kementerian Kesehatan RI, “Hari Malaria Sedunia, Pemerintah Perluas Wilayah Bebas Malaria,” 28 April 2018. [Online]. Available: http://www.depkes.go.id/pdf.php?id=18043000010. [Diakses 10 Desember 2019]. |
| [3] | Breslauer, David N et al. “Mobile phone based clinical microscopy for global health applications.” *PloS one* vol. 4,7 e6320. 22 Jul. 2009, doi:10.1371/journal.pone.0006320 |
| [4] | National Library of Medicine, “Malaria Datasets” [Online]. Available: https://lhncbc.nlm.nih.gov/publication/pub9932. [Diakses 10 Desember 2019]. |
| [5] | He, Kaiming & Zhang, Xiangyu & Ren, Shaoqing & Sun, Jian. (2016). Deep Residual Learning for Image Recognition. 770-778. 10.1109/CVPR.2016.90. |
| [6] | Howard, Jeremy, and Sylvain Gugger. “Fastai: A Layered API for Deep Learning.” Information 11.2 (2020): 108. Crossref. Web. |
| [7] | L. N. Smith, "Cyclical Learning Rates for Training Neural Networks," 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, 2017, pp. 464-472. |
| [8] | Irsyad, Rahadian. (2018). Penggunaan Python Web Framework Flask Untuk Pemula. 10.31219/osf.io/t7u5r. |
| [9] | https://research.google.com/colaboratory/faq.html |
| [10] | https://github.com/BlackrockDigital/startbootstrap-freelancer |
| [9] | Rajaraman S, Antani SK, Poostchi M, Silamut K, Hossain MA, Maude RJ, Jaeger S, Thoma GR. 2018. “Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images”. PeerJ 6:e4568 |
| [10] | Jeremy Jordan, “Setting the learning rate of your neural network. ,” [Online]. Available: hhttps://www.jeremyjordan.me/nn-learning-rate/. [Diakses 25 Maret 2020]. |
| 11 | fastai, “vision.data | fastai,” [Online]. Available: https://docs.fast.ai/vision.data.html#ImageDataBunch. [Diakses 01 April 2020]. |
|  |  |