

Tugas:

1. Pembuatan model klasifikasi v2 (5 hama 5 penyakit)
2. Riset model deteksi organ tanaman
3. Riset model klasifikasi untuk poin 2

Kendala:

1. Beberapa waktu lalu laptop saya mengalami bad sector sehingga hampir seluruh data yang ada di harddisk terhapus dan hal ini menghambat proses penelitian.
2. Struktur folder dataset dalam storage cloud cukup banyak sehingga dalam pengumpulannya membutuhkan waktu lebih.

Progress:

1. Pengumpulan dataset untuk model v2

Akibat hilangnya data dalam harddisk, saya perlu mengumpulkan ulang data tersebut. Sehingga saya kembali membuka dataset output cropper yang ada di google drive. Setelah itu, saya memutuskan untuk melakukan sedikit perubahan pada struktur penyimpanan data yang ada di cloud yang akan diterapkan terpisah dari penyimpanan awal. Hal ini dilatarbelakangi oleh sulitnya proses mengambil dataset dari output cropper (terlalu banyak folder yang perlu diakses serta penamaan file yang terjadi dua kali atau lebih).

a. Pengambilan ulang dataset dari cloud

Berikut hasil pengambilan ulang dataset dan pengelompokkannya:

	Daun		Buah		Batang		Pohon		Instar Muda		Stadia Dewasa		Stadia Larva		Stadia Pupa		Stadia Telur		Host Alternatif	
	Cabai	Tomat	Cabai	Tomat	Cabai	Tomat	Cabai	Tomat	Cabai	Tomat	Cabai	Tomat	Cabai	Tomat	Cabai	Tomat	Cabai	Tomat	Cabai	Tomat
Late Blight	-	3388	-	236	-	67	-	0	-	-	-	-	-	-	-	-	-	-	-	-
Aphid	0	-	0	-	0	-	0	-	58	-	2590	-	8	-	0	-	0	-	0	-
Antraknose	0	-	3394	-	0	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-
Phytophthora Blight	61	-	0	-	126	-	127	-	-	-	-	-	-	-	-	-	-	-	-	-
Bacterial Spot	1138	1899	0	1501	0	19	0	0	-	-	-	-	-	-	-	-	-	-	-	-
Whitefly	0	0	0	0	0	0	0	0	0	0	1544	1378	4	831	0	0	0	0	206	0
Helicoverpa Armigera	0	0	139	1121	0	0	0	0	0	0	25	2	1434	1434	9	8	0	0	0	0
Lalat Buah	0	0	1235	2	0	0	0	0	-	1172	-	316	-	43	-	0	-	46	-	-
Tungau	1378	101	0	0	0	0	0	0	-	1	-	0	-	0	-	0	-	0	-	-
Ulat Grayak	0	7	7	199	0	0	0	0	-	0	-	13	-	2281	-	40	-	3	-	12
Thrips	678	0	0	0	0	0	0	0	0	0	1917	597	9	0	0	0	0	0	0	0
Bacterial Wilt	4	905	0	0	0	0	2558	466	-	-	-	-	-	-	-	-	-	-	-	-
Early Blight	75	2390	0	17	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-
Sehat	2607	2462	4107	2012	0	3	804	0	-	-	-	-	-	-	-	-	-	-	-	-
Begomovirus	3834	2383	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	-

Figure 1. Hasil Pengelompokkan Dataset

Pengelompokkan yang saya coba berasal dari tujuan untuk meningkatkan model serta membandingkan dengan dataset benchmark yakni PlantVillage. Dataset plantvillage yang dijadikan acuan terdiri atas data gambar yang seluruhnya adalah gambar gejala penyakit pada daun. Mengambil ide plantvillage, maka saya terpikir untuk mengelompokkan kelas klasifikasi berdasarkan organ tanaman dan didapatkan seperti pada gambar di atas. Selain dari itu, dengan membagi pengelompokkan berdasarkan organ, kita dapat memantau secara rinci perkembangan dataset lebih dalam lagi.

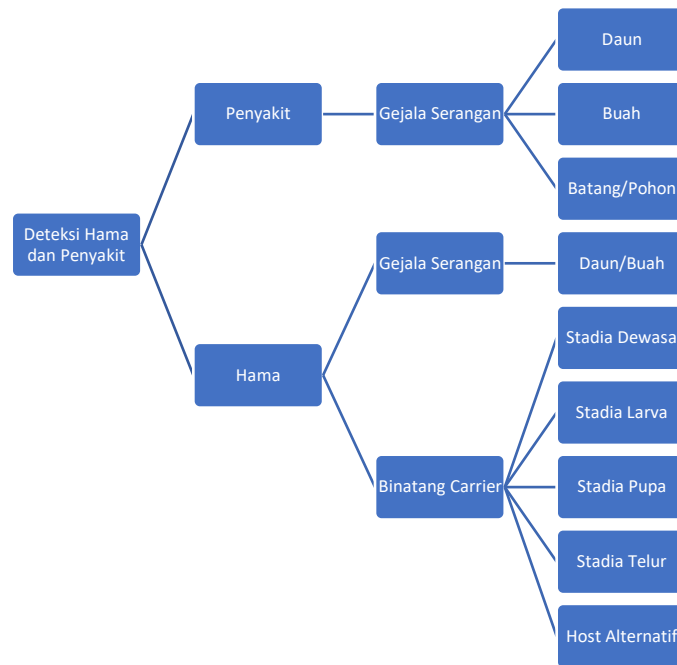


Figure 2 Tujuan Pengelompokkan Dataset

Dengan penetapan tahap penelitian kedua untuk deteksi hama yakni menggunakan 5 hama dan 5 penyakit yakni:

		Infeksi	
		Cabai	Tomat
Penyakit	Antraknose	✓	☒
	Phytophthora Blight	✓	☒
	Late Blight	☒	✓
	Early Blight	☒	✓
	Bacterial Spot	✓	✓
	Bacterial Wilt	✓	✓
	Begomovirus	✓	✓
Hama	Aphid	✓	☒
	Lalat Buah	✓	☒
	Thrips	✓	✓
	Ulat Grayak	✓	✓
	Tungau	✓	✓
	Whitefly	✓	✓
	Helicoverpa Armigera	✓	✓

Figure 3 Korekasi antara Penyakit, Hama dan Objek Infeksi

Perlu untuk ditentukan dari manakan deteksi akan dilakukan untuk setiap penyakit dan hama. Karena untuk sebaran data ada yang bercampur untuk gejala serangan di daun, di buah sehingga fitur dalam gambar yang dimasukkan ke dalam CNN berpotensi menyulitkan proses learning model.

Sehingga, jika menggunakan dataset saat ini deteksi hama dan penyakit dari organ tanaman seperti ini :

Sumber Deteksi	Kondisi Sehat	Penyakit							Hama				
		Phytophthora Blight	Bacterial Spot	Bacterial Wilt	Early Blight	Begomovirus	Late Blight	Antraknose	Ulat Gray	Helicoverpa Armigera	Lalat Buah	Tungau	Thrips
Daun Cabai	✓	✓	✓	✓	✓	✓						✓	✓
Daun Tomat	✓		✓	✓	✓	✓	✓		✓			✓	
Buah Cabai	✓							✓	✓	✓	✓		
Buah Tomat	✓		✓		✓		✓		✓	✓	✓		
Batang Cabai		✓											
Batang Tomat	✓		✓				✓						

Figure 4. Korelasi Sumber Deteksi dan Availability Prediksi Penyakit

Sementara untuk hama seperti ini:

Cabai							
Sumber Deteksi	Aphid	Whitefly	Helicoverpa Armigera	Lalat Buah	Tungau	Ulat Grayak	Thrips
Instar Muda	✓						
Stadia Dewasa	✓	✓	✓	✓	✓		✓
Stadia Larva	✓	✓	✓	✓			✓
Stadia Pupa			✓	✓			
Stadia Telur							
Host Alternatif		✓		✓			

Tomat							
Sumber Deteksi	Aphid	Whitefly	Helicoverpa Armigera	Lalat Buah	Tungau	Ulat Grayak	Thrips
Instar Muda							
Stadia Dewasa		✓	✓			✓	✓
Stadia Larva		✓	✓			✓	
Stadia Pupa			✓			✓	
Stadia Telur						✓	
Host Alternatif						✓	

Figure 5. Korelasi Sumber Deteksi dan Availability Prediksi Hama

Prakiraan pada gambar-gambar di atas diambil dengan melihat setidaknya ada satu gambar yang dikumpulkan sehingga, meski memungkinkan perlu dilihat juga jumlah gambar yang tersedia di dalamnya.

b. Hasil Pengambilan Ulang Dataset

Pengambilan ulang dataset dimulai dengan mengumpulkan fitur dari gambar daun. Kemudian dipilih kategori dengan jumlah kelas lebih dari 1000 gambar. Berikut hasil yang telah dikumpulkan.

Nama	Dataset Partial V1			
	Daun Cabai	Train	Validation	Test
Bacterial Spot	1138	1000	69	69
Tungau	1378	1000	189	189
Sehat	2607	1000	804	803
Begomovirus	3834	1000	1417	1417

Nama	Dataset Partial V2			
	Daun Tomat	Train	Validation	Test
Late Blight	3388	1000	1194	1194
Bacterial Spot	1899	1000	450	449
Early Blight	2390	1000	695	695
Sehat	2462	1000	731	731
Begomovirus	2383	1000	692	691

Figure 6. Komposisi Pengumpulan Dataset

c. Proposal Untuk Update Cropper

Untuk menghindari kontraproduktif terhadap prosedur yang sudah ada untuk update data cropper dalam folder [Cropper](#). Tim cropper dapat melakukan update juga pada file zip dalam folder [CropperV2](#). Dalam folder tersebut terdapat file zip untuk setiap tim cropper seperti gambar di bawah ini.

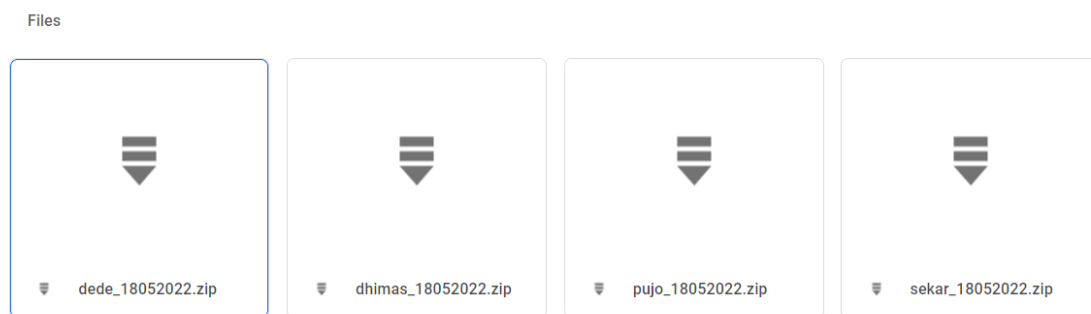


Figure 7. File Zip Tim Cropper

Tim cropper dapat mengunduh setiap miliknya dan menambahkan gambar baru setelah preprocessing tanpa menambahkan folder baru melainkan sesuai dengan nama kelasnya, seperti gambar berikut ini.

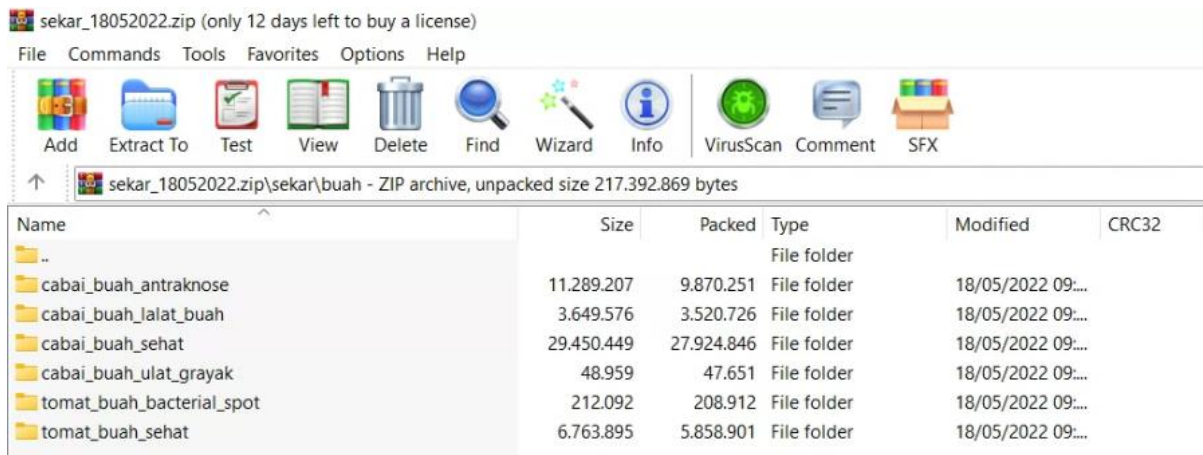


Figure 8. Kelompok Folder Baru

Setelah menambahkan gambar, tim cropper dapat mengupdate nama dataset sesuai dengan tanggal update dan mengganti filze zip yang ada di google drive dengan file zip yang telah di update.

2. Riset Model Deteksi Organ Tanaman

Model deteksi organ tanaman dilakukan untuk menunjukkan proses klasifikasi hama pada tanaman. Deteksi organ tanaman diharapkan dapat membantu mendapatkan ROI spesifik yang berpotensi adanya gejala penyakit atau kerusakan. Dataset yang digunakan berasal dari 100 gambar buah cabai/tomat yang sehat dan 100 gambar cabai/tomat yang memiliki gejala serangan. Kemudian dataset tersebut digabung sehingga model dapat mendeteksi bentuk buah cabai/tomat meski dalam kondisi sehat atau memiliki gejala serangan.

Class	Train	Validation	Test
Buah Tomat	160	20	20
Buah Cabai	160	20	20

Figure 9. Dataset Model Deteksi Organ Tanaman

Model deteksi dibangun dengan menggunakan framework YOLOv4 menggunakan metode transfer learning dari pretrained model yolo conv 137. Pada training pertama dilakukan dengan 4000 epoch dan mampu mendapat mAP 42.6%.

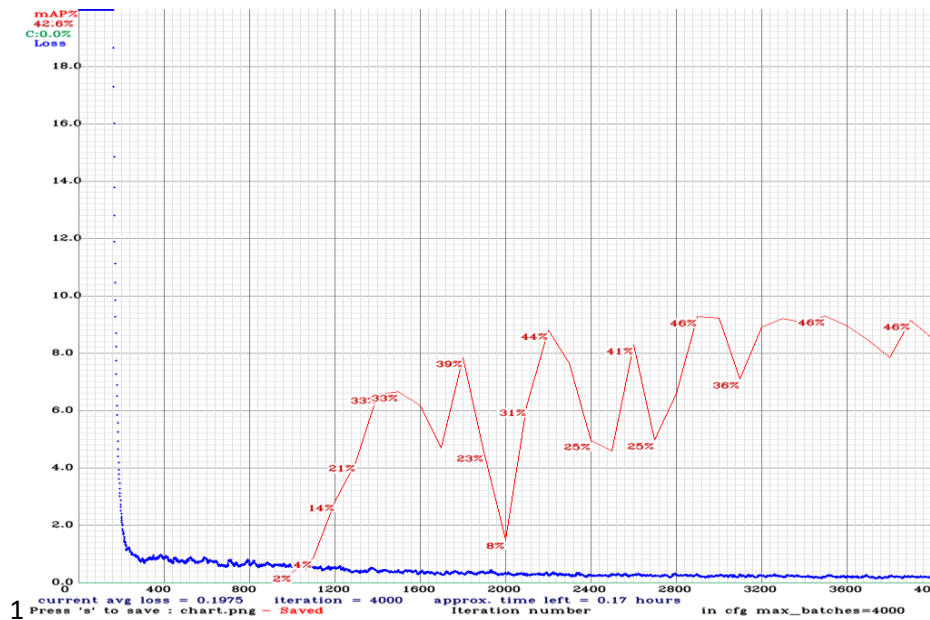


Figure 10. Learning Curve YOLOv4 Training 1

Uji training selanjutnya dilakukan dengan menambahkan epoch lebih banyak yakni 6000 mengikuti best practice yang biasa dilakukan developer lain dan hanya mampu mendapatkan mAP 39.5%.

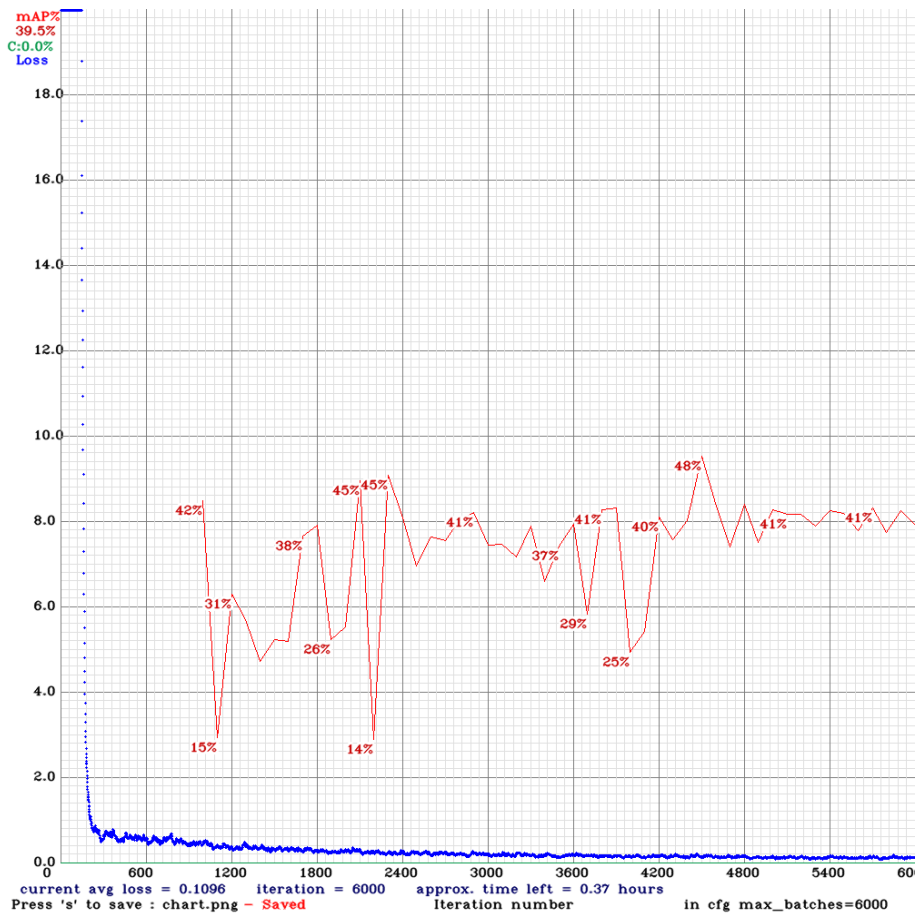


Figure 11. Learning Curve YOLOv4 Training 2

Model ini kemudian dievaluasi menggunakan validation dataset dan mendapatkan f1 score sebesar 0.82 dan f1score dengan test dataset sebesar 0.93.

```
class_id = 0, name = Buah Tomat, ap = 0.00%      (TP = 0, FP = 1)
class_id = 1, name = Buah Cabai, ap = 85.28%     (TP = 16, FP = 1)

for conf_thresh = 0.25, precision = 0.89, recall = 0.76, F1-score = 0.82
for conf_thresh = 0.25, TP = 16, FP = 2, FN = 5, average IoU = 66.63 %

IoU threshold = 50 %, used Area-Under-Curve for each unique Recall
mean average precision (mAP@0.50) = 0.426386, or 42.64 %
Total Detection Time: 1 Seconds
```

Figure 12. Hasil Evaluasi Validation Dataset Training 2

```
class_id = 0, name = Buah Tomat, ap = 0.00%      (TP = 0, FP = 1)
class_id = 1, name = Buah Cabai, ap = 93.88%     (TP = 19, FP = 0)

for conf_thresh = 0.25, precision = 0.95, recall = 0.90, F1-score = 0.93
for conf_thresh = 0.25, TP = 19, FP = 1, FN = 2, average IoU = 69.46 %

IoU threshold = 50 %, used Area-Under-Curve for each unique Recall
mean average precision (mAP@0.50) = 0.469388, or 46.94 %
Total Detection Time: 1 Seconds
```

Figure 13. Hasil Evaluasi Test Data Training 2

Dapat dilihat bahwa class id 0 memiliki jumlah prediksi benar 0, dan hal ini masih dianalisis. Akan tetapi, terdapat galat yang terjadi disepanjang training dua model tersebut yakni ekstensi file yang bermasalah.

Failed to infer label file name (check image extension is supported)

Asked 7 months ago Modified 7 months ago Viewed 232 times

I am working on an object detection project with a custom dataset. The images have the format of .tif. the model I work with to train my model is **YOLOv3**. I labeled all images with [Labelimg](#) application. The objects should be categories into 4 different classes. I use google collab and google drive [Google Drive folders](#). Each folder contains:

- backup ==> the weights of my model will save in this folder after training the darknet.
- custom_weight ==> pre-train weight of YOLOv3 (darknet53.conv.74)
- darknet ==> darknet folder (I costomized the **Makefile** file and in the cfg folder **yolov3.cfg**)

The Overflow Blog

- Open-source is winni and investors (Ep. 44;
- Stack under attack: w handling DDoS attac

Featured on Meta

- Retiring Community- Based for Server Fa

Figure 14. Warning Saat Training YOLOv4

Karena terdapat warning tersebut pada saat training, saya memutuskan untuk merapikan seluruh penamaan file dalam dataset yang diharapkan ke depannya dataset dapat memiliki nama yang tidak bertumpuk seperti ("6. Antraknose(1)_gejalaserangan.jpg.jpg").

a. Proposal Untuk Anotasi Bounding Box

Dengan berjalannya riset untuk deteksi objek, maka diperlukan solusi untuk proses anotasi bounding box untuk gambar. Karena proses anotasi merupakan proses yang sangat menyita waktu.

Rencana Minggu Depan:

1. Agenda selanjutnya adalah setelah mengumpulkan kembali dataset diharapkan dataset partial v1 dan v2 dapat dijadikan data untuk model klasifikasi dengan hasil yang lebih baik.
2. Melakukan anotasi ulang setelah menambahkan dataset deteksi organ tanaman.