Implementasi Algoritma CNN untuk Melakukan Klasifikasi pada Daun Anggur

Salsabila Oktafani¹, Alvin Putra Perdana², Jihan Kamilah³, Nurhikmah Mawaddah Solin⁴

Program Studi Informatika, Universitas Pembangunan Nasional Veteran Jakarta

¹2010511001@mahasiswa.upnvj.ac.id ²2010511011@mahasiswa.upnvj.ac.id ³2010511013@mahasiswa.upnvj.ac.id ⁴2010511026@mahasiswa.upnvj.ac.id

Abstract—Tanaman anggur memiliki berbagai macam jenis. Hal itu dapat dibedakan melalui masing-masing karakteristik dari daun anggur tersebut. Adapun algoritma klasifikasi pada citra yang digunakan ialah CNN (Convolutional Neural Network) dengan arsitektur yang digunakan berupa ResNet50 dan VGGNet19. Data citra yang digunakan pada penelitian ini sebanyak 500 data citra yang termasuk dalam lima jenis kelas yaitu Ak, Ala_Idris, Buzgulu, Dimnit dan Nazli, dan pada tiap jenisnya memiliki 100 citra. Split data dilakukan menggunakan perbandingan 80:20, dengan 80% untuk data training dan 20% untuk data validasi. Penelitian ini mendapatkan hasil evaluasi sebesar 85% pada salah satu model arsitektur yang digunakan.

Kata Kunci—CNN, ResNet50, VGGNet19, Klasifikasi, Anggur.

Abstract—There are various types of grapes. It can be distinguished through each of the characteristics of the grape leaf. The classification algorithm on the image used is CNN (Convolutional Neural Network) with the architecture used in the form of ResNet50 and VGGNet19. The image data used in this study is 500 image data which is included in five types of classes, namely Ak, Ala_Idris, Buzgulu, Dimnit and Nazli, and in each type has 100 images. Split data is done using a ratio of 80:20, with 80% for training data and 20% for validation data. This study obtained an evaluation result of 85% on one of the architectural models used.

Keywords—CNN, ResNet50, VGGNet19, Classification, Grape.

I. PENDAHULUAN

Buah dengan jenis tanaman merambat ini termasuk dalam keluarga Vitaceae. Anggur menjadi salah satu jenis buah yang memiliki nutrisi tinggi. Dapat di konsumsi secara langsung ataupun dilakukan pengolahan terlebih dahulu. Banyak kandungan manfaat dari anggur ini, seperti karbohidrat, serat, lemak, protein, vitamin B1, B2, B3, B5, B6, B9, C, E, dan K, Kalsium, serta masih banyak kandungan lainnya. Anggur juga memiliki banyak khasiat untuk kesehatan, seperti membantu mengurangi masalah sesak nafas, dapat melindungi tubuh dari berbagai infeksi, meningkatkan kesehatan otak manusia, mencegah penyakit kanker payudara, dan sebagainya.

Tak hanya buahnya, daun dari tanaman Vitis vinifera ini juga tak kalah bermanfaat. Bisa dijadikan masker wajah, diolah menjadi makanan, dan lainnya. Adapun kandungan dari daun anggur seperti, karbohidrat, protein, vitamin, dan

kandungan lainnya. Buah anggur tentu memiliki berbagai macam jenis. Hal ini bisa dibedakan dengan bentuk daun yang berbeda. Anggur memiliki berbagai macam bentuk daun dengan karakteristik yang berbeda. Apabila dilihat sekilas, mungkin terlihat sama. Namun, apabila diidentifikasi dari bentuk dan karakter setiap daun, anggur memiliki jenis dan varian daun yang berbeda satu sama lain.

Akibat perkembangan teknologi, membedakan jenis buah anggur melalui karakteristik daunnya bisa menjadi lebih mudah. Untuk membedakan daun anggur ini, dapat menggunakan metode machine learning. Machine learning itu sendiri ialah ilmu yang mempelajari bagaimana komputer bisa belajar dari data guna meningkatkan kecerdasannya. Di penelitian ini, peneliti menggunakan metode Supervised Learning, dengan algoritma CNN (Convolutional Neural Network) dengan arsitektur VGGNet19 dan ResNet50. Dengan menggunakan bantuan dari beberapa library atau package yakni pandas ataupun numpy untuk mengatasi load dataset, keras untuk menangani praproses image, tensorflow untuk membangun model arsitektur yang digunakan, matplotlib.pyplot dan sns untuk visualisasi data, serta sklearn untuk evaluasi model yang telah dibuat. Untuk pengukuran evaluasi, peneliti menggunakan nilai loss, accuracy, precision, dan recall. Dalam penelitian, ini hasil yang diharapkan yaitu mampu melakukan klasifikasi jenis anggur berdasarkan citra bentuk daunnya dengan melakukan perbandingan dari dua arsitektur yang berbeda yakni antara ResNet50 dengan VGGNet19, sehingga akan didapatkannya arsitektur dengan nilai evaluasi terbaik.

II. STUDI PUSTAKA

A. Penelitian Terkait

Penelitian [1] tahun 2020 pada melakukan pengklasifikasian penyakit pada daun padi agar tak menular ataupun menjadi semakin parah, dengan menggunakan algoritma CNN (Convolutional Neural Network). Pada penelitian ini citra dibagi menjadi tiga kategori yakni Bacterial Leaf Blight, Brown Spot dan Leaf Spot, serta digunakannya parameter size 100x100 pixel, kernel size 3x3, epochs 150, learning rate 0,01, varian optimizer-nya adalah Adam, batch size 30, dan perbandingan dataset 90%: 10% pada citra RGB. Hasil yang didapatkan ialah tingkat akurasi sebesar 91,7%.

Penelitian [2] pada tahun 2020 melakukan pengklasifikasian citra daun mangga pada 1761 citra dengan

digunakannya algoritma CNN (*Convolutional Neural Network*). Pada penelitian ini citra dibagi menjadi tiga jenis yakni mangga golek, mangga harum manis, dan mangga manalagi. Serta, digunakannya perbandingan 9 data *training*: 1 data validasi, 4 lapisan konvolusi disertai *maxpooling*, dan 60 epoch. Hasil yang didapatkan ialah tingkat akurasi pada *training set* sebesar 97,72%, dan pada *validation set* sebesar 89,2%.

Penelitian [3] tahun 2021 melakukan pengklasifikasian citra daun pepaya pada 150 citra menggunakan algoritma Naive Bayes dan ekstraksi ciri LBP (*Local Binary Pattern*). Pada penelitian ini citra dibagi menjadi lima jenis yakni daun pepaya Sumatera, Hawai, Cibinong, Bangkok dan California. Seta, dilakukannya pembagian data *training* sebanyak 125 citra, data *testing* sebanyak 25 citra, *resize* citra menjadi 256x256 *pixel*, ekstraksi ciri menggunakan LBP dengan tiap citranya mempunyai 10 fitur. Hasil yang didapatkan ialah tingkat akurasi sebesar 96% pada percobaan pertama, dan 93% pada percobaan kedua.

Penelitian [4] tahun 2022 melakukan pengklasifikasian tingkat kematangan pucuk daun teh varietas Assamica klon GMB 7, menggunakan algoritma CNN (*Convolutional Neural Network*) dengan arsitektur VGGNet19 dan ResNet50. Pada penelitian ini citra dibagi menjadi empat kelas 7 hari, 14 hari, 21 hari dan 28 hari. Serta, tiap percobaan dilakukan pembagian menjadi 80% data *training* dan 20% data *testing*. Hasil yang didapatkan ialah penggunaan arsitektur VGGNet19 menghasilkan performa terbaik dengan akurasi sebesar 97,5% , loss 0.1208, dan precision 97%.

Penelitian [5] tahun 2022 melakukan pengklasifikasian jenis daun Aglaonema menggunakan algoritma CNN (Convolutional Neural Network). Pada penelitian ini menggunakan lima jenis yakni Commutatum, Costatum, Deborah, King of Siam, Snow White, dan terdapat 100 citra di tiap jenis. Serta, dilakukannya resize citra menjadi 227x227 pixel, pembagian data menjadi 80% data training: 20% data testing, dan digunakannya Adam optimizer, dan Stochastic Gradient Descent with Momentum (SGDM) optimizer pada tahap training dengan epochs 15, fungsi aktivasi relu, learning rate 0,0001. Hasil yang didapatkan ialah tingkat akurasi rata-rata sebesar 96%.

B. Tinjauan Pustaka

1. Machine Learning

Pembelajaran mesin atau yang sering disebut dengan *Machine Learning* menjadi bidang untuk membangun, mempelajari, dan memahami suatu data dengan menerapkan metode "*learn*" yaitu belajar. Istilah *Machine Learning* dikemukakan oleh Arthur Samuel pertama kali pada tahun 1959. *Machine Learning* menjadi cabang dari kecerdasan buatan yang difokuskan ke data dan algoritma untuk meniru kinerja belajarnya manusia. Dimana algoritma-algoritma dilatih untuk membuat suatu klasifikasi ataupun prediksi.

Machine Learning mempelajari bagaimana komputer bisa belajar dari data guna meningkatkan kecerdasannya. Ada beberapa kategori dalam machine learning, yaitu Supervised Machine Learning, Unsupervised Machine Learning, dan Semi-supervised Learning.

2. Supervised Learning

Supervised learning menjadi algoritma machine learning dengan menggunakan data latih, yang artinya data latih tersebut memiliki data yang sudah berlabel. Dengan tujuan agar mesin bisa melakukan identifikasi label input dengan menggunakan Fitur yang nanti akan digunakan untuk membuat atau melakukan klasifikasi dan prediksi. Adapun contoh implementasi supervised learning seperti deteksi spam pada email.

3. Convolutional Neural Network (CNN)



Gambar. 1 Ilustrasi Algoritma CNN

Convolutional Neural Network (CNN) menjadi salah satu arsitektur deep learning yang popular guna untuk melakukan object recognition ataupun object classification. CNN menjadi bagian dari ANN yang banyak bekerja pada bidang Citra Visual. CNN menjadi pengembang dari Multi Layer Perceptron (MLP). CNN bisa digunakan untuk melakukan klasifikasi gambar, segmentasi, dan deteksi object. Contoh penggunaan CNN yang membuat popular, ketika digunakan pada mobil Nvidia, yang tentunya butuh suatu proses training.

a. ResNet50

Residual Network, yang sering disingkat sebagai ResNet sering digunakan untuk melakukan klasifikasi pada gambar. Maksud dari 50 pada ResNet 50 yaitu memiliki lapisan layer sebanyak 50. ResNet 50 dianggap menjadi arsitektur CNN yang baik, dibuktikan pada penelitian penyakit otak dengan menggunakan gambar MRI oleh Talo et al pada tahun 2019. [6]

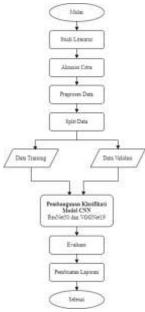
b. VGGNet19

VGGNet19 termasuk arsitektur jaringan yang tercipta dari *Visual Geometry Group*, bentuk dari variasi VGGNet16. VGGNet memiliki lapisan bobot sejumlah 19, dengan 16 lapisan konvolusi dimana 3 lapisan terhubung penuh dan 5 lapisan penyatuan yang sama. [7]

III. METODE PENELITIAN

A. Kerangka Berpikir

Berdasarkan penelitian terkait, tinjauan penelitian, dan permasalahan yang telah dijelaskan. Kerangka berpikir dari penelitian digambarkan melalui bagan dibawah ini:



Gambar. 2 Kerangka Berpikir Penelitian

B. Instrumen Penelitian

Berikut alat bahan perangkat yang digunakan untuk melakukan penelitian dalam melakukan klasifikasi pada daun anggur:

1. Bahan

Dalam penelitian ini, bahan yang digunakan adalah 500 data citra daun anggur. Yang terdapat lima jenis yaitu Ak, Ala_Idris, Buzgulu, Dimnit, Nazli, dan pada tiap jenisnya memiliki 100 citra.

Perangkat yang digunakan dan dibutuhkan dalam penelitian ini diantaranya :

a. Perangkat Lunak

Google Colab

Perangkat lunak Google Colab dipakai untuk menginput dan menjalankan kode program tanpa instalasi teks editor.

Google Document

Digunakan sebagai tempat mengetik laporan secara bersama dalam satu waktu.

Google Drive

Digunakan sebagai tempat menyimpan dataset citra dari Kaggle.

• Microsoft Office Word

Ms. Word sebagai tempat untuk melakukan *finishing* pada laporan.

b. Perangkat keras

• Laptop dengan spesifikasi RAM 12GB

- AMD Quad-Core Processor FX-9800P 3.60GHz
- Mouse

C. Pengumpulan Data

Dalam penelitian ini pengumpulan data dilakukan dengan beberapa cara yaitu :

1. Studi Pustaka

Pada tahapan ini dilakukan guna mendapatkan penjelasan terkait teori-teori yang berhubungan dengan topik penelitian.

2. Open Dataset

Pada penelitian ini citra daun anggur yang digunakan didapatkan dari platform Kaggle, dengan nama dataset Grapevine Leaves Image Dataset yaitu melalui link berikut : https://www.kaggle.com/datasets/muratkokludataset/grapevine-leaves-image-dataset?resource=download.

D. Teknik Analisis Data

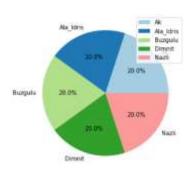
Dalam penelitian ini, teknik analisis data yang digunakan untuk data tersebut adalah dengan menggunakan dua arsitektur CNN, antara lain:

- 1. ResNet50
- 2. VGGNet19

IV. HASIL PENELITIAN DAN PEMBAHASAN

A. Data yang Digunakan

Data yang digunakan dalam penelitian ini terdiri dari 500 citra untuk setiap lima kelas jenis daun anggur, yaitu: Ak, Ala_Idris, Buzgulu, Dimnit dan Nazli. Sehingga, tiap kelas dari daun anggur tersebut mengandung 100 buah citra daun.



Gambar. 3 Diagram Distribusi tiap Kelas

Berikut adalah salah satu contoh citra yang mewakili tiap kelasnya.



Gambar. 4 Citra Daun Anggur Kelas Ak



Gambar. 5 Citra Daun Anggur Kelas Ala_Idris



Gambar. 6 Citra Daun Anggur Kelas Buzgulu



Gambar. 7 Citra Daun Anggur Kelas Dimnit



Gambar. 8 Citra Daun Anggur Kelas Nazli

Setelah itu, data akan di split dengan perbandingan 80:20. Pada data *training* atau data latih jumlahnya sebesar 80% dan untuk data validasi atau data uji sebesar 20%. Jika dengan angka, data training sebanyak 400 citra dan data validasi sebanyak 100 citra.

B. Perhitungan Algoritma

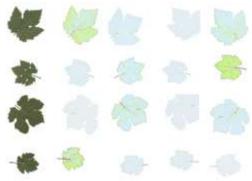
- 1. Praproses Citra
 - a. Membuat *data frame* berdasarkan lokasi tempat penyimpanan dari dataset. Lokasi penyimpanan tersebut berada di folder google drive.



Gambar. 9 Proses Pembuatan Dataframe

 Melakukan pengecekan apakah terdapat missing value pada dataset dengan menggunakan fungsi isna()

- c. Melakukan augmentasi pada masing-masing citra dengan menggunakan *library image generator* dari Keras, antara lain:
 - rotation_range=20, Citra akan dirotasi maksimal sebanyak 20%.
 - zoom_range=0.10, Citra akan di zoom.
 - brightness_range=[0.6,1.4], Tingkat kecerahan citra akan diubah di rentang 0,6 sampai 1,4.
 - *channel_shift_range*=0.7, Saturasi warna pada citra akan diubah.
 - width_shift_range=0.15, Citra akan digeser ke kanan.
 - height_shift_range=0.15, Citra akan digeser ke atas atau ke bawah
 - *shear_range*=0.15, Sudut dari citra akan digeser.
 - *horizontal_flip=*True, Citra akan di *flip* secara horizontal.



Gambar. 10 Citra Daun Anggur setelah Proses Augmentasi

2. Model Convolutional Neural Network

a. ResNet50

Pada tahap *splitting data*, jumlah untuk data *training* sebanyak 400 citra dan untuk data validasi sebanyak 100 citra.

Selanjutnya, untuk menggunakan arsitektur ResNet50 maka argumen yang ditentukan adalah sebagai berikut :

- include _top = false
- $Input_shape = (227,227,3)$
- Weight = imagenet
- Layer.trainable = false
- Pooling = average
- $Input_layer = 5$
- Activation = softmax

b. VGGNet19

- *include* _top = false
- $Input_shape = (227,227,3)$
- Weight = imagenet
- Layer.trainable = false
- Pooling = average
- $Input_layer = 5$
- Activation = softmax

C. Hasil Akurasi

Kedua arsitektur ResNet50 dan VGGNet19 menggunakan parameter pelatihan yang sama guna mendapatkan salah satu model yang lebih optimal. Parameter yang digunakan yaitu parameter optimizer Adam, dan untuk fungsi loss yang digunakan yaitu categorical crossentrophy. Tahapan training data menggunakan epochs sebanyak 50 dan secara langsung menggunakan data validasi untuk mengetahui hasil dari kedua model tersebut. Evaluasi model yang ditentukan adalah loss, accuracy, precision, recall, val_loss, val_accuracy, val_precision, val_recall.

$$\begin{aligned} Accuracy &= \frac{(TP + TN)}{(TP + FP + FN + TN)} \\ Precision &= \frac{TP}{(TP + FP)} \\ Recall &= \frac{TP}{(TP + FN)} \end{aligned}$$

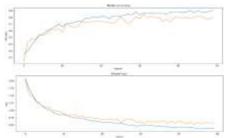
- TP (True Positive Correct result)
- TN (True Negative Correct absence of result)
- FP (False Positive Unexpected result)
- FN (False Negative Missing result)

Model ResNet50 memiliki total parameter sebanyak 23,597,957 sedangkan VGGNet19 memiliki total parameter sebanyak 20,026,949.

a. ResNet50

Berdasarkan hasil yang telah didapatkan setelah menjalankan arsitektur ResNet50, nilai *loss* yang didapatkan sebesar 0.5893 pada *epochs* terakhir dan dan rata-rata dari keseluruhan tahap *training data* sebesar 0.5164.

Hasil lain yang didapatkan pada tahap epochs terakhir adalah *accuracy, precision dan recall*. Secara berturut-turut, hasil yang didapatkan adalah *accuracy* sebesar 0.7700, *precision* sebesar 0.8214 dan *recall* sebesar 0.7200 dan hasil akurasi evaluasi adalah 0.8500, *precision* sebesar 0.8929, *recall* sebesar 0.7500.



Gambar. 11 Grafik Nilai Evaluasi Arsitektur ResNet50

Berikut adalah tampilan grafik dari hasil akurasi dan *loss model training* data menggunakan arsitektur ResNet50.

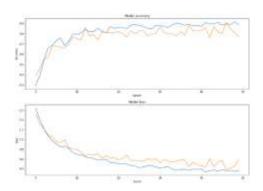
TABLE I. NILAI EVALUASI ARSITERTUR RESNET50

Step/Architecture	ResNet50	
	Last Epochs	Evaluasi Model
Durations (s)	84	17
Loss	0,3641	0,5164
Akurasi	0,8900	0,8000
Precision	0,9254	0,9024
Recall	0,8375	0,7400
Val Loss	0,5893	0,5074
Val Akurasi	0,7700	0,8500
Val Precision	0,8214	0,8929
Val Recall	0,7200	0,7500

Berdasarkan hasil yang telah didapatkan setelah menjalankan arsitektur VGGNet19, nilai *loss* yang didapatkan sebesar 0.5363 pada *epochs* terakhir dan dan rata-rata dari keseluruhan tahap *training data* sebesar 0.4973.

Hasil lain yang didapatkan pada tahap *epochs* terakhir adalah *accuracy*, *precision* dan *recall*. Secara berturut-turut, hasil yang didapatkan adalah *accuracy* sebesar 0.8100, *precision* sebesar 0.8571 dan *recall* sebesar 0.7200 dan hasil akurasi evaluasi adalah 0.8000, *precision* sebesar 0.8295, *recall* sebesar 0.7300.

Berikut adalah tampilan grafik dari hasil akurasi dan *loss model training data* menggunakan arsitektur VGGNet19.



Gambar. 12 Grafik Nilai Evaluasi Arsitektur VGGNet19

TABLE II. NILAI EVALUASI ARSITERTUR VGGNET19

Step/Architecture	VGG19	
	Last Epochs	Evaluasi Model
Durations (s)	322	65
Loss	0,3872	0,4974
Akuras	0,9075	0,8300
Precision	0,9224	0,8736
Recall	0,8325	0,7600

Step/Architecture	VGG19	
	Last Epochs	Evaluasi Model
Val Loss	0,5363	0,5597
Val Akurasi	0,8100	0,8000
Val Precision	0,8571	0,8295
Val Recall	0,7200	0,7300

V. KESIMPULAN

Penelitian klasifikasi daun anggur menggunakan algoritma CNN dengan model arsitektur ResNet50 dan VGGNet19. Didapatkannya hasil evaluasi dengan menggunakan arsitektur ResNet50 yaitu 85%. Sedangkan, dengan menggunakan arsitektur VGGNet19 didapatkan hasil evaluasi 80%. Dilihat dari dua arsitektur yang telah digunakan pada penelitian ini, model arsitektur ResNet50 lah yang menghasilkan performa sistem terbaik.

VI. SARAN

Para peneliti menyadari bahwa terdapat kekurangan dalam penelitian ataupun penyusunan karya tulis ini. Oleh karena itu, kritik serta saran yang membangun dari para pembaca dan pengamat sangat dibutuhkan untuk memperbaiki penelitian-penelitian selanjutnya.

Mengenai penelitian selanjutnya, dibutuhkan pengembangan dalam dataset citra serta jumlah data dalam tiap kelas yang digunakan. Selain itu, model arsitektur yang digunakan masih membutuhkan perkembangan guna memiliki kemampuan mengklasifikasi citra lebih baik. Pada penelitian selanjutnya dapat menggunakan arsitektur lain seperti GoogLeNet, Inception-V3, Squeezenet dan lain-lain.

REFERENSI

- S. R. D. Amiril, "IMPLEMENTASI ALGORITMA CONVOLUTIONAL NEURAL NETWORK PADA KLASIFIKASI PENYAKIT PADI MELALUI CITRA DAUN," UII Yogyakarta, Yogyakarta, 2020.
- [2] F. and R., "KLASIFIKASI JENIS CITRA DAUN MANGGA MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK," Jurnal Ilmiah Teknologi dan Rekayasa, vol. 25, no. 3, pp. 223-238, 2020.

- [3] C. A. Sari and E. H. Rachmawanto, "Fitur Esktraksi LBP dan Naive Bayes dalam Klasifikasi Jenis Pepaya Berdasarkan Citra Daun," *Jurnal Masyarakat Informatika*, vol. 12, no. 2, pp. 102-113, 2021.
- [4] N. Ibrahim, G. A. Lestary, F. S. Hanafi, K. Saleh, N. K. C. Pratiwi, M. S. Haq and A. I. Mastur, "Klasifikasi Tingkat Kematangan Pucuk Daun Teh menggunakan Metode Convolutional Neural Network," ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika, vol. 10, no. 1, pp. 162-176, 2022.
- [5] Y. P. Irawan and I. Susilawati, "Klasifikasi Jenis Aglaonema Berdasarkan Citra Daun Menggunakan Convolutional Neural Network (CNN)," *Journal of Information System and Artificial Intelligence*, vol. 2, no. 2, 2022.
- [6] B. Febriana, "IDENTIFIKASI PENYAKIT DAUN APEL MENGGUNAKAN RESNET 50 DILATED CONVOLUTION NEURAL NETWORK," Institut Teknologi Nasional Bandung, Bandung, 2020.
- [7] D. Hindarto and . H. Santoso, "Plat Nomor Kendaraan dengan Convolution Neural Network," JII: Jurnal Inovasi Informatika Universitas Pradita, vol. 6, no. 2, pp. 1-12, 2021.

Kontribusi

- A. Pencarian Studi Literatur.
 - Jihan Kamilah, Nurhikmah Mawaddah Solin, Salsabila Oktafani.
- B. Pembuatan Kode Program.
 - Alvin Putra Perdana, Jihan Kamilah, Nurhikmah Mawaddah Solin, Salsabila Oktafani.
- C. Penyusunan Laporan.
 - Alvin Putra Perdana, Jihan Kamilah, Nurhikmah Mawaddah Solin, Salsabila Oktafani.
- D. Pembuatan PPT.
 - Jihan Kamilah, Nurhikmah Mawaddah Solin, Salsabila Oktafani.
- E. Presentasi Materi (Video).
 - Alvin Putra Perdana, Jihan Kamilah, Nurhikmah Mawaddah Solin, Salsabila Oktafani.
- F. Upload Laporan, dan Link Video. Alvin Putra Perdana.