

## BAB II LANDASAN TEORI

### 2.1. TINJAUAN PUSTAKA

#### 2.1.1. Analisis Penelitian Terdahulu

Refrensi penelitian sebelumnya yang menggunakan metode *You Only Look Once (YOLO)* dan *Single Shot Multibox Detector (SSD)* untuk klasifikasi citra dan pengolahan objek penelitian dirangkum dalam tabel ini.

**Tabel 2. 1** Penelitian Terdahulu

| No. | Penulis                                      | Tahun | Metode                             | Judul  | Hasil  |
|-----|--|-------|------------------------------------|--|--|
| 1.  | Taufiq Nurhidayat, Patmi Kasih, Ardi Sanjaya | 2019  | KNN (K-Nearest Neighbor)           | Classification of Aglaonema Plants Berdasarkan Corak Daun                                    | Dari hasil pengujian yang dilakukan terhadap aplikasi ini, dengan menggunakan 3 jenis tanaman Aglaonema dan menggunakan data uji sebanyak 90 citra, lalu memperoleh tingkat akurasi sebesar 75,55%.                          |
| 2.  | Satrio Muhammad                              | 2021  | Convolutional Neural Network (CNN) | Klasifikasi Tanaman Aglaonema Berdasarkan Citra Daun Menggunakan Metode Convolutional Neural | Hasil penelitian menunjukkan bahwa pada model yang memiliki background mendapatkan akurasi pengujian mencapai 99% dengan nilai loss sebesar 0.084 serta F1-Score tertinggi sebesar 100% untuk jenis Red Majesty dengan Black |

|    |                              |      |                                  |   |   |
|----|------------------------------|------|----------------------------------|---|---|
|    |                              |      |                                  | Network (CNN)   | Maroon. Sementara pada model yang tidak memiliki background mendapatkan akurasi pengujian mencapai 71%.   |
| 3. | Muhammad Syarif Hidayatulloh | 2021 | You Only Look Once (YOLO)        | Sistem Pengenalan Wajah Menggunakan Metode YOLO ( <i>You Only Look Once</i> )     | Hasil pengujian yang dilakukan pada aplikasi ini, deteksi wajah mencapai akurasi sebesar 100% saat kamera berada pada posisi depan, kanan, dan kiri terhadap wajah. Pengenalan wajah dari berbagai sudut pandang juga memiliki akurasi sebesar 100% . |
| 4. | Henry Tanujaya, Lina         | 2023 | Single Shot Multi Detector (SSD) | Pengenalan Objek Menggunakan Metode Single Shot Multi Detector Pada Bahan Sembako | Dari pengujian ini, dapat disimpulkan bahwa semakin besar jumlah epoch yang digunakan akan semakin kecil tingkat akurasi yang dihasilkan. Sehingga, pada epoch 150 memperoleh akurasi sebesar 86.08%.   |
| 5. | Budi Setiawan                | 2021 | You Only Look Once (YOLO)        | Penerapan Algoritma You Only Look Once (YOLO)                                     | Hasil pengujian pada penelitian ini dalam mendeteksi tanaman Miana mendapatkan akurasi sebesar 87% dengan melakukan   |

|  |  |  |  |  |  |
|--|--|--|--|--|--|
|  |  |  |  | Untuk Deteksi Tanaman Miana Berbasis Android | proses pada data latih (train) dilakukan menggunakan 250 gambar. |
|--|--|--|--|--|--|

### 2.1.2. Perbedaan Analisis Penelitian Terdahulu Dan Saat Ini

Setelah mencermati beberapa penelitian terdahulu mengenai metode dan objek yang digunakan dalam penelitian ini, maka dapat dilihat perbedaan antara penelitian ini dengan penelitian sebelumnya yang tercantum pada tabel dibawah ini.

**Tabel 2. 2** Penelitian Saat Ini

| No. | Penulis        | Tahun | Metode  | Judul   | Hasil   |
|-----|----------------|-------|---|---|---|
| 1.  | Riska Larasati | 2024  | You Only Look Once (YOLO) dan Single Shot Multibox Detector (SSD) | Implementasi Klasifikasi Tanaman Aglaonema Menggunakan Metode YOLO (You Only Look Once) Dan SSD (Single Shot Multibox Detector) | Hasil dari pengujian yang dilakukan pada aplikasi ini, penulis menggunakan jenis tanaman Aglaonema Anjamani, Lady Valentine, Red Legacy, Suksom dan Bukan Aglaonema yang mempunyai data uji coba sebanyak 1000 gambar, menggunakan metode <i>You Only Look Once (YOLO)</i> dan <i>Single Shot Multibox Detector (SSD)</i> dan memperoleh presentase akurasi sebanyak 95% dan 90%. |

Penelitian ini bertujuan untuk mengklasifikasi tanaman Aglaonema, dan seiring berjalannya waktu, metode dan algoritma akan terus berkembang dengan kemampuan yang lebih baik. Dalam banyak penelitian sebelumnya, ditemukan bahwa metode *You Only Look Once (YOLO)* dan *Single Shot Multibox Detector (SSD)* mampu memberikan tingkat akurasi yang lebih baik dalam mengklasifikasikan tanaman Aglaonema. Metode ini juga memungkinkan terjadinya overfitting pada model yang dikembangkan.

## **2.2. TANAMAN AGLAONEMA**

Tanaman hias Aglaonema, juga dikenal sebagai Sri Rejeki di Indonesia, memiliki nama ilmiah Aglaonema. Asal-usul namanya berasal dari Bahasa Yunani, di mana 'agros' yang artinya cahaya dan 'nema' yang artinya benang. Dengan demikian, secara harfiah 'Aglaonema' mengacu pada benang yang bersinar. Keunikan tanaman Aglaonema terletak pada kombinasi warna dan tekstur yang khas dari daunnya. Secara umum, terdapat dua jenis Aglaonema yaitu spesies dan hibrida. Perbedaan antara keduanya terletak pada warna daun, Aglaonema spesies cenderung memiliki warna hijau kehijauan dengan corak hijau kehitaman, sementara Aglaonema hibrida memiliki variasi warna daun yang lebih beragam (GH Muliana, 2022).

Aglaonema merupakan tanaman monokotil yaitu memiliki akar serabut. Akar yang sehat biasanya memiliki warna putih dan berisi, sedangkan untuk akar yang tidak sehat cenderung mengalami perubahan warna menjadi cokelat, mengecil, atau bahkan busuk. Batang Aglaonema merupakan batang yang berair dan lunak, tidak berkayu, berwarna putih, hijau, merah muda hingga merah, serta memiliki nodus (berbuku). Buku atau

nodus pada batang Aglaonema merupakan tempat keluarnya daun baru atau tunas yang memungkinkan tumbuh menjadi cabang baru bila memungkinkan. Daun tanaman Aglaonema memiliki warna yang beranekaragam, seperti warna hijau tua, hijau muda, kuning, merah maroon, pink tua, pink muda, merah cerah hingga putih. Tanaman Aglaonema terkenal karena memiliki daun yang beranekaragam warnanya, daun Aglaonema merupakan bagian dari tanaman Aglaonema yang menonjol dan disukai banyak peminat tanaman hias. Bentuk daun Aglaonema umumnya berbentuk oval membulat dengan ujung runcing, atau oval memanjang langsing, ukuran daun yang bervariasi tergantung dari jenisnya, daun tanaman Aglaonema juga tahan lama dan tidak mudah rontok (GH Muliana, 2022).

### **2.2.1. Anjamani**

Anjamani adalah salah satu varietas Aglaonema yang terkenal di kalangan masyarakat ini berasal dari Thailand. Tanaman hias ini memiliki kombinasi warna hijau cerah dan merah yang mencolok. Bentuk daunnya yang sedikit bulat, memudahkannya dalam penyusunan daunnya sehingga membentuk roset (Zulfita & Hariyanti, 2020).



**Gambar 2. 1** Tanaman Aglaonema Jenis Anjamani

### 2.2.2. Lady Valentine

Aglaonema ‘Lady Valentine’ memiliki daun merah muda dengan semburat hijau tua, tepi helai daunnya rata, ujung daun yang melancip, pelepah daun berwarna hijau, dan batang berbentuk silindris yang berwarna hijau (Alifia et al., 2023).



**Gambar 2. 2** Tanaman Aglaonema Jenis Lady Valentine

### 2.2.3. Red Legacy

Aglaonema ‘Red Legacy’ mempunyai daun berwarna hijau kekuningan dengan bercak hijau gelap yang merata dipermukaannya. Daunnya memiliki bentuk yang bulat namun meruncing di ujungnya. Ciri khas yang membedakan Aglaonema ‘Red Legacy’ adalah adanya garis warna merah lembut yang membatasi tulang daunnya.



**Gambar 2. 3** Tanaman Aglaonema Jenis Red Legacy

#### 2.2.4. Suksom

Aglaonema Suksom memiliki ciri khas yaitu dengan bentuk oval lebar dan ujung daun yang meruncing. Warna daun pada Aglaonema jenis suksom memiliki perpaduan warna antara warna hijau tua dan merah terang, warna hijau tua biasanya hanya terdapat pada bagian tepi daun (GH Muliana, 2022).



**Gambar 2. 4** Tanaman Aglaonema Jenis Suksom

### 2.3. DEEP LEARNING

*Deep Learning* merupakan cabang ilmu dalam bidang *machine learning* yang menggunakan Jaringan Saraf Tiruan (JST) sebagai basisnya atau bisa dianggap sebagai evolusi dari JST. Dalam *deep learning*, komputer dapat melakukan klasifikasi langsung berdasarkan data gambar atau audio (Sarirotul, Agung, Deep Learning Jenis Tumbuhan Cnn.Pdf, 2018).

### 2.4. COMPUTER VISION

*Computer Vision* adalah bidang ilmu komputer yang bertujuan untuk memungkinkan komputer untuk melihat, mengenali, dan memproses gambar dengan

cara yang serupa dengan manusia, serta memberikan respons yang sesuai. Prinsipnya mirip dengan menanamkan naluri dan kecerdasan manusia ke dalam komputer.

Bidang ini erat kaitannya dengan kecerdasan buatan karena komputer perlu mengartikan gambar yang diterimanya dan menjalankan analisis. *Computer Vision* memiliki tujuan yang lebih dari sekedar kemampuan visual, ini juga tentang mengolah informasi yang diperoleh pengamatan secara langsung untuk menghasilkan hasil yang bermanfaat (Hidayatulloh, 2021).

## **2.5. PERBEDAAN FOLDER DAN DATABASE**

Folder adalah lokasi untuk menyimpan berbagai file di dalam komputer. Fungsi utamanya adalah menyimpan file dan mengelompokkannya berdasarkan jenis file tersebut.

Database atau basis data merupakan kumpulan informasi yang tersimpan secara sistematis di dalam komputer, sehingga dapat diakses dan dikelola menggunakan program komputer untuk memperoleh informasi. Konsepnya, basis data adalah kumpulan data yang saling terkait dalam suatu berkas atau file yang mengikuti aturan tertentu untuk membentuk data atau informasi baru. Database adalah kumpulan dari banyaknya data yang berhubungan satu sama lain yang diatur berdasarkan skema atau struktur tertentu. Di dalam komputer, basis data disimpan di perangkat keras penyimpanan, dan dimanipulasi menggunakan perangkat lunak tertentu untuk tujuan tertentu (Andaru, 2020).

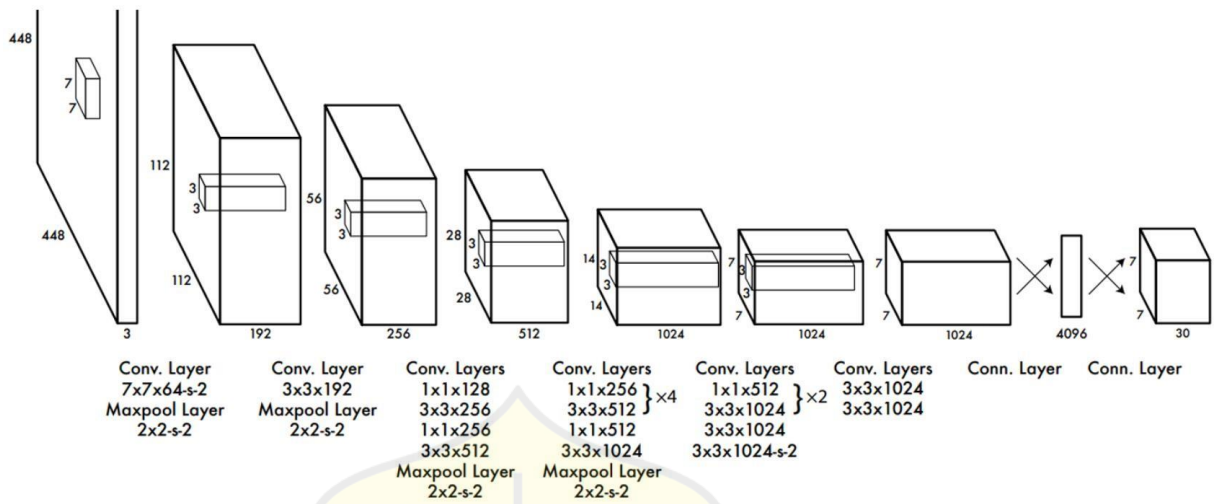


## **2.6. PENGOLAHAN CITRA DIGITAL**

Pengolahan Citra Digital adalah cabang ilmu yang mempelajari bagaimana proses pembentukan, pengolahan, dan analisis citra dengan tujuan menghasilkan informasi yang dapat dimengerti oleh manusia. Pengolahan citra melibatkan manipulasi sinyal di mana gambar merupakan input utamanya, seperti foto atau frame video, dan outputnya bisa berupa gambar yang dimodifikasi atau serangkaian karakteristik atau parameter yang terkait dengan gambar tersebut. Citra digital yang dihasilkan dari proses ini adalah representasi visual dari daun. Citra daun kemudian diproses untuk menghasilkan informasi tentang luas daun menggunakan teknik computer vision (Andrian et al., 2022).

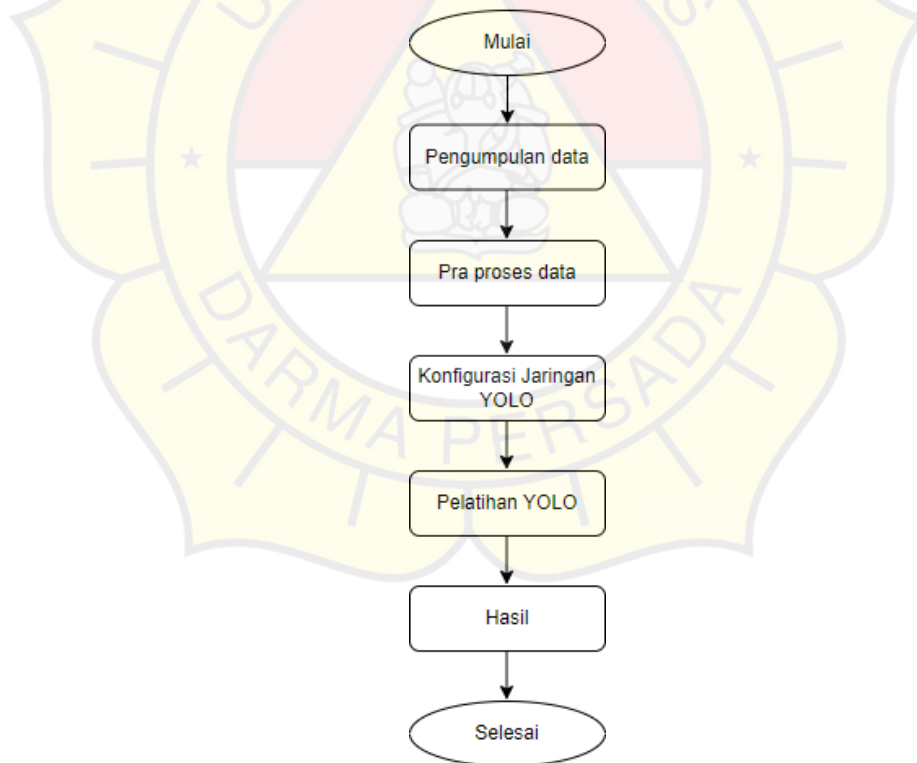
## **2.7. YOU ONLY LOOK ONCE (YOLO)**

*You Only Look Once (YOLO)* merupakan jaringan pendeteksi objek yang bertujuan dalam menentukan posisi suatu objek dalam gambar dan mengklasifikasikannya. Dengan cara menggunakan gambar sebagai data untuk diinput, YOLO menghasilkan vektor yang menunjukkan kotak pembatas serta kelas prediksi sebagai outputnya (Hidayatulloh, 2021).



**Gambar 2. 5** Arsitektur YOLO

(Sumber : Dwiyanto et al., 2022)



**Gambar 2. 6** Skema YOLO

## 1. Pengumpulan Data

Gambar tanaman yang digunakan untuk penelitian ini diperoleh dari dataset yang sebagian dikelola sendiri. Gambar tanaman yang digunakan adalah 1000 gambar dengan format jpg, serta menggunakan 5 kategori yaitu Anjamani, Lady Valentine, Red Legacy, Suksom, dan Bukan Aglaonema.

## 2. Pra Proses Data

Sebelum memulai proses pelatihan, data pra-pemrosesan dilakukan dengan dua langkah utama yaitu pelabelan dan penyesuaian ukuran citra. Pelabelan melibatkan pemberian label kepada setiap citra dalam dataset, dengan penekanan pada penggunaan kotak pembatas dan kelas untuk masing-masing objek dalam gambar. Kemudian, penyesuaian bentuk gambar dilakukan untuk meningkatkan kinerja pada model YOLO dalam pengenalan objek (Hidayatulloh, 2021).

## 3. Konfigurasi Jaringan YOLO

Untuk konfigurasi jaringan YOLO, pada batch size atau iterasi yang digunakan adalah 64, yang berarti setiap iterasi berisi 64 citra. Subdivision menggunakan sebesar 8, kemudian batch dibagi menjadi 8 bagian. Dengan demikian, masing-masing 8 citra per minibatch dikirim ke GPU untuk diproses, dan proses tersebut diulang 8 kali hingga selesai. Iterasi selanjutnya akan dimulai dengan 64 citra baru (Hidayatulloh, 2021).

## 4. Pelatihan YOLO

Proses pelatihan YOLO melibatkan jaringan saraf yang terdiri atas lapisan konvolusi dengan kernel  $3 \times 3$  dan lapisan max-pooling dengan kernel  $2 \times 2$ . Di bawah lapisan konvolusional terakhir, terdapat penggunaan kernel  $1 \times 1$  untuk mengurangi dimensi

data ke format  $13 \times 13 \times 40$ . Dimensi  $13 \times 13$  merepresentasikan ukuran grid, sedangkan angka 40 didapat dari jumlah filter yang digunakan (Hidayatulloh, 2021).

| layer        | filters | size/strd(dil) | input             | output                  |
|--------------|---------|----------------|-------------------|-------------------------|
| 0 conv       | 16      | 3 x 3/ 1       | 416 x 416 x 3 ->  | 416 x 416 x 16 0.158 BF |
| 1 max        |         | 2 x 2/ 2       | 416 x 416 x 16 -> | 208 x 208 x 16 0.003 BF |
| 2 conv       | 32      | 3 x 3/ 1       | 208 x 208 x 16 -> | 208 x 208 x 32 0.399 BF |
| 3 max        |         | 2 x 2/ 2       | 208 x 208 x 32 -> | 104 x 104 x 32 0.001 BF |
| 4 conv       | 64      | 3 x 3/ 1       | 104 x 104 x 32 -> | 104 x 104 x 64 0.399 BF |
| 5 max        |         | 2 x 2/ 2       | 104 x 104 x 64 -> | 52 x 52 x 64 0.001 BF   |
| 6 conv       | 128     | 3 x 3/ 1       | 52 x 52 x 64 ->   | 52 x 52 x 128 0.399 BF  |
| 7 max        |         | 2 x 2/ 2       | 52 x 52 x 128 ->  | 26 x 26 x 128 0.000 BF  |
| 8 conv       | 256     | 3 x 3/ 1       | 26 x 26 x 128 ->  | 26 x 26 x 256 0.399 BF  |
| 9 max        |         | 2 x 2/ 2       | 26 x 26 x 256 ->  | 13 x 13 x 256 0.000 BF  |
| 10 conv      | 512     | 3 x 3/ 1       | 13 x 13 x 256 ->  | 13 x 13 x 512 0.399 BF  |
| 11 max       |         | 2 x 2/ 1       | 13 x 13 x 512 ->  | 13 x 13 x 512 0.000 BF  |
| 12 conv      | 1024    | 3 x 3/ 1       | 13 x 13 x 512 ->  | 13 x 13 x 1024 1.595 BF |
| 13 conv      | 512     | 3 x 3/ 1       | 13 x 13 x 1024 -> | 13 x 13 x 512 1.595 BF  |
| 14 conv      | 48      | 1 x 1/ 1       | 13 x 13 x 512 ->  | 13 x 13 x 48 0.007 BF   |
| 15 detection |         |                |                   |                         |

**Gambar 2. 7** Model YOLO

(Sumber : Hidayatulloh, 2021)

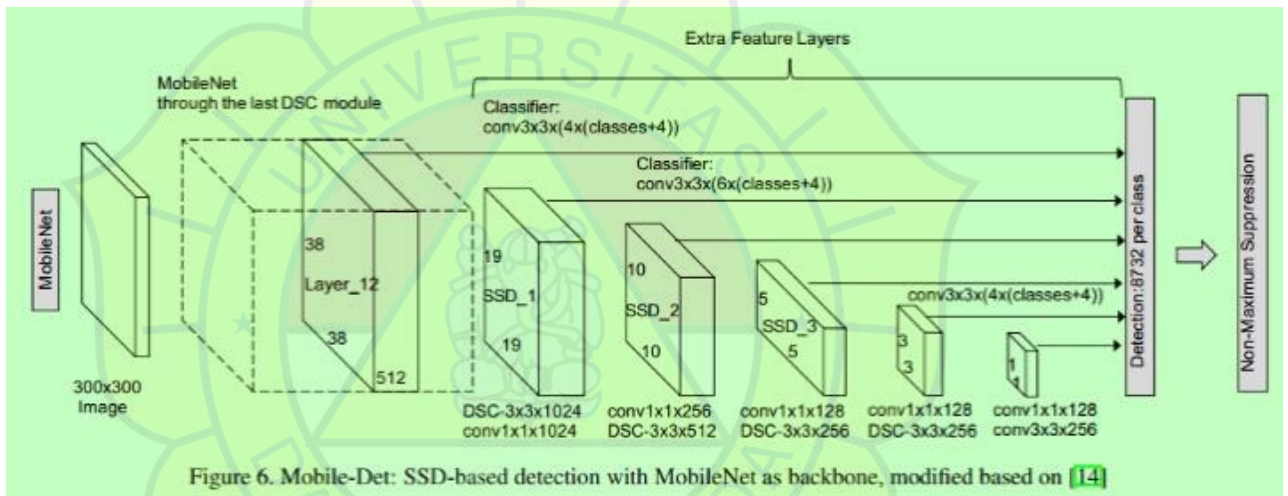
## 5. Hasil

Hasil dari pengujian yang dilakukan dengan dua macam cara yaitu melalui gambar dan melalui realtime dengan kamera handphone. Tujuan dari pengujian ini adalah untuk mengevaluasi tingkat akurasi deteksi objek menggunakan model baru yang telah dilatih sebelumnya.

### 2.8. SINGLE SHOT MULTIBOX DETECTION (SSD)

*Single Shot Multibox Detector (SSD)* adalah sebuah algoritma *deep learning* yang menghasilkan serangkaian kotak standar pada bermacam rasio dan skala aspek untuk setiap lokasi pada peta fitur, yang menggambarkan objek

yang mungkin hadir dalam gambar. Metode ini menggunakan fitur kotak pembatas untuk memperkirakan posisi objek yang terdeteksi. Keunggulan SSD terletak pada kecepatan pemrosesan dan tingkat deteksi yang tinggi, membuatnya ideal untuk aplikasi deteksi objek secara real-time. SSD juga mengintegrasikan lapisan fitur konvolusional tambahan di ujung jaringan dasar untuk memprediksi berbagai rasio aspek objek yang memungkinkan (Tanujaya et al., n.d.).



**Gambar 2. 8** Arsitektur SSD  
(Sumber : academy-ai-Medium)

## 2.9. PERBEDAAN YOLO DAN SSD

Metode You Only Look Once (YOLO) dan Single Shot Multibox Detector (SSD) adalah dua pendekatan utama dalam deteksi objek real-time pada citra. Meskipun keduanya berbagi tujuan yang sama, yaitu mendeteksi dan mengklasifikasi objek dalam gambar, mereka memiliki perbedaan dalam pendekatan dan arsitektur mereka :

## 1. You Only Look Once (YOLO) :

- a. YOLO adalah pendekatan dalam mendeteksi objek end-to-end yang memperlakukan masalah deteksi objek sebagai masalah regresi satu kali.
- b. YOLO membagi gambar ke dalam grid dan untuk setiap sel grid, algoritma langsung memprediksi kotak pembatas (bounding box) dan probabilitas kelas untuk semua objek dalam sel tersebut.
- c. Ini berarti bahwa YOLO menghasilkan prediksi objek secara simultan dan langsung dari seluruh gambar dalam satu langkah inferensi.
- d. YOLO memiliki keunggulan dalam kecepatan karena hanya memerlukan satu langkah inferensi untuk menghasilkan prediksi.

## 2. Single Shot Multibox Detector (SSD) :

- a. SSD juga merupakan pendekatan dalam mendeteksi objek real-time yang menggunakan jaringan saraf tiruan untuk memperkirakan kotak pembatas dan probabilitas kelas secara langsung.
- b. Namun, SSD menggunakan skema multi-scale feature maps untuk mendeteksi objek di berbagai skala. Ini berarti SSD memprediksi kotak pembatas dan probabilitas kelas pada beberapa tingkatan resolusi gambar.
- c. Dengan menggunakan feature maps multi-skala, SSD dapat mendeteksi objek yang berbeda dalam berbagai ukuran dan proporsi dengan lebih baik daripada YOLO.
- d. Meskipun demikian, SSD mungkin lebih lambat daripada YOLO karena memerlukan beberapa langkah inferensi pada feature maps multi-skala.

Kesimpulannya, YOLO dan SSD adalah dua pendekatan yang populer dalam deteksi objek real-time. YOLO memprioritaskan kecepatan dengan menghasilkan prediksi objek dalam satu langkah, sedangkan SSD memperhatikan kemampuan mendeteksi objek dalam berbagai skala dengan menggunakan feature maps multi-skala. Pilihan antara YOLO dan SSD tergantung pada kebutuhan aplikasi, di mana kecepatan dan akurasi dalam deteksi objek menjadi pertimbangan utama.

