

**BAB II**  
**TINJUAN PUSTAKA**

**2.1. Penellitian Terdahulu**

<b>Topik</b>	<b>Metode</b>	<b>Hasil</b>	<b>Referensi</b>
<i>Yoga pose classification: a CNN and MediaPipe inspired deep learning approach</i>	Metode estimasi pose Holistic dari MediaPipe	Penelitian ini membahas tentang klasifikasi pose yoga menggunakan pendekatan <i>deep learning</i> yang terinspirasi oleh CNN dan MediaPipe	[11]
<i>Human posture recognition based on multiple features and rule learning</i>	Algoritma baru berdasarkan fitur dan Game aturan	Penelitian ini membahas algoritma yang dapat mengenali berbagai jenis postur manusia secara efektif	[12]
<i>Vision-based human activity recognition: a survey</i>	Pengenalan aktivitas manusia berbasis visi	Penelitian ini membahas sistem HAR yang mencoba untuk menganalisis dan mengenali aktivitas manusia menggunakan informasi yang diperoleh dari berbagai jenis sensor.	[13]
<i>Survey on Emotional Body Gesture Recognition</i>	Pengenalan gestur tubuh emosional	Penelitian ini membahas tentang penelitian otomatis tentang pengenalan emosi telah menjadi topik penelitian yang sedang tren dalam dekade terakhir.	[14]

<b><i>Deep Learning- Based Human Pose Estimation: A Survey</i></b>	Estimasi pose manusia berbasis <i>Deep Learning</i>	Penelitian ini memiliki tujuan untuk menemukan bagian tubuh manusia dan membangun representasi tubuh manusia (misalnya, kerangka tubuh) dari data input seperti gambar dan <i>video</i>	<b>[15]</b>
--	---	---	-------------



<b><i>Human Body Pose Estimation and Applications</i></b>	Metode estimasi pose Holistic dari MediaPipe	Penelitian ini mempelajari tentang pendekatan <i>real-time</i> untuk deteksi dan pengenalan bahasa isyarat dalam <i>video</i>	[16]
<b><i>Image Recognition Using Machine Learning</i></b>	Penggunaan <i>machine learning</i> untuk pengenalan gambar	Pengenalan gambar adalah aspek penting dari pemrosesan gambar untuk <i>machine learning</i> tanpa melibatkan dukungan manusia pada setiap langkah. Dalam penelitian ini, mereka mempelajari bagaimana klasifikasi gambar dilakukan menggunakan <i>backend</i> imajiner	[17]
<b><i>Deep Learning-Based Human Pose Estimation: A Survey</i></b>	Estimasi pose manusia berbasis <i>Deep Learning</i>	Tujuannya adalah untuk menemukan bagian tubuh manusia dan membangun representasi tubuh manusia (misalnya, kerangka tubuh) dari data input seperti gambar dan <i>video</i>	[18]
<b><i>Human body gesture recognition method based on deep learning</i></b>	Metode berbasis <i>deep learning</i>	Teknologi pengenalan pose manusia selalu menjadi konten penelitian penting di bidang visi komputer dan pengenalan pola	
<b><i>Comparative Analysis of OpenPose, PoseNet, and MoveNet Models for Pose Estimation in Mobile Devices</i></b>	Analisis komparatif antara <i>OpenPose</i> , <i>PoseNet</i> , dan <i>MoveNet</i> untuk estimasi pose pada perangkat	Penelitian ini bertujuan untuk membandingkan performa dari model <i>OpenPose</i> , <i>PoseNet</i> , dan <i>MoveNet</i> dalam melakukan estimasi pose pada perangkat <i>mobile</i> .	[4]

	<i>mobile.</i>		
<b><i>Empirical Mode Decomposition and Hilbert Spectrum for Abnormality Detection in Normal and Abnormal Walking Transitions</i></b>	Deteksi abnormalitas dalam transisi berjalan normal dan abnormal menggunakan <i>Empirical Mode Decomposition</i> dan <i>Hilbert Spectrum</i> .	Penelitian ini bertujuan untuk mendeteksi abnormalitas dalam transisi berjalan menggunakan data yang diperoleh dari PoseNet.	[5]
<b><i>Comparative Analysis of Skeleton-Based Human Pose Estimation</i></b>	Analisis komparatif dari estimasi pose manusia berbasis kerangka.	Penelitian ini bertujuan untuk membandingkan beberapa metode estimasi pose manusia berbasis kerangka, termasuk PoseNet.	[6]
<b>Implementasi Algoritma Blum Blum Shub Pada Kuis Sandi Semaphore Berbasis Augmented Reality</b>	Implementasi Algoritma Blum Blum Shub dalam Game Semaphore berbasis <i>Augmented Reality</i> .	Penelitian ini bertujuan untuk membangun aplikasi Game gerakan semaphore dan kuis semaphore dengan menggunakan teknologi <i>augmented reality</i> berbasis android guna memberikan Game yang menarik dan interaktif.	[7]
<b>DESAIN GAME EDUKASI GERAKAN SEMAPHORE BERTEKNOLO</b>	Desain game edukasi gerakan semaphore berteknologi	Penelitian ini bertujuan untuk melatih pemain dalam mengingat dan memahami sandi semaphore, pembuatan desain game edukasi dilakukan dengan menggunakan	[8]

<b>GI KINECT MENGUNAKA N FRAMEWORK MECHANICS DYNAMICS AESTHETICS</b>	Kinect menggunakan <i>framework Mechanics Dynamics Aesthetics.</i>	pendekatan <i>Mechanics Dynamic Aesthetics.</i>	
<b>Studi Klasifikasi Gerakan Semaphore menggunakan Fuzzy Mamdani dari Data IMU Sensor</b>	Studi klasifikasi gerakan Semaphore menggunakan Fuzzy Mamdani dari data IMU Sensor.	Penelitian ini bertujuan untuk menilai gerakan latihan dengan menggunakan <i>fuzzy logic</i> dengan mamdani <i>inference system</i> .	[9]
<b>Studi Klasifikasi Gerakan Semaphore menggunakan Fuzzy Mamdani dari Data IMU Sensor</b>	Studi klasifikasi gerakan Semaphore menggunakan Fuzzy Mamdani dari data IMU Sensor.	Penelitian ini bertujuan untuk menilai gerakan latihan dengan menggunakan <i>fuzzy logic</i> dengan mamdani <i>inference system</i> .	[10]

Tabel 2. 1. 1 Penelitian Terkait.

## 2.2. PoseNet

Dalam pengembangan sistem ini, algoritma yang digunakan adalah PoseNet. PoseNet adalah model *machine learning* yang dapat memperkirakan pose manusia dalam gambar dan *video* secara real-time. Dengan PoseNet, kita dapat melacak 17 titik kunci pada tubuh manusia, yang dapat digunakan untuk memonitor dan mengoreksi gerakan tubuh pengguna saat melakukan gestur semaphore.

Posenet adalah sebuah model *machine learning* yang dirancang untuk mendeteksi dan melacak pose manusia dalam gambar atau video. Model ini dikembangkan oleh Google dan berfokus pada estimasi pose tubuh manusia. Pose dalam konteks ini merujuk pada posisi dan orientasi berbagai bagian tubuh manusia, seperti kepala, bahu, lengan, dan kaki.

Berikut adalah beberapa poin yang menjelaskan Posenet dengan lebih detail:

### 2.2.1. Arsitektur Model:

PoseNet umumnya menggunakan arsitektur *deep neural network Convolutional Neural Network* (CNN) untuk melakukan estimasi pose. CNN adalah jenis arsitektur yang baik untuk tugas pengolahan gambar karena kemampuannya untuk mengekstrak fitur-fitur hierarkis dari data gambar.

### 2.2.2. Keamanan dan Privasi:

Model PoseNet dikembangkan dengan memperhatikan keamanan dan privasi. Dalam banyak kasus, pengolahan gambar dilakukan secara lokal pada perangkat tanpa mengirimkan data gambar ke server eksternal, sehingga menjaga privasi pengguna.

### 2.2.3. Deteksi Poin Kunci:

Salah satu fitur utama PoseNet adalah kemampuannya untuk mendeteksi poin kunci pada tubuh manusia. Poin kunci ini dapat mencakup bagian-bagian tubuh seperti kepala, bahu, siku, pergelangan tangan, pinggul, lutut, dan pergelangan kaki.

### 2.2.4. Estimasi Kepercayaan:

Model PoseNet juga memberikan informasi tentang sejauh mana model yakin terhadap estimasinya. Setiap poin kunci disertai dengan skor kepercayaan yang mengindikasikan seberapa baik model yakin pada posisi tersebut.

#### **2.2.5. Aplikasi:**

PoseNet banyak digunakan dalam berbagai aplikasi, termasuk *augmented reality*, *fitness tracking*, pengenalan gerakan, dan analisis postur ergonomis. Dalam *augmented reality*, misalnya, PoseNet dapat digunakan untuk menempatkan objek virtual secara realistis dalam lingkungan yang direkam.

#### **2.2.6. Transfer Learning:**

Model PoseNet sering kali menggunakan konsep *transfer learning*, di mana model yang telah dilatih pada dataset besar dapat diadaptasi untuk tugas khusus pose manusia dengan melatihnya pada dataset yang lebih kecil dan spesifik.

#### **2.2.7. Open Source:**

Beberapa implementasi PoseNet bersifat *open source*, yang berarti kode sumbernya dapat diakses dan dimodifikasi oleh pengembang. Hal ini memfasilitasi penelitian dan pengembangan lebih lanjut di komunitas.

### **2.3. YOLO**

YOLO (*You Only Look Once*) mencakup pemahaman tentang cara kerja algoritma ini dalam deteksi objek pada citra. YOLO adalah pendekatan yang inovatif untuk deteksi objek secara *real-time*, dan berikut adalah beberapa konsep dasar yang membentuk landasan teori YOLO:

#### **2.3.1. Single Shot Detection (SSD):**

YOLO menggunakan pendekatan SSD, yang berarti deteksi objek dilakukan dalam satu tahap, tidak seperti pendekatan yang memerlukan beberapa tahap seperti



R-CNN (*Region-based Convolutional Neural Network*). SSD memungkinkan YOLO untuk mencapai tingkat kecepatan tinggi dalam deteksi objek.

### **2.3.2. Grid System:**

YOLO membagi citra menjadi sel-sel atau grid. Setiap sel bertanggung jawab untuk mendeteksi objek dalam area tersebut. Sebagai contoh, jika citra dibagi menjadi grid 3x3, setiap sel bertanggung jawab untuk deteksi objek di bagian tertentu dari citra.

### **2.3.3. Bounding Box Prediction:**

Setiap sel dalam grid memprediksi kotak pembatas (*bounding box*) yang mengelilingi objek. Prediksi *bounding box* mencakup koordinat relatif terhadap ukuran sel tersebut.

### **2.3.4. Class Prediction:**

YOLO juga memprediksi probabilitas kelas objek yang terdapat dalam *bounding box* tersebut. Ini berarti untuk setiap *bounding box*, model memberikan skor probabilitas untuk setiap kelas yang mungkin.

### **2.3.5. Non-Maximum Suppression (NMS):**

Setelah prediksi dilakukan, teknik NMS digunakan untuk mengurangi redundansi deteksi. NMS menghapus kotak pembatas dengan probabilitas rendah atau yang memiliki IoU (*Intersection over Union*) tinggi dengan kotak pembatas lain yang memiliki probabilitas lebih tinggi.

### **2.3.6. Loss Function:**

YOLO menggunakan fungsi kerugian (*loss function*) yang menggabungkan kerugian lokal (*localization loss*) dan kerugian kelas (*classification loss*). Ini membantu dalam melatih model untuk memberikan prediksi *bounding box* yang akurat dan probabilitas kelas yang tepat.

### **2.3.7. Darknet Framework:**



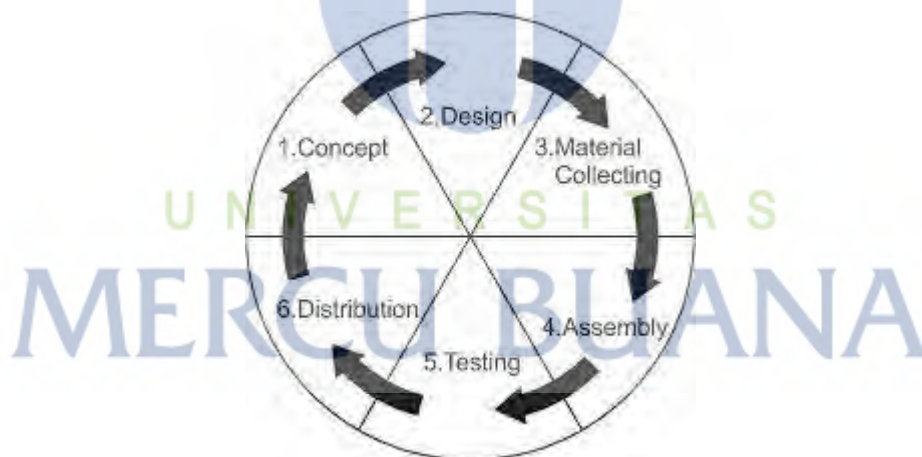
*Darknet* adalah *framework* yang dikembangkan oleh pembuat YOLO, Joseph Redmon. *Darknet* dikhususkan untuk tugas-tugas *vision* AI dan menyediakan implementasi YOLO.

### 2.3.8. Versi YOLO:

Terdapat beberapa versi YOLO, seperti YOLOv1, YOLOv2 (atau disebut YOLO9000), YOLOv3, dan seterusnya. Setiap versi memiliki perbaikan dan peningkatan kinerja dari versi sebelumnya.

Memahami konsep-konsep di atas membantu dalam merancang dan melatih model YOLO untuk deteksi objek dengan akurasi dan kecepatan tinggi. Implementasi YOLO pada aplikasi deteksi objek seperti gambar atau video memanfaatkan arsitektur jaringan saraf konvolusional (CNN) untuk memproses dan memahami konten visual secara mendalam.

## 2.4. Metode MDLC



Gambar 2.3 *Multimedia Development Life Cycle*

MDLC (*Multimedia Development Life Cycle*) adalah metode pengembangan sistem yang cocok untuk pengembangan sistem berbasis multimedia. MDLC terdiri dari enam tahap, yaitu:

**2.4.1. *Concept* (Pengonsepan):** Tahap ini adalah tahap untuk menentukan tujuan dan kepada siapa multimedia ditujukan (*audiens identification*). Selain itu

menentukan jenis aplikasi (presentasi, interaktif, dan lain-lain) dan tujuan aplikasi (hiburan, pembelajaran, dan lain-lain). Dasar aturan untuk perancangan juga ditentukan pada tahap ini misalnya ukuran, target.

**2.4.2. Design (Perancangan):** Tahap pembuatan spesifikasi meliputi arsitektur proyek, gaya, tampilan dan kebutuhan material atau bahan untuk program. Spesifikasi dibuat serinci mungkin sehingga pada tahap berikutnya yaitu material collecting dan assembly, pengambilan keputusan baru tidak diperlukan lagi.

**2.4.3. Material Collecting (Pengumpulan Bahan):** Tahap pengumpulan bahan yang sesuai dengan kebutuhan yang dikerjakan. Bahan-bahan tersebut antara lain seperti clip-art, graphic, animasi, video, audio.

**2.4.4. Assembly (Pembuatan):** Tahap pembuatan semua obyek atau bahan multimedia dibuat.

**2.4.5. Testing (Pengujian):** Tahap ini melibatkan uji *black box* dan uji kualitas. Dalam uji *black box*, peneliti melakukan validasi fitur yang telah dibangun dimana hasilnya setiap fitur tersebut dapat berjalan saat aplikasi digunakan.

Metode MDLC ini banyak digunakan pada penelitian terkait aplikasi multimedia atau mobile aplikasi. Namun, metode ini tidak tepat untuk pengembangan jangka pendek karena metode ini memiliki idealis pada setiap tahap, dimana pengerjaan setiap tahap dapat dilakukan bilamana tahap sebelumnya sudah selesai dikerjakan, sehingga membutuhkan waktu yang lebih panjang dalam tahap pengembangan aplikasi menggunakan metode MDLC.

## 2.5. Library p5.js

p5.js adalah *library* JavaScript untuk coding kreatif, dengan fokus pada membuat coding dapat diakses dan inklusif untuk seniman, desainer, pendidik, pemula, dan siapa saja. p5.js gratis dan *open-source* karena kami percaya bahwa perangkat lunak, dan alat untuk mempelajarinya, harus dapat diakses oleh semua orang[19].

p5.js memiliki seperangkat fungsi gambar lengkap. Namun, tidak terbatas pada kanvas gambar. Dapat menganggap seluruh halaman *browser* sebagai sketsa , termasuk objek HTML5 untuk teks, input, *video*, *webcam*, dan suara[19].

Sebuah *library* p5.js dapat berupa kode *JavaScript* apa pun yang memperluas atau menambahkan ke fungsionalitas inti p5.js. Ada dua kategori *library*. *Main Library* adalah bagian dari distribusi p5.js, sementara perpustakaan yang dikontribusikan dikembangkan, dimiliki, dan dipelihara oleh anggota komunitas p5.js[20].

## 2.6. *Library ml5.js*

ml5.js adalah perpustakaan *JavaScript* yang dibangun di atas TensorFlow.js yang memberikan akses ke berbagai algoritma *machine learning* dan *deep learning* di dalam *browser*. Perpustakaan ini bertujuan untuk membuat *machine learning* dapat diakses oleh komunitas *JavaScript* yang besar[21].

ml5.js bertujuan untuk membuat *machine learning* dapat diakses oleh berbagai audiens, termasuk seniman, coder kreatif, dan pelajar. Perpustakaan ini menyediakan akses ke algoritma dan model *machine learning* di *browser*, dibangun di atas TensorFlow.js. ml5.js adalah antarmuka tingkat tinggi yang ramah dan *open source* untuk TensorFlow.js, sebuah perpustakaan untuk menangani operasi matematika yang dipercepat oleh GPU dan manajemen memori untuk algoritma *machine learning*[22].

ml5.js menyediakan akses langsung di *browser* ke model pra-latih untuk mendeteksi pose manusia, menghasilkan teks, menata gambar dengan gambar lain, membuat musik, deteksi nada, hubungan kata bahasa Inggris yang umum, dan banyak lagi. ml5.js bertujuan mendukung pemahaman publik yang lebih luas tentang *machine learning* dan mendorong keterlibatan yang lebih dalam dengan komputasi etis, pengumpulan data yang bertanggung jawab, dan aksesibilitas serta keragaman orang dan perspektif dalam teknologi dan seni[22].