

## **BAB II**

### **TINJAUAN LITERATUR**

#### **2.1 Penelitian Terkait**

Penelitian ini memaparkan kajian teoretis dan empiris yang menjadi landasan dalam penelitian implementasi *eye tracking* untuk deteksi kecurangan pada simulasi ujian berbasis *computer vision*. Kajian pustaka disusun secara sistematis dengan menelaah konsep-konsep fundamental yang saling berkaitan, meliputi teori dan prinsip kerja *eye tracking*, arsitektur dan algoritma *computer vision*, karakteristik perilaku visual yang mengindikasikan kecurangan, serta sistem pengawasan ujian (*proctoring system*) yang telah dikembangkan sebelumnya. Selain itu, bab ini juga mengulas penelitian-penelitian terdahulu yang relevan sebagai acuan perbandingan sekaligus justifikasi atas pendekatan yang digunakan dalam penelitian ini.

Penelitian oleh Dilini tahun 2021, penelitian ini mengusulkan sistem pendeteksi kecurangan ujian daring berbasis browser dengan memanfaatkan pelacakan gerak mata melalui kamera web bawaan perangkat. Sebuah plugin browser dikembangkan untuk merekam data tatapan mata peserta ujian, yang kemudian dianalisis menggunakan metode *One Class Support Vector Machine* (OCSVM) untuk mendeteksi perilaku anomali, seperti seringnya menatap ke luar layar yang mengindikasikan potensi kecurangan. Pendekatan ini dirancang sebagai solusi otomatis dan murah tanpa perlu perangkat keras tambahan, serta telah diuji coba pada mahasiswa dengan tingkat akurasi sekitar 89,53%. Meskipun demikian,

sistem ini masih memiliki keterbatasan, seperti ketergantungan pada kualitas kamera dan kondisi pencahayaan, serta belum mampu mendeteksi semua jenis kecurangan. Secara keseluruhan, metode ini menawarkan alternatif proctoring yang praktis dan dapat dikembangkan lebih lanjut untuk berbagai platform ujian berbasis web (Dilini *et al.*, 2021).

Penelitian oleh Bimantoro dan kawan-kawan tahun 2024, penelitian ini mengembangkan sistem manajemen pembelajaran (LMS) untuk kursus online yang dilengkapi dengan kemampuan deteksi kecurangan ujian secara otomatis menggunakan model MediaPipe Face Mesh dari TensorFlow.js. Sistem bekerja dengan mengakses kamera web peserta, kemudian menganalisis 468 titik penanda wajah secara real-time untuk mendeteksi perilaku mencurigakan seperti menoleh ke kanan, kiri, atas, atau bawah. Selain itu, sistem juga memantau penyalahgunaan aplikasi seperti membuka tab baru atau berpindah antar aplikasi, dan akan menghentikan ujian secara otomatis jika terdeteksi kecurangan. LMS ini dirancang untuk tiga jenis pengguna, yaitu admin, instruktur, dan peserta, serta telah diuji menggunakan metode *blackbox testing* yang menunjukkan hasil sesuai harapan. Penelitian ini bertujuan untuk meningkatkan integritas dan kualitas ujian daring melalui pendekatan deteksi wajah yang akurat, cepat, dan tidak memerlukan instalasi perangkat tambahan (Bimantoro, Hidayattullah and Af'idah, 2024).

Penelitian yang dilakukan oleh Yaqub dan kawan-kawan tahun 2023. Penelitian ini mengusulkan sistem pengawasan ujian daring yang menjaga privasi peserta dengan cara menyembunyikan bagian wajah selain area mata, sehingga pihak ketiga tidak dapat melihat identitas peserta namun masih bisa memantau

gerak mata untuk mendeteksi kecurangan. Sistem menggunakan pelacakan mata (gaze tracking) untuk mengidentifikasi perilaku mencurigakan, seperti ketika peserta membaca dari layar komputer lain (misalnya ChatGPT). Proses verifikasi identitas dilakukan melalui pengenalan wajah dengan Siamese Network dan OCR untuk membaca kartu identitas, sedangkan deteksi kecurangan dilakukan dengan menganalisis arah pandangan menggunakan model iTracker yang telah dikalibrasi secara personal. Hasil uji coba pada tiga video simulasi menunjukkan bahwa metode gabungan antara deteksi anomali berbasis hash gambar dan pelacakan mata menghasilkan akurasi hingga 90,3%, serta mampu berjalan pada komputer biasa dengan kecepatan 31–35 frame per detik. Penelitian ini merupakan langkah awal menuju sistem pengawasan daring yang tetap menghormati privasi peserta, meskipun masih memerlukan pengembangan lebih lanjut seperti penggunaan dataset yang lebih besar dan penambahan deteksi anomali suara (Yaqub, Mohanty and Suleiman, 2023).

Penelitian lain oleh Zuo, Chai and Goh tahun 2024, penelitian ini mengembangkan sistem deteksi kecurangan ujian secara otomatis dengan memanfaatkan algoritma YOLOv8 yang ditingkatkan menggunakan mekanisme perhatian (attention mechanism). Metode ini mampu mendeteksi kecurangan secara real-time melalui rekaman video ruang ujian, meskipun hasil akhir tetap memerlukan penilaian manual karena tidak semua perilaku tidak biasa menandakan kecurangan. Penelitian ini membuka peluang penerapan di bidang pendidikan, pelatihan korporat, hingga sistem pengawasan keamanan (Zuo, Chai and Goh, 2024).

Penelitian oleh Lamba dan Sharma tahun 2024, penelitian ini mengembangkan sistem pendeteksi kecurangan ujian daring multimodal berbasis deep learning yang menggabungkan berbagai teknologi untuk memantau perilaku peserta ujian secara real-time. Sistem ini menggunakan YOLO untuk deteksi wajah dan jumlah orang, MediaPipe untuk mengekstraksi 468 titik landmark wajah, L2CS-Net untuk pelacakan arah mata, SixDRepNet untuk estimasi pose kepala, serta Silero VAD untuk analisis audio. Semua data tersebut kemudian diproses menggunakan arsitektur hibrida CNN-BiLSTM yang mampu menangkap pola temporal dari perilaku mencurigakan seperti kehadiran orang lain, perhatian yang teralihkan, atau komunikasi suara di luar layar. Berdasarkan pengujian pada 18 subjek dengan durasi video 8–20 menit, model mencapai akurasi 87,5%, F-score 0,8762, dan AUC 0,8795, dengan kemampuan spesifisitas tinggi (93,75%) yang menandakan sistem jarang salah menandai perilaku normal sebagai kecurangan. Sistem juga mendeteksi perpindahan tab browser melalui JavaScript dengan asumsi akurasi 100%. Meskipun demikian, penelitian ini mengakui bahwa data dikumpulkan dari partisipan yang berakting sehingga mungkin tidak sepenuhnya natural. Ke depan, sistem ini dapat ditingkatkan dengan penambahan fitur privasi serta penanganan kasus-kasus pinggiran yang masih menghasilkan kesalahan (S. Lamba and N. Sharma, 2024).

## **2.2 Pengolahan Citra Digital**

Pengolahan citra digital (*digital image processing*) merupakan serangkaian teknik komputasi yang digunakan untuk memanipulasi dan menganalisis citra dalam format digital guna menghasilkan informasi yang lebih bermakna. Secara

fundamental, citra digital direpresentasikan sebagai matriks dua dimensi yang terdiri dari elemen-elemen piksel (*picture element*), di mana setiap piksel memiliki nilai intensitas yang merepresentasikan kecerahan atau warna pada posisi spasial tertentu. Pada citra *grayscale*, nilai intensitas piksel berkisar antara 0 hingga 255, di mana nilai 0 merepresentasikan warna hitam penuh dan nilai 255 merepresentasikan warna putih penuh. Sementara itu, pada citra berwarna ruang warna BGR (*Blue, Green, Red*), setiap piksel diwakili oleh tiga nilai kanal warna yang masing-masing juga berkisar antara 0 hingga 255 (Hardini *et al.*, 2023). Rangkaian operasi dalam pengolahan citra biasanya mengalir secara berurutan, dimulai dari akuisisi, pre-processing, segmentasi, ekstraksi ciri, hingga interpretasi atau klasifikasi. Tujuan akhir dari keseluruhan proses ini adalah agar sistem komputer mampu "memahami" isi citra secara otomatis, atau setidaknya menyajikan informasi visual dalam bentuk yang lebih mudah dianalisis oleh manusia (Liaw *et al.*, 2020).

### **2.3 Kecurangan Akademik pada Ujian Berbasis Komputer**

Kecurangan akademik merupakan perilaku yang melanggar integritas pendidikan dan telah menjadi permasalahan serius di berbagai jenjang pendidikan. Menurut (Zayyinul Mushthofa, 2021), perilaku menyontek tidak hanya merupakan pelanggaran etika, tetapi juga telah menjadi kebiasaan yang mengakar sejak tingkat sekolah dasar hingga perguruan tinggi, tanpa memandang kemampuan akademik maupun jenis kelamin peserta didik. Kondisi ini menunjukkan bahwa kecurangan bukan semata persoalan individual, melainkan problem sistemik yang memerlukan penanganan secara teknologi.

Meluasnya penyelenggaraan Computer Based Test (CBT) dan ujian daring pasca pandemi Covid-19 semakin memperumit tantangan pengawasan. Dalam lingkungan ujian berbasis komputer, minimnya kehadiran fisik pengawas memberikan peluang yang lebih besar bagi peserta untuk melakukan kecurangan (Mawih, 2025). Berbagai modus kecurangan dalam ujian berbasis komputer meliputi: mengalihkan pandangan ke sumber jawaban lain, berkomunikasi dengan orang di sekitar, serta memanfaatkan perangkat digital secara sembunyi-sembunyi. Di antara semua indikator perilaku tersebut, pola pergerakan bola mata merupakan sinyal yang paling sulit disembunyikan. Ketika peserta mengalihkan pandangannya ke sumber jawaban yang bukan layar ujian, arah tatapan matanya akan berubah secara signifikan dari pola normal seseorang yang sedang berkonsentrasi mengerjakan soal.

## **2.4 Eye Tracking**

Eye tracking atau pelacakan mata adalah teknologi yang memungkinkan pengukuran dan pencatatan posisi serta gerakan mata secara presisi. Teknologi ini bekerja dengan mendeteksi posisi pupil dan pantulan kornea untuk menentukan arah pandang seseorang terhadap suatu objek atau layar (Yogasara *et al.*, 2023). Dalam konteks pengawasan ujian, eye tracking menjadi solusi yang menjanjikan karena mampu mendeteksi perilaku mencurigakan seperti mengalihkan pandangan ke luar layar, yang tidak dapat terdeteksi oleh sistem pengawasan berbasis orientasi kepala semata.

### 2.4.1 Komponen Utama Eye Tracking

Sistem eye tracking pada umumnya terdiri dari beberapa komponen fungsional yang bekerja secara terintegrasi. Komponen pertama adalah detektor wajah yang bertugas melokalisasi posisi wajah dalam setiap frame video. Komponen kedua adalah ekstraktor landmark yang mengidentifikasi titik-titik anatomis penting pada wajah, khususnya area mata. Komponen ketiga adalah estimator gaze yang menghitung arah pandang berdasarkan posisi relatif iris terhadap batas mata (Lv *et al.*, 2026).

### 2.4.2 Jenis Gerakan Mata

Gerakan mata manusia dapat diklasifikasikan ke dalam beberapa jenis yang memiliki karakteristik dan implikasi berbeda dalam analisis perilaku visual. Pemahaman terhadap jenis-jenis gerakan mata ini menjadi fondasi dalam perancangan sistem deteksi kecurangan berbasis eye tracking (Birawo and Kasprowski, 2022).

- 1) Fiksasi (Fixation): Kondisi di mana pandangan mata tertahan pada suatu titik dalam durasi tertentu, umumnya antara 100 hingga 500 milidetik. Fiksasi mengindikasikan bahwa seseorang sedang memproses informasi visual dari titik tersebut.
- 2) Sakade (Saccade): Gerakan cepat dan balistik bola mata dari satu titik fiksasi ke titik fiksasi lainnya. Sakade berlangsung sangat singkat (20-200 milidetik) namun dapat menjadi indikator perpindahan perhatian yang mencurigakan apabila terjadi terlalu sering ke arah lateral.

- 3) Gerakan Halus (Smooth Pursuit): Gerakan mata yang mengikuti objek bergerak secara halus, berbeda dengan sakade yang bersifat tiba-tiba.
- 4) Nistagmus: Gerakan mata involunter yang berulang secara ritmis, biasanya tidak relevan dalam konteks ujian namun dapat memengaruhi akurasi estimasi gaze.

## 2.5 Computer Vision

Computer vision merupakan cabang kecerdasan buatan yang memungkinkan komputer untuk menginterpretasikan dan memahami konten visual dari dunia nyata, seperti gambar dan video. Dalam konteks penelitian ini, computer vision digunakan untuk memproses setiap frame video peserta ujian guna mengekstrak informasi perilaku visual secara otomatis dan sistematis (Irfansyah *et al.*, 2025).

### 2.5.1 Pra-pemrosesan Citra (Image Preprocessing)

Pra-pemrosesan citra merupakan tahap awal yang krusial dalam pipeline computer vision. Tujuan utamanya adalah meningkatkan kualitas citra dan menyeragamkan kondisi masukan agar algoritma deteksi berikutnya dapat bekerja secara optimal. Tahapan pra-pemrosesan yang diterapkan dalam penelitian ini meliputi:

- 1) Resize: Penyeragaman dimensi frame ke resolusi target (640×480 piksel) untuk memastikan konsistensi komputasi dan mengoptimalkan kecepatan pemrosesan.



- 2) Konversi Ruang Warna: Transformasi dari format BGR (Blue-Green-Red) yang digunakan OpenCV ke format RGB (Red-Green-Blue) yang diperlukan MediaPipe, atau ke grayscale untuk proses tertentu.
- 3) Gaussian Blur: Penerapan filter Gaussian untuk mereduksi noise pada citra. Gaussian blur bekerja dengan merata-ratakan nilai piksel berdasarkan distribusi Gaussian, sehingga menghaluskan detail yang tidak diperlukan tanpa menghilangkan tepi objek utama.
- 4) Normalisasi Intensitas Piksel: Penyesuaian rentang nilai intensitas piksel ke skala standar (0-255) untuk meningkatkan konsistensi di berbagai kondisi pencahayaan.

### **2.5.2 Deteksi Wajah**

Deteksi wajah adalah proses mengidentifikasi dan melokalisasi keberadaan wajah manusia dalam sebuah citra atau frame video. Dalam penelitian ini, deteksi wajah menjadi tahap fundamental yang menentukan keberhasilan seluruh pipeline, karena seluruh analisis perilaku visual bergantung pada keakuratan lokalisasi area wajah dan mata peserta ujian.

## **2.6 MediaPipe Face Mesh**

MediaPipe Face Mesh adalah solusi machine learning yang dikembangkan oleh Google untuk mendeteksi 468 titik landmark wajah tiga dimensi secara akurat dan efisien dari sebuah citra atau video. Pustaka ini memanfaatkan arsitektur jaringan saraf tiruan yang telah dilatih pada dataset wajah berskala besar, sehingga

mampu beroperasi secara real-time bahkan pada perangkat komputasi dengan spesifikasi terbatas (Effendi and Widyaningrum, 2026).

Keunggulan utama MediaPipe Face Mesh dibandingkan pendekatan konvensional adalah kemampuannya mendeteksi landmark wajah secara tiga dimensi, sehingga estimasi pose kepala dan arah pandang dapat dilakukan lebih akurat. Selain itu, dengan mengaktifkan parameter `refine_landmarks`, MediaPipe dapat mengekstrak titik-titik landmark iris secara spesifik, yang menjadi dasar estimasi gaze vector dalam penelitian ini.

### 2.6.1 Landmark Iris untuk Estimasi Gaze

MediaPipe Face Mesh menyediakan empat titik landmark iris untuk masing-masing mata (indeks 469-472 untuk mata kanan dan 474-477 untuk mata kiri). Posisi sentroid keempat titik tersebut merepresentasikan pusat iris, yang kemudian digunakan untuk menghitung posisi relatif iris terhadap sudut-sudut mata. Nilai rasio posisi horizontal iris terhadap lebar mata menjadi dasar estimasi arah pandang (Amadea, Bachtiar and Pangestu, 2022).

### 2.6.2 Eye Aspect Ratio (EAR)

Eye Aspect Ratio (EAR) adalah metrik yang diperkenalkan oleh (Bakunah and Mohammed Baneamoon, 2021) untuk mendeteksi kondisi mata tertutup (kedipan) secara real-time. EAR dihitung berdasarkan rasio jarak vertikal terhadap jarak horizontal mata menggunakan persamaan berikut:

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2 \times \|p_1 - p_4\|}$$

di mana p1 hingga p6 merupakan titik-titik landmark kelopak mata. Nilai EAR mendekati nol ketika mata tertutup dan bernilai sekitar 0,25-0,30 dalam kondisi normal. Nilai EAR di bawah threshold tertentu (pada penelitian ini 0,20) mengindikasikan bahwa mata dalam kondisi menutup, yang digunakan dalam perhitungan PERCLOS.

## 2.7 Fitur Perilaku Visual dalam Deteksi Kecurangan

Ekstraksi fitur perilaku visual merupakan inti dari sistem deteksi kecurangan berbasis eye tracking. Fitur-fitur ini merepresentasikan pola perilaku pandang mata peserta ujian yang dapat menjadi indikator kecurangan. Tabel 2.1 merangkum fitur-fitur perilaku visual yang digunakan beserta definisi dan threshold deteksinya.

**Tabel 2.1 Fitur Perilaku Visual dalam Sistem Deteksi Kecurangan**

| Fitur                         | Definisi                                                                      | Cara Pengukuran                                                      | Threshold Kecurangan                                        |
|-------------------------------|-------------------------------------------------------------------------------|----------------------------------------------------------------------|-------------------------------------------------------------|
| Fiksasi (Fixation)            | Kondisi di mana pandangan mata berhenti pada satu titik dalam durasi tertentu | Durasi dan frekuensi pandangan pada zona layar                       | Fiksasi di luar layar > 3 detik berturut-turut              |
| Sakade (Saccade)              | Gerakan cepat bola mata dari satu titik fiksasi ke titik lainnya              | Jumlah perubahan arah pandang yang terdeteksi per menit              | Frekuensi sakade lateral > 15 kali per sesi                 |
| Arah Pandang (Gaze Direction) | Arah vektor pandangan mata terhadap bidang layar komputer                     | Posisi relatif iris terhadap sudut mata (horizontal)                 | Pandangan ke kiri atau kanan melampaui threshold rasio 0,25 |
| PERCLOS                       | Percentage of Eye Closure: rasio waktu kelopak mata dalam kondisi menutup     | Rasio frame dengan EAR di bawah nilai threshold terhadap total frame | PERCLOS > 30% mengindikasikan tidak fokus atau mengantuk    |

*Sumber: Diolah dari (Yogasara et al., 2023)*

### **2.7.1 PERCLOS (Percentage of Eye Closure)**

PERCLOS merupakan metrik yang mengukur proporsi waktu di mana kelopak mata berada dalam kondisi menutup, sebagai indikator tingkat kewaspadaan (alertness) seseorang. Nilai PERCLOS dihitung sebagai rasio jumlah frame dengan EAR di bawah threshold terhadap total frame yang diproses. Nilai PERCLOS yang tinggi mengindikasikan bahwa peserta ujian berada dalam kondisi mengantuk atau tidak fokus, yang dapat berkorelasi dengan perilaku mencurigakan.

### **2.7.2 Metode Deteksi Berbasis Threshold (Rule-Based)**

Metode deteksi berbasis threshold atau rule-based merupakan pendekatan deterministik di mana keputusan klasifikasi dibuat berdasarkan seperangkat aturan eksplisit yang membandingkan nilai fitur dengan ambang batas yang telah ditetapkan. Berbeda dengan pendekatan machine learning yang memerlukan proses pelatihan model, metode rule-based bersifat transparan, mudah diinterpretasikan, dan dapat diimplementasikan tanpa memerlukan dataset berlabel dalam jumlah besar.

Dalam konteks penelitian ini, sistem menerapkan tiga aturan deteksi kecurangan yang didasarkan pada kajian literatur dan karakteristik perilaku visual yang mengindikasikan kecurangan. Sistem menandai seorang peserta sebagai terindikasi curang apabila memenuhi satu atau lebih dari kondisi berikut: (1) pandangan berada di luar area layar secara berturut-turut selama lebih dari 3 detik, (2) frekuensi sakade lateral melebihi 15 kali per sesi pengujian, atau (3) nilai PERCLOS melampaui 30 persen dari total waktu pengujian.

Kelebihan pendekatan rule-based dalam konteks penelitian ini antara lain adalah transparansi dalam pengambilan keputusan (setiap keputusan dapat ditelusuri langsung ke nilai fitur dan aturan yang berlaku), tidak memerlukan proses pelatihan model sehingga lebih mudah diterapkan, serta mampu memberikan penjelasan yang jelas mengenai alasan suatu peserta terdeteksi sebagai curang.

## 2.8 Evaluasi Kinerja Sistem dengan Confusion Matrix

Evaluasi kinerja sistem klasifikasi biner seperti deteksi kecurangan umumnya dilakukan menggunakan confusion matrix, yang merupakan tabel kontingensi yang merangkum hasil prediksi sistem terhadap label aktual (ground truth). Confusion matrix memberikan gambaran yang komprehensif mengenai kinerja sistem, tidak hanya dari sisi ketepatan prediksi secara keseluruhan, tetapi juga dari kemampuan mendeteksi kecurangan yang sebenarnya dan menghindari kesalahan identifikasi.

**Tabel 2.2 Struktur Confusion Matrix Sistem Deteksi Kecurangan**

|                         |        | Prediksi Positif (Curang)                                               | Prediksi Negatif (Normal)                                 |
|-------------------------|--------|-------------------------------------------------------------------------|-----------------------------------------------------------|
| Aktual Positif (Curang) | Curang | True Positive (TP): Sistem benar mendeteksi curang                      | False Negative (FN): Sistem gagal mendeteksi curang       |
| Aktual Negatif (Normal) | Normal | False Positive (FP): Sistem salah mengira peserta normal sebagai curang | True Negative (TN): Sistem benar mengenali peserta normal |

### 2.8.1 Metrik Evaluasi

Berdasarkan nilai True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) yang diperoleh dari confusion matrix, dihitung empat metrik evaluasi utama sebagai berikut:

- 1) Akurasi (Accuracy): Mengukur proporsi prediksi yang benar dari keseluruhan data. Akurasi dihitung menggunakan persamaan:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

- 2) Presisi (Precision): Mengukur proporsi prediksi curang yang benar dari seluruh prediksi curang yang dihasilkan sistem. Presisi yang rendah mengindikasikan banyaknya peserta normal yang salah diklasifikasikan sebagai curang.

$$Precision = TP / (TP + FP)$$

- 3) Recall (Sensitivity): Mengukur proporsi peserta yang benar-benar curang yang berhasil dideteksi oleh sistem. Recall yang rendah mengindikasikan banyaknya kasus kecurangan yang tidak terdeteksi.

$$Recall = TP / (TP + FN)$$

- 4) F1-Score: Rata-rata harmonik dari presisi dan recall, yang memberikan gambaran keseimbangan antara kemampuan mendeteksi kecurangan dan menghindari kesalahan klasifikasi.

$$F1-Score = 2 \times (Precision \times Recall) / (Precision + Recall)$$

- 5) Spesifisitas (Specificity): Mengukur proporsi peserta normal yang berhasil dikenali dengan benar oleh sistem, sehingga sistem tidak salah menuduh peserta jujur.

$$Specificity = TN / (TN + FP)$$