1. Pendahuluan

1.1 Latar Belakang

Wallhack adalah jenis cheat yang digunakan dalam game FPS (first-person shooter) yang memungkinkan pemain untuk melihat melalui dinding atau objek lain yang biasanya tidak tembus pandang. Cheat ini memberikan keuntungan yang signifikan bagi pemain yang menggunakannya, karena mereka dapat melihat posisi dan gerakan lawan yang tersembunyi di balik penghalang, sehingga meningkatkan kemampuan mereka untuk menghindari atau menyerang musuh dengan lebih efektif. Penggunaan wallhack sangat merugikan bagi integritas permainan, karena memberikan keuntungan yang tidak adil dan merusak pengalaman bermain bagi pemain lain yang bermain secara jujur, Selain itu, Wallhack tidak direkomendasikan dalam real-game karena sangat bertentangan dengan prinsip fairplay dan dapat merusak komunitas pemain yang merasa dirugikan akan kehilangan minat untuk bermain permainan tersebut yang pada akhirnya dapat mengurangi jumlah pemain aktif dan mempengaruhi tidak hanya reputasi, keberlangsungan juga monetisasi dari developer itu sendiri.

Adversarial attack atau serangan adversarial adalah teknik di mana penyerang dengan sengaja menciptakan atau memodifikasi data input dengan tujuan untuk menipu atau mengecoh sistem yang berbasis machine learning, seperti deep neural network. Dalam konteks ini, serangan adversarial berupa penambahan noise yang dihasilkan saat augmentasi data. Tujuan dari penambahan noise ini adalah untuk mengelabui sistem, seperti sistem deteksi kecurangan pada permainan, sehingga menghasilkan prediksi yang salah atau merusak kinerja sistem yang bergantung pada pemrosesan data. Serangan adversarial pada lingkungan permainan dapat berdampak serius, seperti memberikan keuntungan yang tidak adil kepada pemain yang melakukan kecurangan atau merusak integritas kompetisi online khususnya pada game First Person Shooter dimana reaksi pemain sangat diandalkan.

Pada Tahun 2011 Galli dan koleganya melakukan penelitian Cheat Detection Famework untuk game Unreal Tournament III Menggunakan metode machine Learning dalam paper tersebut Galli dan koleganya menggunakan beberapa metode seperti SVM (Support Vector Machine) , Naïve bayes , Random Forest, Neural Network dan Breiman Random Forest [1] namun saat ini sudah banyak sistem anti-deteksi yang berfokus pada fitur yang bersifat non-visual [2] .Oleh karena itu, penting untuk mengembangkan metode yang efektif untuk mendeteksi dan mengatasi serangan ini guna menjaga keadilan dan integritas dalam lingkungan permainan. Bahkan persentase kecil dari pemain yang curang dapat merusak pengalaman bermain game bagi sebagian besar pemain: jika 6% pemain melakukan kecurangan, kemungkinan bertemu dengan setidaknya satu pemain curang adalah 42,7% dalam pertandingan 5 vs 5 (seperti CS:GO atau Valorant) dan 49,4% dalam pertandingan 6 vs 6 (seperti Overwatch). [2] Telah Dilakukan Penelitian Terkait Pendeteksian Adversarial Attack menggunakan DNN (Deep Neural Network)

Pada tahun 2021 A. Jonnalagadda dan koleganya melakukan penelitian , Penelitian ini fokus pada pengembangan metode deteksi kecurangan secara visual dalam game kompetitif. Metode yang diusulkan didasarkan pada penggunaan deep neural network untuk mendeteksi pemain yang menggunakan cheat dengan hasil IBP Bound maksimum 72% dan hasil minimum 55% , Iterative maksimum 99% minimum 58%, FGSM 62% dan Hasil PalingBesar 80% [2].

Pada tahun 2020 Pinto dan koleganya menguji dan memvalidasi kemampuan sistem dalam mendeteksi kecurangan pada pemain yang sebelumnya tidak pernah dikenal. Dalam eksperimen, model yang dikembangkan berhasil mencapai akurasi rata-rata sebesar 99,2% untuk triggerbot [2], triggerbot sendiri adalah triggerbot adalah jenis perangkat lunak atau skrip yang digunakan dalam beberapa video game, khususnya dalam permainan FPS (first-person shooter) , triggerbot sendiri bekerja dengan mengaitkan kode perangkat lunak dengan input dari mouse atau tombol tertentu pada keyboard.Ketika triggerbot mendeteksi bahwa crosshair atau bidikan pemain melewati karakter lawan atau target, itu secara otomatis memicu tindakan menembakkan senjata dengan sangat cepat dan akurat, memberikan keuntungan pemain dalam merespons dan mengalahkan lawan dengan kecepatan yang lebih tinggi daripada pemain manusia biasa. dan 98,9% untuk aimbot [2], Aimbot adalah jenis perangkat lunak atau skrip yang digunakan dalam video game, terutama dalam permainan FPS (first-person shooter), untuk memberikan keunggulan kompetitif kepada pemain. Aimbot dirancang untuk

meningkatkan keakuratan menembak pemain dengan secara otomatis mengarahkan dan mengunci bidikan atau crosshair mereka pada target lawan. , Aimbot bekerja dengan menggunakan pemrograman yang rumit untuk mengidentifikasi posisi target lawan dalam permainan. Setelah target terdeteksi, aimbot secara otomatismengarahkan crosshair pemain ke target tersebut, memungkinkan pemain untuk menembak dengan tingkat akurasi yang tinggi dan konsisten. Beberapa aimbot bahkan dapat menyesuaikan pengaturan sensitivitas dan faktor lainnya untuk meningkatkan keakuratan tembakan , dua jenis kecurangan yang umum terjadi. Pendekatan ini juga dapat diterapkan pada berbagai jenis permainan atau metode input, serta tugas-tugas lain yang terkait dengan pemodelan aktivitas manusia.[3] .

Pada Tahun 2017 Feinman dan koleganya mengembangkan metode deteksi adversarial yang implisit yang tidak tergantung pada algoritma serangan yang digunakan. Penulis melakukan evaluasi terhadap metode ini menggunakan berbagai dataset standar termasuk MNIST dan CIFAR-10, dan menunjukkan bahwa metode tersebut umumnyaberlaku dengan baik pada berbagai arsitektur dan serangan. Hasil penelitian menunjukkan bahwa ROCAUC sebesar 85-93% dapat dicapai dalam sejumlah tugas klasifikasi standar dengan menggunakan kelas negatif yang terdiri dari sampel-sampel normal dan sampel- sampel noise.[4] Untuk mengatasi tantangan ini, penelitian telah dilakukan untuk mengembangkan teknik dan pendekatan dalam mendeteksi kecurangan dalam game dan menghadapi adversarial attack dalam lingkungan game. Penelitian ini menggunakan metodevisual atau computer vision, di mana deteksi bounding box yang diambil dari replay game pemain akan digunakan sebagai dasar klasifikasi untuk menentukan apakah pemain tersebutmenggunakan cheat atau tidak. Untuk mengatasi masalah ini penelitian relah dilakukan dalam mendeteksi kecurangan dalam game. Penelitian ini menggunakan metode visual dimana replay game dari pemain digunakan sebagai dasar klasifikasi umtuk menentukan apakah pemain tersebut menggunakan cheat atau tidak, Metode visual seperti ini masih jarang digunakan dalam sistem deteksi kecurangan, terutama sistem dengan basis framework YOLO , khusus nya YOLOv8 , framework ini memiliki potensi yang bagus dalam mendeteksi cheat secara visual namum dalam konteks Wallhack Detection masih sangat terbatas, guna meningkatkan efektivitas deteksi cheat dalam game

1.2 Rumusan Masalah

1. Bagaimana cara mendeteksi kecurangan dalam permainan First-Person Shooter (FPS) yang menggunakan wallhack dengan memanfaatkan algoritma deteksi objek YOLOv8 dan metode visual lain dalam konteks serangan adversarial?
2. Apa perbedaan dalam akurasi deteksi antara model YOLOv8 yang dilatih dengan data yang telah ditambahkan adversarial attack (noise) dan model YOLOv8 yang dilatih dengan data asli tanpa augmentasi?
3. Bagaimana performa model deteksi lainnya, seperti SVM, Decision Tree, Naive Bayes, dan K-means, dalam mengidentifikasi cheat jika dibandingkan dengan YOLOv8 dalam konteks deteksi wallhack yang terpengaruh oleh adversarial attack?

1.3 Tujuan

1. Mengevaluasi efektivitas YOLOv8 dalam mendeteksi wallhack dalam permainan FPS dengan membandingkan hasil deteksi menggunakan data yang telah ditambahkan noise atau adversarial attack. Fokus utama adalah pada perbandingan antara model yang dilatih dengan dataset yang telah dimodifikasi dengan noise dan model yang dilatih tanpa modifikasi tersebut.
2. Membandingkan kinerja model YOLOv8 yang dilatih dengan dataset yang telah ditambahkan noise dan model yang dilatih tanpa noise untuk menilai apakah augmentasi data melalui adversarial attack meningkatkan kemampuan deteksi terhadap wallhack.
3. Membandingkan efektivitas berbagai model klasifikasi dalam mendeteksi cheat yang terpengaruh oleh adversarial attack, serta mengukur perbedaan akurasi di antara model-

model tersebut

1.4 Batasan Masalah

1. Penelitian ini hanya fokus pada deteksi wallhack dan tidak mencakup jenis kecurangan lain dalam permainan FPS, seperti aimbot atau triggerbot.
2. Dataset yang digunakan dalam penelitian ini terdiri dari video gameplay dengan dua kategori: pemain yang menggunakan wallhack dan pemain yang tidak menggunakan cheat. Dataset ini mencakup total 2000 gambar, dengan 1600 gambar untuk pelatihan (800 cheat dan 800 bersih) dan 200 gambar untuk pengujian (100 cheat dan 100 bersih). Data ini diambil dari rekaman video gameplay yang sudah ada, tanpa melibatkan data gameplay baru yang mungkin memiliki kecurangan atau cheatterbaru.
3. Penelitian ini hanya menggunakan teknik augmentasi data yang telah tersedia dalam Roboflow untuk menghasilkan adversarial attack (noise). Tidak ada eksplorasi lebih lanjut terhadap teknik adversarial attack yang lebih kompleks atau metode serangan yang berbeda.
4. Fokus utama penelitian ini adalah pada deteksi kecurangan menggunakan YOLOv8 sebagai model deteksi objek. Perbandingan dilakukan antara hasil deteksi menggunakan model YOLOv8 dengan dan tanpa adversarial attack, tanpa melibatkan metode deteksi lain di luar YOLOv8 dalam analisis.
5. Penelitian ini membatasi penilaian kinerja sistem hanya pada metrik akurasi deteksi, tanpa mempertimbangkan metrik lain seperti kecepatan deteksi atau penggunaan sumber daya komputasi.
6. Implementasi sistem deteksi hanya dilakukan pada platform dan lingkungan yang ditentukan, yaitu pada data yang sudah ada dan menggunakan alat serta framework yang telah dipilih (misalnya YOLOv8 dan Roboflow). Penelitian ini tidak mencakup pengujian di platform permainan nyata atau di lingkungan yang berbeda dari yang telah ditetapkan.
7. Meskipun penelitian ini mencakup penggunaan model klasifikasi seperti SVM, Decision Tree, Naive Bayes, dan K-means, evaluasi utama difokuskan pada perbandingan antara YOLOv8 dengan dan tanpa adversarial attack, bukan pada perbandingan langsung antara berbagai model klasifikasi.