كشف الغش في الامتحانات عبر الإنترنت بناءً على تحليل الفيديو باستخدام التعلم العميق

طالبة الدراسات العليا: م. بشرى السطلي ماجستير علوم الويب – الجامعة الافتراضية

اشراف الدكتور: محمد بسام الكردي

الملخص:

تهدف هذه الدراسة إلى تطوير نظام ذكي لرصد حالات الغش في الامتحانات الإلكترونية بالاعتماد على تحليل الفيديو المباشر باستخدام تقنيات التعلم العميق والرؤية الحاسوبية. تم تطوير ثلاثة نماذج تصنيفية مستقلة تختلف في نوع المدخلات والبنية المعمارية، ثم تحويلها إلى صيغة ONNX (Open Neural Network Exchange) لضمان قابلية التشغيل عبر بيئات مختلفة.

ولتحقيق استجابة زمنية فورية، نُشر النظام عبر واجهة ويب تفاعلية متعددة المسالك (Multithreading)، مع دمج وحدات إضافية لكشف الكائنات المشبوهة باستخدام YOLOv9بعد تدريبها على قاعدة بيانات خاصة بالامتحانات مما حسّن دقة الكشف والتعرف على سلوكيات الغش. كما تم تتبع حركة الفم عبر مؤشر MAR (Mouth Aspect Ratio) لرصد إذا كان الطالب يتحدث أثناء الامتحان.

أظهرت النتائج تفوق النموذج الهجين على النماذج الأخرى محققاً دقة تصنيف بلغت 98.8% مع غياب كامل للأخطاء السلبية (False Negative=0) مما يعزز موثوقية النظام في بيئات الامتحانات الحساسة ويدعم اعتماده في البيئات التعليمية الرقمية.

الكلمات المفتاحية:

كشف الغش، الامتحانات عبر الانترنت، تحليل الفيديو، YOLOv9، مؤشر نسبة أبعاد الفم MAR

Online exams cheating detection based on video analysis using deep learning

Abstract

This study aims to develop of an intelligent system for detecting cheating in online examinations through real-time video stream analysis using deep learning and computer vision techniques. Three independent classification models were developed, each differing in input types and architectural structures, and subsequently converted into the ONNX (Open Neural Network Exchange) format to ensure cross-platform compatibility.

To achieve real-time responsiveness, the system was deployed through a multi-threaded interactive web-based interface. Additional modules were integrated to enhance functionality. These include suspicious objects detection using YOLOv9, retrained on a custom examination dataset to improve detection accuracy and better identify cheating behaviours, as well as mouth movement tracking based on the Mouth Aspect Ratio (MAR) indicator to determine whether students were speaking during the exam.

The experimental results demonstrated that the hybrid model outperformed the other models, achieving a classification accuracy of 98.8% with a complete absence of false negatives. These findings reinforce the reliability of the system in sensitive examination settings and support its adoption in digital educational platforms.

Keywords

Cheating detection, Online exams, Video Analysis, YOLOv9, Mouth Aspect Ratio (MAR).

1. **مقدّمة:**

2. **هدف البحث:**

يهدف هذا البحث إلى تطوير نظام ذكي ومتكامل لرصد حالات الغش في الامتحانات الإلكترونية بالاعتماد على تقنيات التعلم العميق والرؤية الحاسوبية، وذلك من خلال تحقيق الأهداف التالية:

- **تطوير ثلاثة نماذج تصنيفية مستقلة** تختلف في نوع المدخلات والبنى المعمارية، تشمل نموذجاً يعتمد على السمات الهندسية المستخرجة من النقاط المرجعية للوجه، ونموذجاً يعتمد على الصورة الأصلية للوجه مدعوماً بآلية انتباه قنوي لاستخلاص السمات البصرية، إضافة إلى نموذج هجين متعدد المسارات يجمع بين المدخلين.
- **دمج وحدات إضافية داعمة** تشمل كشف الكائنات المشبوهة (مثل وجود أشخاص، هواتف، كتب) باستخدام YOLOv9 وتتبع حركة الفم باستخدام مؤشر نسبة أبعاد الفم (MAR) لرصد حالات التحدث غير المصرّح بها وحساب وتتبع اتجاه الرأس (Head Pose: Yaw/Pitch/Roll) لرصد حالات الانحراف غير الطبيعي أثناء الامتحان.
- **تقييم أداء النماذج وفق مقاييس معيارية** مثل الدقة (Accuracy) والاستدعاء (**Recall**)، وتحديد النموذج الأكثر فعالية في بيئات الامتحانات الحساسة مع التركيز على تقليل الأخطاء السلبية (False Negative) التي قد تؤثر على مصداقية النظام.
- **اختبار قابلية النظام للتطبيق العملي في بيئات تعليمية حقيقية** من خلال قياس زمن الاستجابة (Latency) ومعدل الإطارات في الثانية (FPS) لضمان ملاءمته للاستخدام الفعلي في الزمن الحقيقي.

وبناءً على ذلك يسعى هذا البحث إلى تعزيز النزاهة الأكاديمية من خلال توفير حل عملي وموثوق يحد من محاولات الغش ويدعم اعتماد البيئات التعليمية الرقمية مع قابلية التوسع والتخصيص بما يتناسب مع متطلبات المؤسسات التعليمية المختلفة.

3. دراسات مرجعيّة:

تزايد الاهتمام في السنوات الأخيرة بتطوير أنظمة ذكية لرصد حالات الغش في الامتحانات الإلكترونية نتيجة الانتشار الواسع للتعليم عن بُعد والاختبارات الرقمية. وقد تنوعت الدراسات في هذا المجال بين نماذج هندسية تعتمد على السمات الوجهية وأخرى بصرية تعتمد على الصور إضافة إلى نماذج هجينة تجمع بين أكثر من مصدر بيانات وأنظمة متكاملة تسعى إلى دمج وحدات متعددة في إطار واحد.

الدراسات الهندسية:

في دراسة [3] تم تطوير نظام لرصد السلوك غير الطبيعي في الامتحانات الإلكترونية باستخدام تحليل الفيديو عبر Mediapipeحيث تم تحويل النقاط المرجعية للوجه وتحويلها إلى متجه سمات هندسية مكون من 171 بُعداً باستخدام المسافات الاقليدية بين النقاط. وقد حقق النظام دقة بلغت 77.9% مع زمن استدلال منخفض جداً 0.01 ثانية لكل إطار، مما يتيح المعالجة في الزمن الحقيقي. ورغم سرعة النظام إلاّ أنّ دقته تبقى محدودة نسبياً ولا يتضمن وحدات داعمة مثل كشف الأغراض أو تتبع التحدث مما يحد من شموليته. بينما ركزت الدراسة [5] على التحقق من هوية الطالب باستخدام KNN وكشف الهاتف عبر YOLOV3 وتتبع العين باستخدام Mediapipe Iris وحققت دقة بلغت 93.9% في الكشف عن جميع أنواع الغش بمعدل FAR ثابت 5% ورغم أن الشبكة العصبونية الالتفافية CNN توفر دقة أعلى بكثير منKNN ، إلا أنه لم يتم استخدامها في هذا النظام بسبب متطلباته المحدودة مثل إضاءة جيدة في قاعة الامتحان ومسافة محددة بين الطالب والكاميرا مما يجعل KNN مناسبة للعمل ضمن هذه المتطلبات.

الدراسات البصرية:

اقترحت دراسة [4] بنية متعددة المراحل تشمل التحقق من الهوية باستخدام FaceNet وتتبع حركة الرأس عبر خوارزمية Lucas Kanade وتقدير اتجاه النظر باستخدام نموذج AAM

(Active Appearance Model)، وقد حقق النظام F1-Score=0.94 على بيانات مأخوذة من أكثر من 1000 جلسة امتحانية. ومع ذلك لم تقدم الدراسة تفاصيل دقيقة حول زمن المعالجة أو معدل الأخطاء السلبية، كما أنّ اعتمادها على خوارزميات متفرقة دون دمج ضمن نموذج موحد يحد من كفاءتها في الزمن الحقيقي. في المقابل أجرى [9] تحليلاً مقارناً شاملاً لمجموعة من نماذج التعلم العميق المطبقة في بيئات المراقبة الإلكترونية للامتحانات حيث تم اختبار نماذج مثل EfficientNet، ResNet، MobileNetV2 وYOLOv5 على مجموعتي بيانات متخصصتين OEP,OP وقد أظهرت النتائج أنّ نموذج EfficientNetB1 حقق أعلى دقة بلغت 94.59% على مجموعة بيانات OP بينما حقق YOLOv5 أداءً متميزاً بمعدل98.3%= mAP@0.5 كما حقق ResNet50-CBAM دقة بلغت 93.61% على مجموعة OEP. ورغم أن هذه الدراسة وفرت مؤشرات كمية دقيقة وأظهرت قوة النماذج في الكشف عن سلوكيات الغش، إلا أنها اقتصرت على المقارنة النظرية للأداء دون تقديم نظام متكامل قابل للتطبيق المباشر في بيئات امتحانية فعلية.

الدراسات الهجينة:

أظهرت دراسة [6] فعالية عالية في دمج البيانات المرئية والصوتية باستخدام الشبكات العصبونية الالتفافية (CNN) حيث حققت دقة بلغت 99.8% للكاميرا الأمامية و98.7% للكاميرا الخلفية. ورغم قوة الدمج بين المصادر المتعددة إلاّ أنّ النظام يعاني من تعقيد البنية والحاجة لموارد حسابية كبيرة، مع غياب تفاصيل دقيقة حول زمن المعالجة أو قدرة التوسع للتعامل مع عدد كبير من المستخدمين بشكل متزامن. أما دراسة [8] فقد أكدت على أهمية دمج تقنيات متعددة مثل Deep Learning وMachine Learning لتحقيق دقة أعلى وصلت إلى نحو 96%، إلاّ أنّ هذا الدمج زاد التعقيد الحسابي وصعّب التطبيق في الزمن الحقيقي، كما لم تُذكر تفاصيل عن زمن الاستجابة أو قدرة النظام على التوسع مما يحد من إمكانية تطبيقه في بيئات امتحانية واسعة النطاق. من ناحية أخرى طور [10] إطاراً هجيناً

يجمع بين YOLOv7 لكشف الأغراض وDeepFace للتعرف على الوجه ضمن تطبيق ويب تفاعلي في الزمن الحقيقي ويدعم تعدد الكاميرات والتوسع الافقي. وقد أظهر النظام دقة عالية مع زمن معالجة منخفض نسبياً. ورغم أنّ هذه الدراسة لا تستهدف كشف الغش في الامتحانات الإلكترونية بشكل مباشر، إلاّ أنها تقدم نموذجاً تقنياً متقدماً يمكن الاستفادة من بنيته في هذا المجال وخصوصاً فيما يتعلق بآليات التكامل بين وحدات الرؤية الحاسوبية والواجهات التفاعلية في بيئات تعليمية رقمية.

الأنظمة المتكاملة:

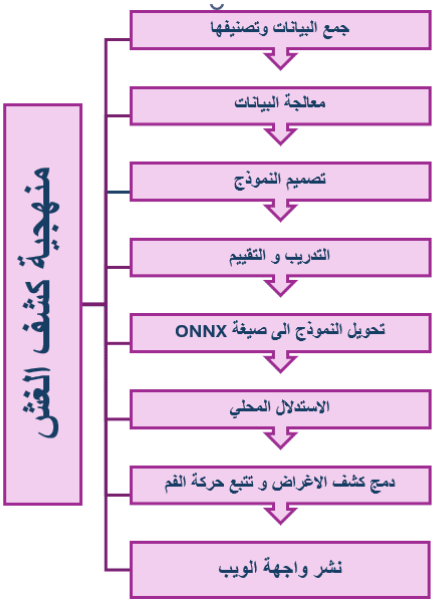
قدمت دراسة [7] نظاماً شاملاً يعتمد على YOLO لكشف الهواتف والكتب إضافة إلى تحليل النظر وحالة الفم. ركّز النظام على الاكتشاف المباشر للسلوكيات المشبوهة دون تصنيف الطلاب، ولم يقدم مؤشرات دقيقة للأداء (مثل الدقة أو زمن الاستجابة) مما يحد من إمكانية تقييم كفاءته بشكل كامل.

من خلال مراجعة الأدبيات يتضح أن أنظمة كشف الغش في الامتحانات الإلكترونية شهدت تطوراً ملحوظاً باستخدام تقنيات الذكاء الصنعي والرؤية الحاسوبية، حيث تنوعت بين نماذج تعتمد على السمات الهندسية وأخرى على الصور البصرية، إضافة إلى محاولات دمج البيانات المتعددة لتحقيق دقة أعلى. ورغم أنّ هذه الدراسات قدمت نتائج جيدة، إلا أنها مازالت تعاني من قيود واضحة مثل محدودية السمات المدخلة، غياب التكامل بين الوحدات المختلفة، الحاجة إلى موارد حسابية كبيرة، أو التركيز على المقارنة النظرية دون تقديم نظام متكامل قابل للتطبيق العملي في البيئات التعليمية. هذه الفجوات البحثية تبرز الحاجة إلى نظام شامل يجمع بين الدقة العالية والقدرة على العمل في الزمن الحقيقي مع قابلية التوسع والنشر في بيئات امتحانية فعلية. ومن هنا تأتي أهمية هذا البحث الذي يسعى إلى تقديم إطار عملي متكامل يعتمد على نموذج هجين يدمج بين المدخلات الهندسية والبصرية، ومدعوماً بوحدات إضافية مثل كشف الكائنات وتتبع حركة الفم، بما يعزز النزاهة الاكاديمية ويوفر حلاً موثوقاً وقابلاً للتطبيق في المؤسسات التعليمية الرقمية.

4. مواد وطرق البحث:

4.1 منهجية البحث:

تم تنفيذ هذه الدراسة وفق منهجية تجريبية تعتمد على تحليل الفيديو المباشر باستخدام نماذج تصنيفية مبنية على تقنيات التعلم العميق بالإضافة إلى وحدتين داعمتين: وحدة لكشف الأغراض المشبوهة باستخدام YOLOووحدة تتبع حركة الفم باستخدام مؤشرMAR نسبة أبعاد الفم. وقد شملت مراحل البحث جمع البيانات، المعالجة المسبقة، تصميم النماذج، تدريبها وتقييمها، تحويلها إلى صيغة قابلة للنشر، وأخيراً دمجها ضمن واجهة تشغيل تفاعلية تعمل في الزمن الحقيقي. تم التركيز على بناء نظام متكامل يجمع بين الدقة، الكفاءة الزمنية، والمرونة التشغيلية مع قابلية التوسع والتكامل في بيئات الامتحانات الرقمية.



الشكل 1 المخطط العام لمنهجية كشف الغش

- **المخطط العام للمنهجية:**

يوضح الشكل 1 المنهجية العلمية والتطبيقية لبناء نظام ذكي لكشف حالات الغش في الاختبارات عبر الانترنت، بالاعتماد على تحليل الفيديو المباشر للجلسة وتوظيف تقنيات التعلم العميق. تشمل المنهجية مراحل متكاملة تبدأ بجمع البيانات وتصنيفها، مروراً بالمعالجة المسبقة واستخراج السمات الهندسية للوجه، ثم تصميم النموذج التصنيفي وتحويله إلى صيغة معيارية قابلة للنشر ONNX، وصولاً إلى دمج آليات كشف الأغراض وتتبع حركة الفم، وانتهاءً بمرحلة الاستدلال الفوري ونشر النظام عبر واجهة ويب تفاعلية.

4.2 الأدوات والمكتبات والتعاريف اللازمة

تم تطوير المنهجية المقترحة باستخدام لغةPython (الإصدار3.8 أو أعلى) ضمن بيئة تطوير تشملVS Code وJupyter Notebook مع دعم وحدة معالجة رسومية (GPU(CUDA على نظام التشغيل Windows. اعتمدت التجارب على مكتبات الذكاء الصنعي ومعالجة الفيديو والصور، بما في ذلك مكتبة PyTorch لتصميم وتدريب النموذج، ONNX Runtime للتشغيل عبر منصات متعددة، وYOLOv9لاكتشاف الوجوه والكائنات. كما استخدمت OpenCV وDlib *لمعالجة الفيديو والصور* واستخراج النقاط المرجعية للوجه، *إضافة إلى* Pandas وSeaborn/*Matplotlib لحفظ النتائج وتحليلها وعرض مؤشرات الأداء*. ساهم هذا التكامل بين الأدوات والمكتبات في تحقيق أداء عالٍ في الزمن الحقيقي مع قابلية التوسع والتطبيق في بيئات تعليمية متنوعة.

4.3 جمع البيانات وتقسيمها:

تم استخدام مجموعة بيانات عامة تُعرف باسم OEP [12] والتي تحتوي على تسجيلات فيديو ل 24 طالباً من جامعة ولاية ميشيغان أثناء أداء امتحان رياضيات وتشمل حالات غش وغير غش. تتضمن المجموعة تسجيلات من كاميرتين wearcam وwebcam وقد تم الاعتماد في هذه الدراسة على فيديوهاتwebcam فقط بمتوسط زمن قدره 17 دقيقة لكل طالب. بلغت نسبة الإطارات المستخرجة من قاعدة البيانات العامة حوالي 90% تقريباً بينما مثلت الإطارات المحلية الخاصة ما يقارب 10% من إجمالي البيانات وذلك لتغطية سيناريوهات إضافية مثل استخدام الكتب والأوراق. تم استخراج الإطارات من الفيديوهات وتصنيفها يدوياً باستخدام واجهة تفاعلية تم تطويرها خصيصاً لهذا الغرض. ثم حفظ الإطارات المصنفة في مجلدات منفصلة حسب الفئة (غش/ ليس غش).

أما عملية التقسيم فقد اعتمدت على خوارزمية مخصصة تضمن استقلالية المجموعات وتمنع تسرب المعلومات بين بيانات التدريب والاختبار. حيث جُمعت الإطارات في مجموعات متسلسلة استناداً إلى الأرقام المدرجة في أسماء الملفات بحيث تمثل كل مجموعة سلسلة زمنية لمشارك واحد. بعد ذلك تم خلط عشوائي على مستوى المجموعات (وليس على مستوى الإطارات الفردية) لضمان توزيع غير متحيز، وتم توزيع المجموعات بالكامل على أحد التقسيمين حتى الوصول إلى

النسبة المستهدفة وهي 70% للتدريب و30% للاختبار. وبذلك بلغ عدد عينات التدريب 2718 غش و1699 ليس غش بينما بلغ عدد عينات الاختبار 1665 غش و1314 ليس غش وهو ما يعكس درجة من التوازن النسبي بين الفئتين. ولضمان عدالة النموذج بشكل أكبر أثناء التعلم تمت إضافة أوزان للفئات داخل دالة الخسارة واستخدام WeightedRandomSampler في مرحلة التدريب لمعالجة أي انحياز متبقٍ.

4.4 المعالجة المسبقة للبيانات:

خضعت الصور لعدة خطوات معالجة تهدف إلى تحسين جودة المدخلات وتوحيدها، وقد تم اعتماد ثلاث منهجيات مختلفة بحسب طبيعة النموذج المستخدم:

- **النموذج الهندسي:**

في هذا النموذج شملت المعالجة المسبقة إعادة تحجيم الإطارات إلى أبعاد موحدة، تدويرها لمحاكاة زوايا التقاط مختلفة، تم استخدام كاشف Haarلاستخراج الوجه[13] ،حيث يعتمد على خصائص بسيطة من المستطيلات السوداء والبيضاء تستخدم لتمييز الفروقات في شدة الإضاءة بين مناطق الوجه (مثل العينين، الأنف، والفم). تُحسب هذه الخصائص بكفاءة عالية باستخدام ما يعرف بالصورة التكاملية، ثم تدمج عبر خوارزمية AdaBoost في مصنّف قوي منظم على شكل Cascade of Classifiers لزيادة سرعة التنفيذ. بعد تحديد المستطيل المحيط بالوجه باستخدام كاشف Haar جرى استخراج النقاط المرجعية (landmarks) وحساب المسافات الإقليدية بين النقاط المختارة لتكوين متجه سمات هندسية ثابت الطول، مع استبعاد العينات غير الصالحة التي لم يُكشف فيها وجه أو لم تُستخرج منها معالم صحيحة.

- **النموذج البصري:**

في هذا النموذج لم يتم استخدام كاشف Haar أو استخراج السمات الهندسية بل تم الاعتماد على مكتبة RetinaFace لكشف الوجه بدقة أعلى، بعد اقتصاص الوجه، يتم إعادة تحجيمه إلى أبعاد ثابتة (224*224) ثم تطبيع قيم البكسلات وتحويلها إلى صيغ مناسبة لإدخالها في الشبكة

العصبونية. تم حفظ العينات الصالحة فقط في ملفات npz مع التصنيفات، بينما استُبعدت الصور التي لم يُكشف فيها الوجه.

- **النموذج الهجين:**

في هذا النموذج جرى المزج بين منهجيات حديثة وكلاسيكية، حيث استخدم RetinaFace لكشف الوجه بدقة عالية ثم تمرير الوجه بعد اقتصاصه إلى وحدة استخراج السمات التي تعتمد على dlib لاستخراج النقاط المرجعية (مع استخدام كاشفHaar كخطة بديلة عند فشل الكشف). بعد ذلك يتم حساب المسافات الإقليدية بين النقاط المختارة لتكوين متجه سمات هندسية ثابت الطول. وبذلك يجمع النموذج بين قوة الكشف الحديثة والتمثيل الهندسي المبسط للوجه.

4.5 تصميم النماذج التصنيفية:

تم تطوير ثلاث بنى معمارية مختلفة تختلف في نوعية المدخلات وطريقة المعالجة:

- **النموذج الهندسي (Geometric Feature-Based Model)**

يعتمد على متجه السمات العددية المستخرجة من النقاط المرجعية للوجه (15 بُعداً) باستخدام شبكة التفافية أحادية البُعد (1D-CNN) مع كتل ResidualSE ويتميز بقدرته على تفسير القرار بناءً على العلاقات الهندسية الواضحة.

- **النموذج البصري (Image-Based CNN Model)**

يعتمد على الصورة الاصلية للوجه دون معالجة هندسية باستخدام شبكة ثنائية البُعد (2D-CNN) مع كتل ResidualSE ويتيح التعلم المباشر من الأنماط البصرية الدقيقة.

- **النموذج الثالث: (Hybrid Model)**

يجمع بين المدخلين عبر مسارين CNN لمعالجة الصورة ومسار MLP للسمات الهندسية يتم دمجهما في طبقة تصنيف مشتركة (Fully Connected) لإنتاج القرار النهائي مما يوفر توازناً

بين قابلية التفسير والدقة والمرونة. يوضح الجدول 1 الفروقات الأساسية بين النماذج من حيث نوع المدخلات، البنية، قابلية التفسير، والمرونة، حيث يظهر النموذج الهندسي تفوقاً في الشفافية، بينما يحقق النموذج البصري دقة أعلى في اكتشاف الأنماط، ويقدم النموذج الهجين تكاملاً بين الاثنين.

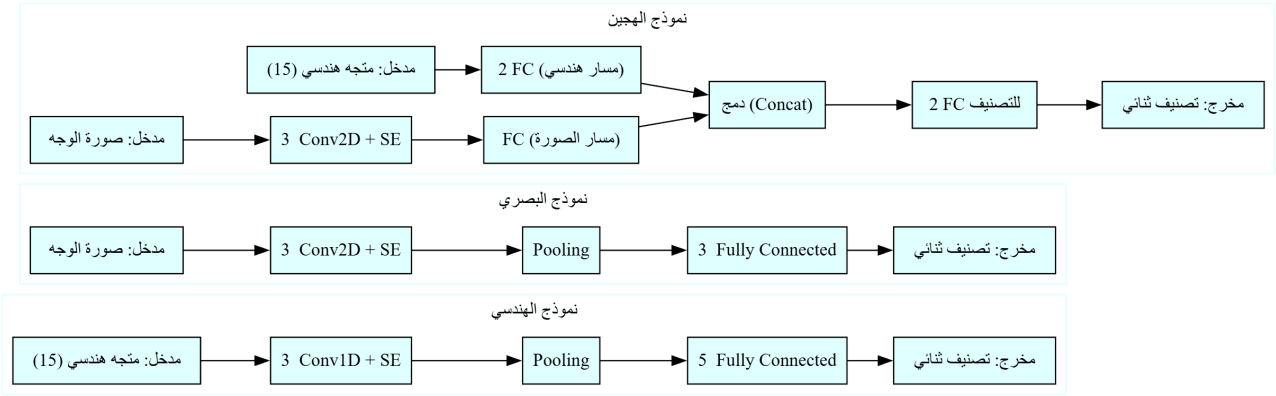
العنصر	النموذج الهندسي	النموذج البصري	النموذج الهجين
المدخلات	متجه هندسي (15 بُعداً)	صورة الوجه الأصلية	صورة الوجه+ متجه هندسي
نوع الشبكة	1D CNN+ ResidualSE	2D CNN+ ResidualSE	2D CNN+MLP
عدد الطبقات	3 Conv1D + 5 FC	3 Conv2D +3 FC	3Conv2D (مسار البصري) +2 FC (مسار الهندسي) +2 FC دمج + (للتصنيف)
حجم المرشحات	32,64,128	32,64,128	32,64,128
حجم النوافذ	3	3*3	3*3
حجم النموذج (ملف الوزن)	4.90 MB	13.69MB	7.52 MB
عدد المعاملات	1270480	3579312	1959152
زمن التنفيذ	الأسرع	الأبطأ	متوسط
قابلية التفسير	عالية	منخفضة	متوسطة–عالية

قابلية التطبيق	يتطلب معالجة مسبقة	مباشر على الصور	مرن وقابل للتكيف
المرونة والتكامل	محدود	محدود	عالي
الفرق الأساسي	يعتمد على العلاقات العددية بين النقاط المرجعية	يتعلم من الأنماط البصرية الدقيقة في الصور	يدمج السمات العددية والبصرية في قرار واحد

جدول 1 مقارنة النماذج الثلاث

الرسم المعماري للنماذج:

يوضح الشكل 2 البنية المعمارية لكل نموذج على حدة بدءاً من نوع المدخل مروراً بمراحل المعالجة وصولاً إلى الإخراج النهائي (تصنيف ثنائي: غش /ليس غش).



الشكل 2 الرسم المعماري للنماذج

4.6 التدريب والتقييم:

تم تدريب النماذج المقترحة باستخدام خوارزمية الانحدار العكسي (Backpropagation) مع خوارزمية التحسين Adam وقد تم ضبط المعلمات الأساسية على النحو التالي:

معدل التعلم (Learning Rate) =0.0001 لضمان استقرار عملية التدريب وتقليل احتمالية تخطي القيم المثلى.

حجم الدفعة (Batch Size) =64 لتحقيق توازن بين سرعة التدريب واستقرار التدرجات.

عدد الحقب (Epochs)=450 مع استخدام آلية التوقف مبكر Early Stopping لتجنب الإفراط في التدريب.

أما بالنسبة لدالة خسارة فقد تم اعتماد تقاطع الانتروبي (Cross-Entropy Loss) مع أوزان للفئات (Class Weights) لمعالجة مشكلة عدم التوازن بين العينات. ورغم أن هذه الدالة معيارية في مهام التصنيف متعدد الفئات إلا أنها تُستخدم أيضاً في حالة التصنيف الثنائي عندما يكون للنموذج مخرجان (غش/ ليس غش). في هذه الحالة تطبق داخلياً دالة Softmax لتحويل المخرجات إلى احتمالات ثم تحسب الخسارة. هذا الاختيار لا يعني أن المهمة متعددة الفئات بل هو متوافق تماماً مع تصميم النموذج الحالي الذي يعتمد على مخرجين ويُعد مكافئاً رياضياً لاستخدام Sigmoid مع Binary Cross-Entropy مع ميزة إضافية في سهولة تطبيق أوزان للفئات وتوسعة النموذج لاحقاً إذا تطلب الامر.

مؤشرات التقييم:

تم تقييم النموذج بعد كل حقبة تدريبية بالاعتماد على مصفوفة الارتباك (Confusion Matrix) وهي مصفوفة ثنائية الأبعاد من الشكل 2*2 في حالة التصنيف الثنائية، وتُستخدم مصفوفة الارتباك لمقارنة التنبؤات بالقيم الحقيقية.

الحقيقة/التنبؤ	القيمة الحقيقية=1	القيمة الحقيقية=0
تنبؤ=1	True Positive (TP)	False Positive (FP)
تنبؤ=0	False Negative (FN)	True Negative (TN)

حيث تم حساب مؤشرات الأداء التالية من مصفوفة الارتباك:

الدقة (Accuracy): هي نسبة العينات المصنفة بشكل صحيح إلى اجمالي العينات:

$$\text{Accuracy} = \frac{\text{عدد التنبؤات الصحيحة}}{\text{إجمالي عدد العينات}} = \frac{TP + TN}{TP + TN + FP + FN}$$

مقياس (F1 Score): وهو المتوسط التوافقي بين الدقة والاسترجاع، ويُستخدم لتقييم الأداء في حالة عدم توازن الفئات:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

$$\text{Recall} = \frac{TP}{TP + FN} \qquad \text{Precision} = \frac{TP}{TP + FP}$$

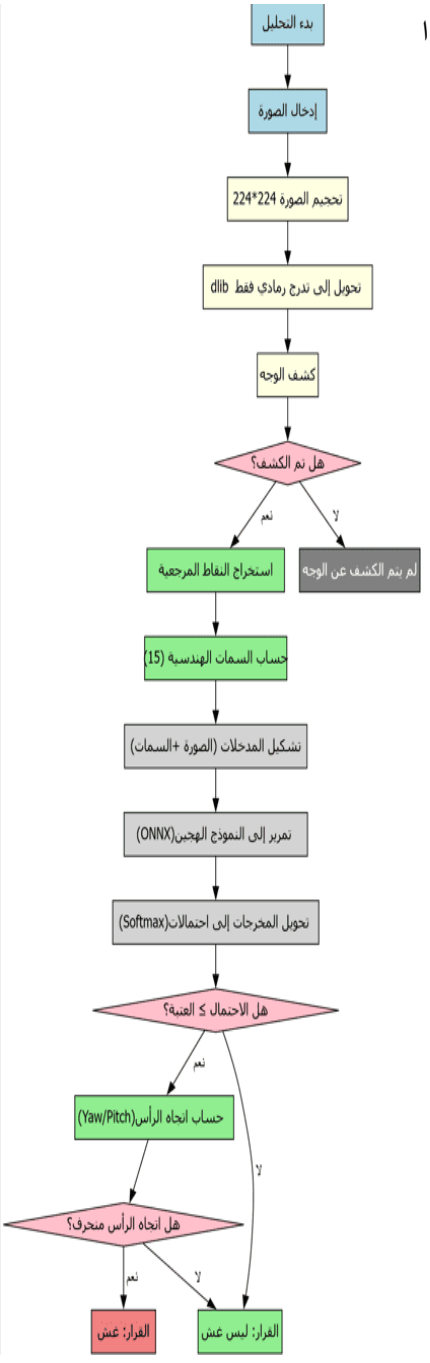
لتفادي الإفراط في التعلّم، تم تطبيق تقنية التوقف المبكر (Early Stopping) حيث يُنهى التدريب تلقائياً إذا لم يتحسن الأداء خلال عدد محدد من الحقبات المتتالية، كما تمت مراقبة ديناميكية التعلم عبر رصد قيم انحدار التدرج (GradientNorms) ولضمان تعميم النتائج وعدم اعتمادها على تقسيم معين للبيانات، تم تطبيق التحقق المتقاطع (Cross-Validation) باستخدام طريقة K-Fold بعدد الطيات=10.وقد جرى التقسيم على مستوى الجلسة/الفيديو لكل طالب بحيث يُخصص كل فيديو بالكامل إما للتدريب أو للاختبار مما يضمن استقلالية الأشخاص والجلسات ويمنع أي تداخل أو تسرب للمعلومات بين المجموعتين. في كل طية، يتم تدريب النموذج على 90% من البيانات وتقييمه على 10% المتبقية، ثم يُحسب متوسط دقة الطيات. بعد اكتمال التدريب تم حفظ أداء النموذج في ملف بصيغة CSV، وتوثيق أفضل حقبة تدريبية بناءً على أقل خسارة اختبار وأعلى دقة اختبار. ثم تحويلها إلى صيغة معيارية ONNX لتسهيل النشر والتكامل مع بيئات مختلفة ودعم التشغيل على محركات استدلال متعددة بكفاءة عالية.

4.7 الاستدلال المحلي:

يعتمد الاستدلال المحلي على معالجة الصورة المدخلة واستخراج مجموعة من السمات الهندسية للوجه باستخدام نموذج كشف المعالم الوجهية القائم على الانحدار.

تبدأ العملية بتحجيم الصورة إلى أبعاد قياسية (224*224) وتحويلها إلى صيغة تدرج رمادي بهدف تبسيط عملية الكشف عن المعالم وتقليل الضوضاء البصرية وهو ما يتناسب مع نموذج dlib المستخدم لاستخراج النقاط المرجعية. أما النماذج الأخرى مثل RetinaFace و YOLOv9 فقد استُخدمت بصيغتها الأصلية مع الصور الملونة (RGB) دون تعديل وذلك للحفاظ على دقتها في الكشف عن الوجوه والأجسام حيث أن هذه النماذج

تعتمد على المعلومات اللونية في إعداداتها الافتراضية. بعد تحديد موقع ا



الشكل 3 مخطط Flowchart للاستدلال

النقاط المرجعية الخاصة بالعينين والفم ويُحسب منها مجموعة من العلاقات الهندسية مثل المسافة بين مركزي العينين، عرض وارتفاع الفم ومتوسط المسافات داخل كل عين بالإضافة إلى مسافات إضافية بين العينين والفم. ينتج عن هذه العمليات متجه سمات هندسية مكون من 15 قيمة يستخدم مع الصورة الاصلية كمدخلات للنموذج الهجين بصيغة ONNX. تُمرر هذه البيانات إلى النموذج بعد إعادة تشكيلها، وتحول المخرجات إلى توزيع احتمالي باستخدام دالة Softmax، مما يتيح تفسيراً احصائياً للنتائج. وفي مرحلة الاستدلال لا يقتصر القرار النهائي على مخرجات النموذج فقط بل يُضاف شرط يعتمد على اتجاه الرأس المستخرج من النقاط المرجعية (Yaw, Pitch)فإذا أظهر التحليل أن الرأس منحرف بشكل واضح (يمين، يسار، أعلى، أسفل) تصنّف الحالة ضمن فئة "غش" حتى لو كان الاحتمال الإحصائي أقل من العتبة المحددة. أما في الحالات الطبيعية فيُتخذ القرار بناءً على مقارنة الاحتمال الناتج مع العتبة، حيث تُصنّف الصورة ضمن الفئة غش أو ليس غش وفقاً لذلك. الشكل 3 يوضح مراحل الاستدلال المحلي (inference) بدءاً من ادخال الصور وصولاً إلى القرار النهائي غش/ليس غش.

4.8 دمج كشف العناصر وتحليل النشاط الفموي:

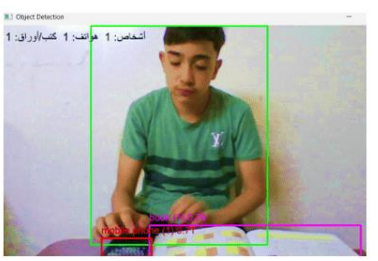
- ***كشف الأغراض:***

في هذه المرحلة، يُنفذ النظام عملية الكشف البصري على الإطارات الواردة من الكاميرا، حيث تُحلل الصورة وتُستخرج الكائنات ذات الصلة وفقاً لعتبات ثقة محددة لكل صنف إذ تم اعتماد 0.5 للأشخاص، 0.3 للهواتف المحمولة، 0.2 للكتب والأوراق وذلك بعد تجارب تحقق أولية لضبط التوازن بين الدقة والاستدعاء لكل فئة. بعد تصنيف الكائنات المكتشفة وتحديد مواقعها داخل

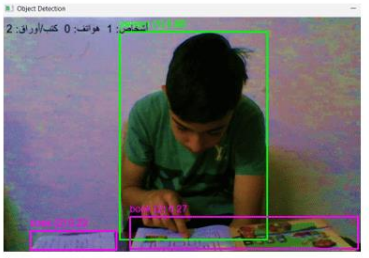
الإطار، تعرض النتائج بصرياً عبر مربعات محيطة وتسميات توضيحية ملوّنة، كما يُضاف ملخص عددي للكائنات في أعلى الشاشة مع مراعاة الترتيب الصحيح للنص العربي وشكل الحروف لضمان وضوح العرض ودقته اللغوية.

تم الاعتماد في هذه المرحلة على تقنية التعلم بالنقل (Transfer Learning) من خلال إعادة تدريب نموذج YOLOv9 المدرب مسبقاً على مجموعة COCO بشكل مخصص لفئة الكتب/الأوراق باستخدام قاعدة بيانات خاصة بالبحث جرى تجميعها وتوسيمها يدوياً عبر أداة LabelImg بينما فئات الأشخاص والهواتف المحمولة تم الاعتماد فيها مباشرة على النموذج الأصلي المدرب مسبقاً على مجموعةCOCO دون إعادة تدريب إضافي. وقد تم تقسيم قاعدة البيانات بنسبة 70% للتدريب، 20% للتحقق، 10% للاختبار لضمان تقييم مستقل للنموذج على بيانات غير مستخدمة في التدريب أو التحقق. هذا الإعداد ساهم في تحسين قدرة النموذج على التعرف بدقة أكبر على العناصر المرتبطة بسلوكيات الغش في الامتحانات [14]

الاشكال 4-6 أمثلة على نتائج الكشف البصري، حيث تم التعرف على الأشخاص، الهاتف، والكتاب ضمن الإطار، مع عرض الثقة والتسميات باللغة العربية.



الشكل 4 كشف شخص وكتاب وهاتف

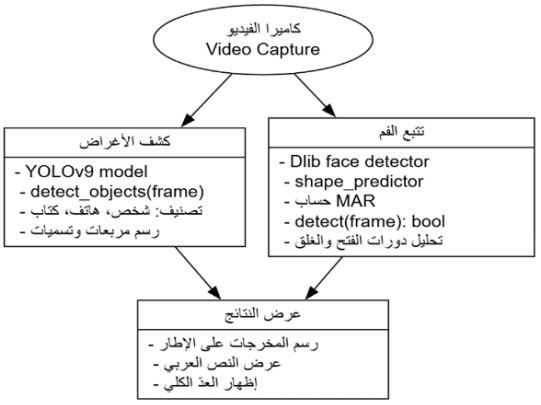


الشكل 5 كشف شخص وكتب/أوراق عدد2

الشكل 6 كشف شخصين وكتاب

- **تحليل حركة الفم:**

في هذه المرحلة، يُعتمد على تحليل الهندسي لمعالم الوجه بهدف الكشف عن نشاط فموي الذي قد يرتبط بالكلام أو السلوكيات غير المرغوبة في سياقات المراقبة. يتم اولاً استخراج وجه المستخدم من الإطار المرئي، ثم تحديد النقاط المرجعية المحيطة بالفم، ليُحسب مؤشر نسبة فتح الفم (MAR) باستخدام معادلة رياضية تعتمد على المسافات الاقليدية بين النقاط الرأسية والجانبية للفم.



الشكل7 دمج كشف الاغراض مع تحليل حركة الفم

يُعرف هذا المؤشر وفق العلاقة:

$$MAR = \frac{A + B + C}{2D}$$

حيث:

- A المسافة بين النقطة 61 و67، وتمثل الجزء الداخلي العلوي والسفلي للشفة اليمنى.
- B المسافة بين النقطة 62 و 66، وتمثل الجزء الأوسط الداخلي للشفتين.
- C المسافة بين النقطة 63 و 65، وتمثل الجزء الداخلي الايسر للشفتين.
- D عرض الفم من الزاوية اليمنى إلى الزاوية اليسرى.

تعكس القيم A,B,C مدى فتح الفم عمودياً، بينما تمثل القيمة D اتساع الفم افقياً. وبناءً على هذه النسبة يتم تحديد حالة الفم (مفتوح أو مغلق) وفق عتبة محددة مسبقاً قيمتها 0.25 فإذا تجاوزت قيمة MAR هذه العتبة يُعتبر الفم في حالة فتح، وإذا كانت أقل منها يُعتبر الفم مغلقاً. ولم يقتصر التحليل على قيمة MAR وحدها، بل تم الاعتماد على النمط الزمني لحركة الفم مما أتاح التمييز بين الحالات المختلفة:

- فم مغلق: عندما تكون قيمة $MAR \leq 0.25$ ولم تُسجل أي دورات فتح/إغلاق.

- فم مفتوح طبيعي: عندما تتجاوز قيمة MAR العتبة لفترة قصيرة أو بشكل غير متكرر.
- تحدث: عندما يُرصد تكرار دورات فتح وإغلاق الفم بشكل متتابع خلال فترة زمنية قصيرة.

يتيح هذا النهج التمييز بين السلوكيات الطبيعية (مثل التثاؤب أو فتح الفم بشكل طبيعي دون تحدث) والسلوكيات المرتبطة بالكلام مما يعزز دقة النظام ويمنع تفسير الحركات الطبيعية أو غير الإرادية كحالات غش.

الشكل 8حالة الفم مفتوح لفترة زمنية قصيرة وبشكل غير مكرر

الشكل 9 يوضح حالة التحدث حيث تم رصد دورات فتح وإغلاق بشكل مكرر

الشكل 10حالة الفم مغلق قيمة مؤشر MAR≤0.25

توضح الأشكال (8–10) تحليل حركة الفم باستخدام مؤشر MAR حيث تُظهر الصورة الأولى قيمة MAR=0.29 لفترة زمنية قصيرة ودون رصد تكرار دورات فتح/إغلاق مما يدل على أن الفم مفتوح بشكل طبيعي دون تحدث. بينما تُظهر الصورة الثانية قيمة المؤشر MAR=0.28 مع رصد ثلاث دورات فتح/إغلاق الفم متتابعة وهو ما يدل على حالة التحدث. أما الصورة الثالثة ظهر مؤشر MAR=0.03وهو أقل من العتبة المحددة مما يعكس حالة الفم المغلق.

4.9 نشر النظام عبر واجهة ويب تفاعلية:

تعتمد المنهجية في هذا النظام على تصميم خط معالجة متكامل لمراقبة الامتحانات عبر الفيديو وتحليل السلوك البصري بهدف كشف حالات الغش. يبدأ النظام بالتقاط الإطارات من كاميرا الويب وإرسالها إلى وحدة المعالجة، حيث يخضع كل إطار لسلسلة من العمليات تشمل: تصغير حجم الإطار، كشف العناصر (أشخاص، هواتف، كتب)، وتحليل حركة الفم. بعد ذل ك يتم الاستنتاج السلوكي باستخدام نموذج التصنيف.

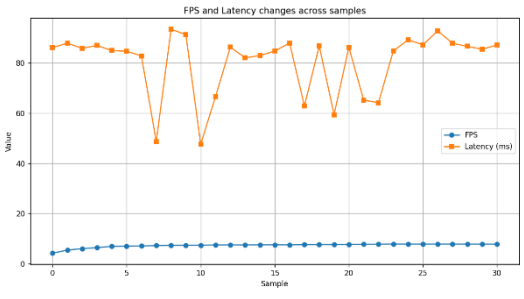
تشمل معايير القرار النهائي: وجود أكثر من شخص في الإطار، ظهور هاتف أو كتاب، التحدث، ونتيجة النموذج السلوكي. تحفظ النتائج في ملف بصيغة CSV يحتوي على المعلومات التالية: توقيت الإطار– رقم الإطار– عدد العناصر المكتشفة – معدل الإطارات في الثانية– وزمن الاستجابة.

يوفر النظام واجهة ويب تفاعلية تتيح تسجيل الدخول، بدء الامتحان، عرض الفيديو المباشر، وإنهاء الامتحان، مع بث فوري لنتائج المعالجة، كما يعتمد على المعالجة المتزامنة (multithreading) لضمان الأداء في الزمن الحقيقي. ويُظهر هذا التصميم كيف يمكن لتقنيات الرؤية الحاسوبية والتعلم العميق ان تتكامل لتقديم حلول عملية قابلة للتطبيق في بيئات تعليمية، مع تعزيز مرونة النشر عبر تحويل النموذج إلى صيغة ONNX وإتاحة إمكانيات التوسع والتكامل مع أنظمة متعددة.

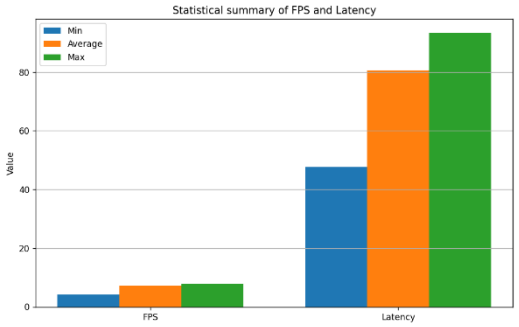
5. النتائج ومناقشتها:

في هذه الفقرة يتم عرض نتائج الاختبارات العملية التي أجريت على النظام المقترح بعد الانتهاء من تطويره وتنفيذه وفقاً للمنهجية الموضحة سابقاً. تهدف هذه النتائج إلى تقييم أداء النظام في بيئة واقعية، من خلال تحليل قدرته على كشف الكائنات المرتبطة بسلوكيات الغش الأكاديمي، ورصد المؤشرات السلوكية مثل التحدث أثناء الامتحان وذلك باستخدام تقنيات الرؤية الحاسوبية والنماذج الذكية المدربة مسبقاً.

تم قياس معدل الإطارات في الثانية (FPS) وزمن الاستجابة (Latency) أثناء التشغيل الفعلي للنظام حيث أظهرت النتائج أن النظام يعمل بزمن شبه لحظي محققاً معدل إطارات يتراوح بين 5–8 إطار/ثانية بمتوسط يقارب 8 إطار/ثانية بينما تراوح زمن الاستجابة بين 47–93 مللي ثانية بمتوسط يقارب 80 مللي ثانية. كما هو موضح في الشكلين (11–12) تم تلخيص القيم الدنيا والعظمى والمتوسطة لكل من FPS و Latency.



الشكل 11 مخطط خطي يوضح تغير قيم معدل الإطارات (FPS) وزمن الاستجابة (Latency) عبر العينات أثناء التنفيذ الفعلي للنظام



الشكل 12مخطط أعمدة يوضح القيم الإحصائية الدنيا والعظمى والمتوسطة لكل من معدل الإطارات (FPS) وزمن الاستجابة (Latency)

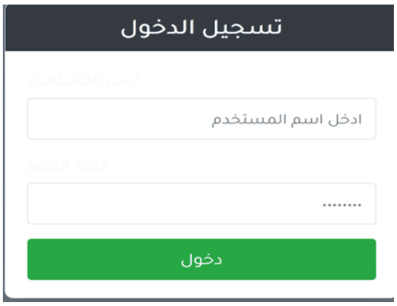
هذه النتائج تؤكد قدرة النظام على العمل في الزمن الحقيقي ومتابعة سلوك الطالب بشكل مستمر أثناء الامتحان.

5.1 صفحات الويب في النظام:

تم تطوير واجهة ويب تفاعلية لتمكين المستخدم من التفاعل المباشر مع النظام، بدءاً من تسجيل الدخول وحتى إنهاء الجلسة. توضح الصور المرفقة المراحل المختلفة لتشغيل النظام وتشمل:

أ- صفحة تسجيل الدخول login/:

تتيح للمستخدمين الدخول إلى النظام باستخدام بيانات اعتماد بسيطة (اسم المستخدم وكلمة مرور).



الشكل 13 صفحة تسجيل الدخول

ب- الصفحة الرئيسية /:

تظهر بعد تسجيل الدخول وتوفّر خيارات لبدء الامتحان أو تسجيل الخروج.

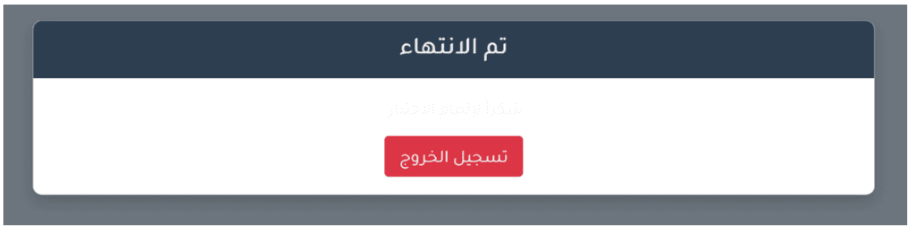
الشكل 14 الصفحة الرئيسية لنظام كشف الغش

ت- صفحة الامتحان exam/:

يتم تفعيل الكاميرا وتحليل الإطارات باستخدام YOLOو MARو Inference مع عرض مباشر للفيديو وتوصيف الحالة (غش / ليس غش) باللغة العربية.

ث- صفحة إنهاء الامتحان end_exam/:

تستخدم لإيقاف جميع العمليات (الكاميرا، المعالجة، الكتابة) ومسح الجلسة وتوجيه المستخدم إلى صفحة النهاية.



الشكل 15 صفحة إنهاء الامتحان

ج- صفحة تسجيل الخروج logout/:

تستخدم لمسح الجلسة يدوياً بحيث تعيد المستخدم إلى صفحة تسجيل الدخول وتوقف النظام. وحفظ جميع النتائج في ملف inference_records.csv حيث يتضمن عدد الأشخاص، وجود الهاتف أو الكتاب، حالة الفم، نتيجة النموذج، والتصنيف النهائي.

أظهرت الصور المرفقة حالات الغش في الزمن الحقيقي، حيث يتم تصنيف الإطارات تلقائياً، وتُعرض الحالة بشكل مرئي واضح، مما يعكس فعالية النظام في بيئة تشغيل واقعية.

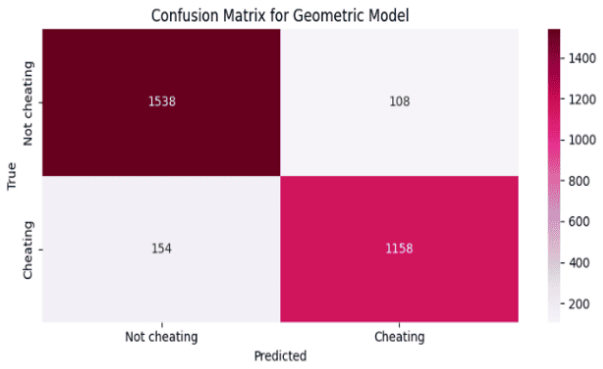
5.2 التقييم:

في هذه المرحلة يتم تحليل الأداء وذلك من خلال المؤشرات كمية ونوعية حيث يتم تقييم كل نموذج على مجموعة الاختبار بعد الوصول إلى أفضل أداء خلال التدريب وتم توثيق النتائج النهائية لكل نموذج على حدا وحفظها في ملفات بصيغة CSV.

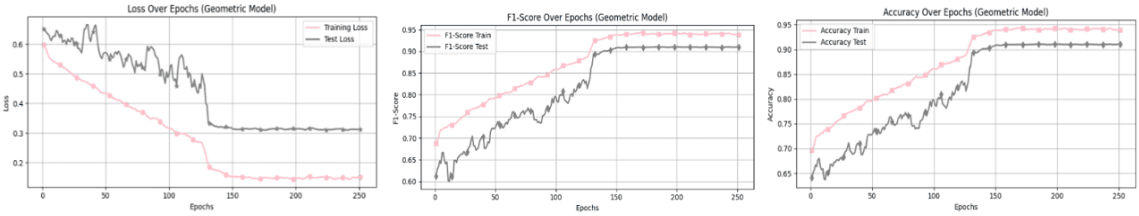
- **تقييم أداء النموذج الهندسي:**

أظهر النموذج الهندسي أداءً جيداً نسبياً في التمييز بين حالات الغش والحالات الطبيعية حيث بلغت دقة الاختبار 91.1% ومعامل F1–score نحو 89.9% مما يعكس توازناً مقبولاً بين الدقة النوعية والاسترجاع.

كما توضح مصفوفة الارتباك الشكل 16 توزيع التنبؤات الصحيحة والخاطئة، حيث حقق النموذج استرجاعاً (Recall) بنسبة 88.3% مع معدل الإنذارات كاذب منخفض بلغ 6.5%.



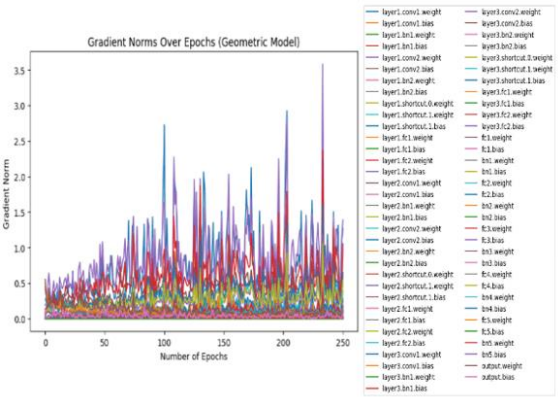
الشكل 16 مصفوفة الارتباك (النموذج الهندسي)



الشكل 17 منحنيات الأداء للنموذ

وتبين منحنيات الأداء (الشكل 17) الخاصة بالدقة والخسارة و F1 استقراراً تدريجياً بعد منتصف التدريب حيث استقرت قيم الدقة حول 0.91 على بيانات الاختبار مع انخفاض ملحوظ في الخسارة التدريبية.

أما منحنيات التدرجات الشكل18 قد اظهر استقراراً واضحاً عبر الطبقات وغياباً لمشكلات التلاشي أو

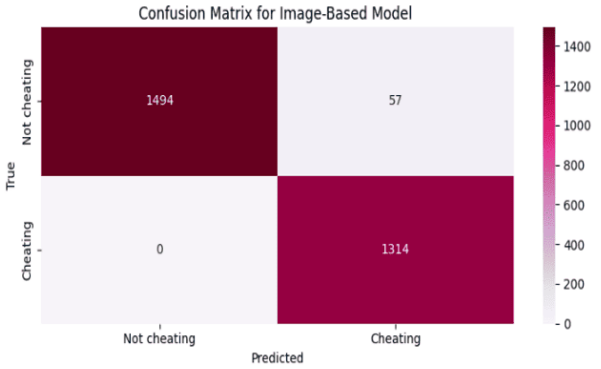


الشكل 18 منحنيات التدرجات النموذج الهندسي

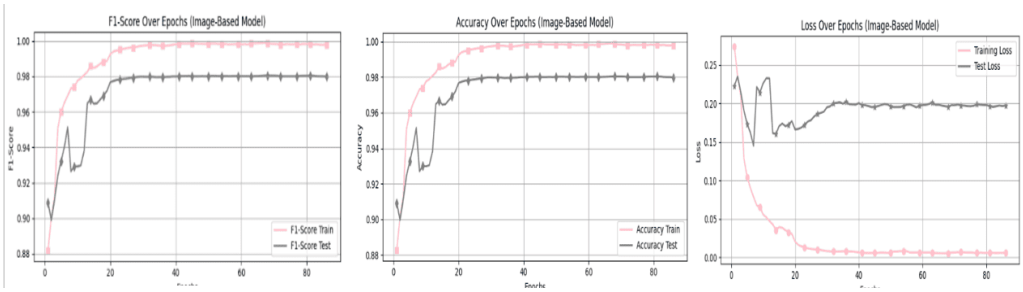
الانفجار مما يعزز الثقة في البنية الداخلية للنموذج. ورغم كفاءته التشغيلية العالية، الا أن النموذج أقل دقة في التعامل مع الحالات المعقدة.

- **تقييم النموذج البصري:**

حقق النموذج المعتمد على الصور اداءً متفوقاً مقارنة بالنموذج الهندسي، حيث بلغت دقته الكلية 98% ومعامل F1 نحو 97.9% مع استرجاع مثالي Recall=100% ودقة نوعية Precision=95.8% وتوضح مصفوفة الارتباك (الشكل 19) أن النموذج لم يُغفل أي خالة غش (FN=0) بينما بقيت الإنذارات الكاذبة منخفضة (FP=57) مما يعزز موثوقيته في البيئات الحساسة.



الشكل 19 مصفوفة الارتباك (النموذج البصري)



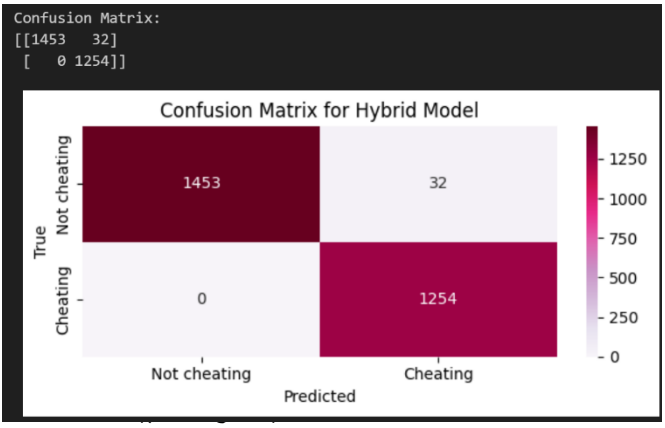
الشكل 20 منحنيات الأداء (النموذج البصري)

كما تبين منحنيات الأداء (الشكل 20) أن النموذج يتعلم بسرعة ويستقر تدريجياً، حيث ارتفعت قيم الدقة و F1 بشكل واضح مع انخفاض الخسارة دون مؤشرات على فرط التعلّم. ويظهر تحليل التدرجات الشكل 21 ديناميكية تعلم متوازنة مع بعض التذبذبات في طبقات BatchNorm لكنها لم تؤثر على الاستقرار العام للنموذج.

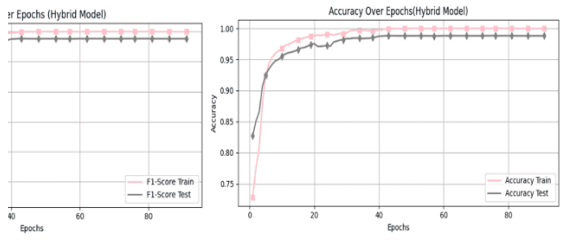
هذا الأداء القوي يجعله خياراً ممتازاً في التطبيقات التي تتطلب دقة عالية رغم حاجته إلى موارد حسابية كبيرة.

- **تقييم النموذج الهجين:**

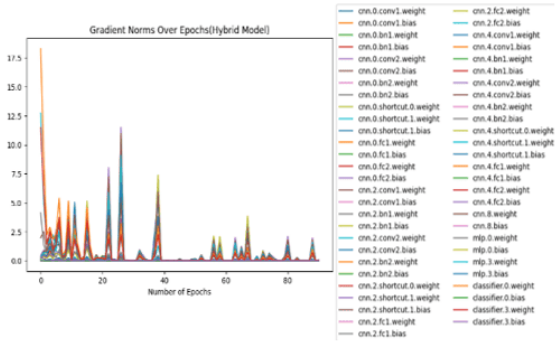
اظهر النموذج الهجين أفضل أداء بين النماذج الثلاثة حيث بلغت دقته 98.8% ومعامل F1 نحو 98.7% مع استرجاع مثالي 100% ودقة نوعية بلغت 97.5% وتوضح مصفوفة الارتباك الشكل 22 أن النموذج لم يغفل أي حالة غش (FN=0) مع عدد إنذارات كاذبة منخفض (FP=32) مما يعكس قدرة فائقة في التمييز بين الفئتين.



الشكل 22 مصفوفة الارتباك (النموذج الهجين)



الشكل 23 منحنيات الأد



الشكل 24 تحليل التدرجات (النموذج الهجين)

تبين منحنيات الأداء (الشكل 23) أن النموذج تعلّم بسرعة واستقر مبكراً حيث انخفضت الخسارة بشكل ملحوظ خلال أولى الحقبات التدريبية، بينما ارتفعت قيم الدقة و F1 إلى مستويات شبه مثالية.

ويؤكد تحليل التدرجات الشكل 24 ان عملية التعلم موزعة بشكل متوازن عبر مكونات النموذج المختلفة دون علامات على الاشي أو انفجار التدرجات. هذا التوازن بين دقة عالية وكفاءة تشغيلية يجعل النموذج الهجين الخيار الأمثل للتطبيقات التي تتطلب استجابة موثوقة وسرعة تنفيذ مع الحفاظ على أعلى مستويات الأداء.

- **المقارنة الكمية بين النماذج:**

النموذج	نوع المدخلات	الدقة Accuracy	الاسترجاع Recall	الدقة النوعية Precision	F1 Score
النموذج الهندسي	سمات هندسية	%91.1	%88.3	%91.5	%89.9
النموذج البصري	صور الوجه	%98	%100	%95.8	%97.9
النموذج الهجين	صورة+ سمات	%98.8	%100	%97.5	%98.7

جدول 2 المقارنة الكمية بين النماذج

أظهرت النماذج الثلاثة (الهندسي، البصري، والهجين) اداءً متبايناً يعكس اختلاف بنيتها ومدى تعقيدها. النموذج الهندسي تميز بكفاءة تشغيلية عالية وسرعة في التنفيذ مما يجعله مناسباً للأنظمة محدودة الموارد او التطبيقات التي تتطلب استجابة سريعة دون الحاجة إلى معالجة بصرية معقدة. ورغم أن دقته بلغت 91.1% إلا أنه أظهر بعض القصور في اكتشاف الحالات المعقدة. في المقابل النموذج المعتمد على الصور حقق اداءً قوياً حيث بلغ معامل F1 نحو 97.9% مع استرجاع مثالي 100% وغياب كامل للأخطاء السلبية (FN=0) مما يجعله خياراً ممتازاً في البيئات التي تتطلب دقة عالية في كشف الغش خاصة عندما تكون الصور متاحة بجودة مناسبة. ومع ذلك يتطلب موارد حسابية أكبر وزمن تنفيذ أطول. أما النموذج الهجين جمع بين مزايا النموذجين السابقين حيث حقق أعلى دقة كلية بلغت 98.8% واسترجاعاً مثالياً وغياب كامل للأخطاء السلبية (FN=0) إضافة إلى توازن واضح في توزيع التدرجات عبر مكوناته المختلفة. كما أظهر النموذج استقراراً تدريجياً في الأداء منذ الحقب الأولى مما يدل على فعالية التصميم البنيوي في دمج السمات البصرية والهندسية. ويجعل هذا التوازن النموذج الهجين الخيار الأمثل للتطبيقات التي تتطلب دقة عالية واستقراراً سريعاً دون التضحية بالكفاءة التشغيلية.

وبناءً على هذه النتائج يمكن اختيار النموذج الأنسب وفقاً لمتطلبات النظام وقيود الموارد مع مراعاة طبيعة البيانات المتاحة وأهمية تقليل الأخطاء في التصنيف لضمان موثوقية الأداء.

5.3 المقارنة مع الدراسات الحديثة:

بعد عرض نتائج التقييم يقدّم هذا القسم مقارنة موجزة بين النظام المقترح وبعض الدراسات المماثلة في المجال وذلك لتوضيح الفروقات الأساسية في الأداء والبنية التنفيذية.

يوضح الجدول 3 أن النظام المقترح يتفوق من حيث دقة التصنيف، التكامل بين الوحدات، ودعم التشغيل في الزمن الحقيقي إضافة إلى مرونته العالية في التوسع مقارنة بالدراسات السابقة.

الوظيفة	دراسة [3]	دراسة [4]	دراسة [8]	دراسة [9] مقارنة نظرية	النظام المقترح في هذا البحث
دقة التصنيف	%77.9	F1 =0.94	%96	93- %98	%98.8 الهجين %98 البصري %91.1 الهندسي
Precision	%81.4	-	_	_	%97.5 الهجين %95.8 البصري %91.5 الهندسي
Recall	%76.51	-	_	_	%100 الهجين+ البصري %88.3 الهندسي
الأخطاء السلبية (FN)	غير موضحة	غير موضحة	غير موضحة	غير موضحة	FN=0 غياب كامل للأخطاء السلبية في كل من النموذج البصري والهجين
نوع المدخلات	نقاط مرجعية للوجه باستخدام Mediapipe	صورة الوجه+ حركة الرأس+ اتجاه النظر	مدخلات متعددة (DL+ML)	صور+ نماذج CNN/+ YOLO	صورة الوجه أو السمات الهندسية أو كليهما حسب النموذج المختار
التحقق من الهوية	لا يوجد تحقق من الهوية	نعم باستخدام FaceNet	غير مذكور	غير مذكور	تحقيق مبدئي عبر بيانات اعتماد ثابتة

	النظام يبدأ من تحليل الفيديو مباشرة				(اسم المستخدم، كلمة المرور)
الوحدات الداعمة	–	تتبع الرأس والنظر	–	–	كشف الأغراض + تتبع حركة الفم + تتبع اتجاه الرأس
التكامل	محدود (فقط نقاط الوجه)	متوسط (عدة خوارزميات متفرقة)	جزئي	نظري	عالي جداً
دعم التزامن بين المهام	لا	لا	غير مدعوم بسبب التعقيد الحسابي وصعوبة التطبيق في الزمن الحقيقي	غير مذكور (مقارنة نظرية فقط)	نعم عبر فصل التقاط الإطارات، التحليل، والعرض

بنية التنفيذ	خطية بدون فصل بين المهام	غير مذكور	معقدة (دمج تقنيات متعددة دون إطار موحد)	غير مذكور (تحليل نظري للأداء)	متعددة المسارات Multi-threaded
قابلية النشر في الزمن الحقيقي	مدعوم (تم اختبارها على 20 مستخدماً افتراضي)	غير موضحة	غير مدعوم	غير موضح	مدعومة عبر ONNX وFlask مع معدل إطارات تقريباً 8 إطار/ثانية وزمن استجابة يقارب 80 مللي ثانية
مرونة التوسع والتخصيص	محدودة	متوسطة	ضعيفة	متوسطة	عالية مع قابلية إضافة وحدات جديدة بسهولة

جدول 3 جدول المقارنة بين الدراسات السابقة والنظام المقترح

تظهر المقارنة بين الدراسات السابقة والنظام المقترح ان التطور الحقيقي في مجال كشف الغش في الامتحانات الإلكترونية لا يتحقق عبر الاعتماد على نماذج بسيطة او خوارزميات متفرقة بل من خلال تصميم متكامل يجمع بين الرؤية الحاسوبية والتحليل السلوكي ضمن بنية تنفيذية متعددة المسارات.

وقد انعكس هذا التكامل في تفوق النموذج الهجين من حيث دقة التصنيف، والقدرة على الاسترجاع، إضافة إلى مرونته العالية في التوسع وامكانية دمج وحدات جديدة بسهولة. وبذلك يقدّم النظام

المقترح خطوة عملية نحو بيئة امتحانية أكثر عدلاً وموثوقية مع جاهزية تشغيلية تميزه عن النماذج السابقة.

6. التوصيات والخاتمة:

6.1 التوصيات:

استناداً إلى نتائج التقييم والمقارنة، يمكن تقديم التوصيات التالية لتطوير أنظمة كشف الغش في الامتحانات الإلكترونية:

- **تعزيز التكامل بين النماذج:** يُوصى باعتماد بنى هجينة تجمع بين السمات الهندسية والبصرية والسلوكية لما لها من أثر واضح في تحسين دقة التصنيف وتقليل الأخطاء السلبية.
- **دمج وحدات داعمة ذكية:** مثل كشف الأغراض باستخدام نماذج YOLO الحديثة وتتبع حركة الفم باستخدام مؤشرات هندسية مثل MARبالإضافة إلى حساب وتتبع اتجاه الرأس مما يتيح رصد أنماط الغش غير المباشرة بشكل أكثر دقة وموثوقية.
- **اعتماد بنية تنفيذية متعددة المسالك (Multi-threaded):** لضمان كفاءة زمنية واستجابة لحظية خاصة في البيئات متعددة المستخدمين او ذات الحمل العالي.
- **التركيز على تقليل الإنذارات الكاذبة (False Positives):** من خلال تحسين خوارزميات التصنيف وتوسيع قاعدة البيانات التدريبية لتشمل سيناريوهات متنوعة.
- **توسيع نطاق التقييم:** ليشمل اختبارات في بيئات واقعية متعددة بما يضمن التحقق من فعالية النظام في ظروف التشغيل المختلفة.
- **توفير واجهات تفاعلية مرنة:** تتيح للمراقب البشري التدخل عند الحاجة دون الاعتماد الكامل على القرار الالي مما يعزز من موثوقية النظام.
- **تصميم قابل للتوسعة والتحديث:** يُفضل اعتماد بنية برمجية مرنة تسمح بدمج نماذج أحدث مستقبلاً (مثل إصدارات YOLO المستقبلية او مؤشرات سلوكية إضافية) مما يضمن استدامة النظام وتطوره دون الحاجة لإعادة بناءه من الصفر.

6.2 الخاتمة:

في ختام هذا البحث تم تطوير نظام ذكي متكامل لكشف حالات الغش في الامتحانات الإلكترونية من خلال الجمع بين تقنيات الرؤية الحاسوبية والتحليل السلوكي ضمن بنية تنفيذية متعددة المسالك. لتحقيق ذلك تم تصميم وتقييم ثلاثة نماذج تصنيفية (هندسي، بصري، هجين) إلى جانب دمج وحدات داعمة لكشف الأغراض وتتبع التحدث. وقد أظهرت النتائج أن النموذج الهجين المقترح يحقق أداء تصنيفي قوي مع استرجاع مثالي ودقة نوعية مرتفعة مما يعكس قدرته على التمييز بين السلوكيات الطبيعية والغش في بيئات الامتحانات الرقمية. كما اثبتت البنية متعددة المسالك كفاءتها في تقليل زمن الاستجابة مما يجعل النظام مناسباً للتطبيق العملي حيث بلغ متوسط معدل الإطارات حوالي 8 إطار/ثانية ومتوسط زمن الاستجابة يقارب 80 مللي ثانية وهو ما يجعل النظام قادراً على العمل في الزمن الحقيقي حتى في البيئات التعليمية ذات الحمل التشغيلي المرتفع.

يمثل هذا النظام اساساً متقدماً لتطوير أدوات مراقبة ذكية أكثر شمولاً ومرونة، قادرة على رصد أنماط الغش المباشرة وغير المباشرة. ويمكن البناء على هذه النتائج مستقبلاً لتوسيع نطاق التطبيق نحو بيئات تعليمية أكثر تعقيداً مع دمج تقنيات إضافية مثل تحليل الصوت، تتبع العين، والتعلم التكيفي بما يعزز من موثوقية وكفاءة أنظمة كشف الغش في العصر الرقمي.

7. المراجع:

[1] MAHLANGU, V.P. 2018–The Good, the Bad, and the Ugly of Distance Learning. IntechOpen.

[2] KADTHIM, R.K. & ALI, Z.H. 2022– Survey: Cheating Detection in Online Exams. International Journal of Engineering Research and Advanced Technology (IJERAT), Vol. 8, No. 1, DOI: 10.31695/IJERAT.2022.8.1.1

[3] NGO, D.A., NGUYEN, T.D., DANG, T.L.C., LE, H.H., HO, T.B., NGUYEN, V.T.K. & NGUYEN, T.T.H. 2024–Examining Monitoring System: Detecting Abnormal Behavior. arXiv preprint arXiv:2402.12179, 19 Feb. 2024 .

[4] GOPANE, S., KOTECHA, R., OBHAN, J. & PANDEY, R.K. 2024– Cheat Detection in Online Examinations Using Artificial Intelligence. ASEAN Engineering Journal, Vol. 14, No. 1, pp. 121–128.

[5] ASKER, B.H. & AL-ALLAF, A.F. 2022– Detecting Cheating in Electronic Exams Using the Artificial Intelligence Approach. International Journal of Mechanical Engineering, Vol. 7, No. 2, pp. 999–1007.

[6] KADDOURA, S. & GUMAEI, A. 2022– Towards Effective and Efficient Online Exam Systems Using Deep Learning-Based Cheating Detection Approach. Intelligent Systems with Applications, Vol. 16.

[7] KUMAR, B.S., DIVYA SRI, M., VINAY, K.R., SRI, K.S.K. & JESSICA, K. 2024– Online Proctoring with Face Analysis and Object Recognition Using YOLO. Journal of Advanced Zoology, Vol. 45, Issue S2, pp. 18–24 .

[8] ERDEM, B. & KARABATAK, M. 2025– Cheating Detection in Online Exams Using Deep Learning and Machine Learning. Applied Sciences (MDPI), Vol. 15, No. 1, Article 400, pp. 1–21, 3 Jan. 2025 .

[9] ESSAHRAUI, S., LAMAAKAL, I., MALEH, Y., ELMAKKAOUI, K., BOUAMI, M.F., OUAHBI, I., ALMOUSA, M., ALQAHTANI, A.A.S. & EL-

LATIF, A.A.A. 2025– Deep Learning Models for Detecting Cheating in Online Exams. Computers, Materials & Continua, Vol. 85, No. 2, pp. 3152–3183, 23 Sep. 2025 .

[10] RAMAMOORTHY, R., ADITHI, S., ANTONY, J., ASHIKA, K. & BASAVARAJ, B. 2025– Design and Implementation of an AI–Powered Hybrid Detection Framework for Real–Time Object and Face Analysis. International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), Vol. 14, Issue. 11.

[11] ONNX RUNTIME DOCUMENTATION. 2025– ONNX Runtime Official Documentation. Available at: https://onnxruntime.ai/docs/

[12] ATOUM, Y., LIN, X., LIU, S., DING, H., HAN, H. & XU, L. 2015– Automated Online Exam Proctoring. IEEE Transactions on Multimedia, Vol. 19, No. 7, pp. 1609–1624, Dec. 2015 .

[13] SAFFOUR, W., MUFDI, F. & HAMMOUD, T. 2021– Improving Viola–Jones Algorithm's Performance in Face Detection for Digital Images and Videos. Homs University Journal, Vol. 43, No. 12, pp. 99–128 .

[14] SAAD ELDIN, N. & AL–ATASSI, Y.A.S. 2020– Applying Transfer Learning in Image Classification. Homs University Journal, Vol. 42, No. 23, pp. 11–52 .