

اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية

المهندس محمد أحمد علي¹

إشراف: الدكتورة اليدا إسبر²
الدكتور سهيل الحمود³

الملخص

تهدف هذه الدراسة إلى تطوير نظام آلي لاكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات، باستخدام تقنيات الرؤية الحاسوبية ونموذج YOLO v11. يركز النظام على كشف أربعة عناصر أساسية داخل الورقة الامتحانية: المعلمات العمودية والمعلمات الأفقية وصندوق رقم الطالب وصندوق نموذج الامتحان. كما يتضمن آلية تصحيح هندسي مزدوج (Affine + Perspective) تعتمد على مجموعة من المعلمات المرجعية ضمن 69 معلمة داخلية، بهدف تسوية الورقة وإزالة الاتحرافات الناتجة عن ظروف التصوير غير المثالية.

تم تدريب النموذج على أوراق من مجموعة بيانات حقيقة من جامعة حمص، وتضم حوالي 9400 ورقة امتحانية تم تصويرها باستخدام ماسحات ضوئية وكاميرات هواف محمولة واستخدمت أداة Label Studio في عملية تسمية البيانات وسرعت العملية بالاعتماد على التسمية المدعومة بالتعلم الآلي (ML-Assisted Labeling). أظهرت النتائج أن متوسط الدقة (mAP) عند عتبة التقطيع على الاجتماع (IoU=0.5) بلغ 99.5%， في حين وصل المتوسط الحسابي للدقة عند عتبات IoU ضمن النطاق [0.5:0.95] إلى 93.64% حيث بينت هذه النتائج كفاءة النظام في التكيف مع اختلاف تصميمات أوراق الإجابة وتوعي ظروف التصوير.

الكلمات المفتاحية: التعرف الآلي على العلامات (OMR)، تسمية البيانات (Data Labeling)، تصحيح الميل، الرؤية الحاسوبية.

¹ طالب دكتوراه - قسم هندسة البرمجيات ونظم المعلومات - كلية الهندسة المعلوماتية - جامعة حمص - سوريا

² أستاذ مساعد - قسم هندسة البرمجيات ونظم المعلومات - كلية الهندسة المعلوماتية - جامعة حمص - سوريا

³ مدرس - قسم هندسة البرمجيات ونظم المعلومات - كلية الهندسة المعلوماتية - جامعة حمص - سوريا

Detection and Analysis of Structural Elements in Multiple-Choice Answer Sheets Using YOLO v11 – An Applied Study on Real-World Data

Eng. Mohammad Ahmad Ali

Supervision: Dr. Alida Isber Dr. Suhel Hammoud

ABSTRACT

This study aims to develop an automated system for detecting and analyzing structural elements in multiple-choice answer sheets using computer vision techniques and the YOLO v11 model. The system focuses on detecting four essential components within the exam sheet: vertical markers, horizontal markers, the student ID box, and the exam test form box. It also incorporates a dual-stage geometric correction mechanism (Affine + Perspective) based on a set of 69 internal reference markers, with the goal of rectifying the sheet and eliminating distortions caused by suboptimal imaging conditions.

The model was trained on a real-world dataset from the University of Homs, comprising approximately 9,400 exam sheets captured using scanners and mobile phone cameras. The data labeling process utilized the Label Studio tool, and the annotation process was expedited by leveraging machine learning-assisted labeling (ML-Assisted Labeling). The results showed that the mean Average Precision (mAP) at an Intersection over Union threshold of 0.5 (IoU=0.5) reached 99.5%, while the mean Average Precision across IoU thresholds in the range [0.5:0.95] reached 93.64%. These results demonstrate the system's effectiveness in adapting to variations in answer sheet designs and diverse imaging conditions.

Keywords: YOLO v11, answer sheets, skew correction, computer vision, Label Studio

1. مقدمة

أصبحت الحاجة إلى أنظمة تصحيح مؤتمته وموثقة لأوراق الإجابة متعددة الخيارات أمر ملح وهام خصوصاً مع الاهتمام المتزايد في التحول الرقمي حول العالم والتقدم المتسارع له بكل القطاعات بما فيها القطاع الأكاديمي وفي مؤسسات التعليم العالي ومع التضخم الكبير لأعداد الطلاب. وتعتمد أنظمة الامتحانات على تقنيات تقليدية في معالجة صور أوراق الإجابة وتقييمها رغم توافر تقنيات التعرف الآلي على العلامات (OMR) منذ عشرات السنوات، بسبب اعتمادها على الكثير من الشروط التي تحد من استخدامها في الواقع ومن دقة تقييمها كأن تكون الورقة ممسوحة بدقة مثالية [1].

تعيق دقة هذه الأنظمة مجموعة من التحديات يجعلها معرضة لأخطاء متكررة في المؤسسات التعليمية التي تعتمد لغات مختلفة ويكون النظام غير مجهز مسقاً للتعامل معها مثل مشاكل الانحراف الهندسي، وتنوع تخطيط الورقة، والتصوير غير المثالي باستخدام الهواتف المحمولة [1,9]. بالإضافة إلى اعتماد العديد منها على تقنيات OCR لاستخلاص رقم الطالب أو النموذج [2].

وتظهر مجموعة من التحديات عند تطوير نماذج ذكاء اصطناعي تعيق العمل على تطوير النماذج الذكية في القطاع التعليمي مثل عملية تجهيز بيانات التدريب من خلال تسمية عناصر الورقة يدوياً كمشكلة حقيقة تتطلب وقتاً وجهداً مضاعف [3].

تقدم هذه الدراسة مقترن عملي متكملاً للتعرف على العناصر الأساسية في أوراق الإجابة متعددة الخيارات باستخدام الذكاء الاصطناعي وتجهيزها للتصحيح الآلي لاحقاً، يجمع النظام المقترن بين قوة أداء نموذج YOLOv11 للكشف الدقيق [4]، وتقنيات الرؤية الحاسوبية في التصحيح الهندسي (Affine + Perspective) [5,6]، وخوارزميات متخصصة لاستخراج رقم الطالب ونموذج الامتحان مباشرة حسب تظليل الفقاعات.

اخترنا النظام على مجموعة من أوراق إجابة امتحانية حقيقة من ضمن 9400 ورقة جمعت من جامعة حمص، غنية بتنسيقات طباعة مختلفة وتم تصويرها باستخدام ماسحات ضوئية وكاميرات موبайл في ظروف إضاءة متنوعة.

يتميز هذا البحث عن غيره بعده مزايا تشمل جمع قاعدة بيانات واقعية في ظروف متعددة، استخدام 9 معلمات عمودية و 6 معلمات أفقية من أصل 69 معلمة داخلية متوزعة على طول

اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية

وعرض الورقة لضمان دقة تصحيح الميلان، ودمجنا أداة Label Studio مع نموذج YOLO لتسمية البيانات (Semi-auto Labeling) [7]، لتوليد مخرجات منظمة وقابلة للتكامل البرمجي مع أنظمة إدارة التعليم.

يعلم النظام على التعرف على الانحراف الفعلي لكل ورقة بناءً على معطياتها الخاصة ويقدم حلولاً مرنة لتصحيحه وتجاوز حالات التصوير غير المثالية بدليلاً عن اعتمادها أن الفرضيات الهندسية محققة افتراضياً عند التصحيح.

يهدف البحث إلى سد فجوة علمية حقيقة في مجال تصحيح أوراق OMR، عبر بناء نظام ذكي متكامل قادر على التعامل مع أوراق متنوعة ومصورة في ظروف مختلفة، وتحليلها بدقة عالية دون الحاجة لتدخل بشري، مما يمهد لتبني أنظمة تصحيح ذكية أكثر موثوقية في المؤسسات التعليمية.

2. أهداف وأهمية البحث

تهدف هذه الدراسة إلى بناء نظام ذكي للتعرف على العناصر البنوية في أوراق الإجابة متعددة الخيارات، يستخدم تقنيات الذكاء الاصطناعي ويراعي التحديات الواقعية المرتبطة بتنوع تصميم الورقة وظروف التصوير المختلفة ويسعى إلى تطوير إطار تقني قابل للتطبيق العملي في المؤسسات التعليمية، حيث يتحقق ما يلي:

- تطوير نموذج كشف ذكي باستخدام YOLOv11 لتحديد العناصر البنوية الهامة في الورقة الامتحانية، بما يشمل المعلومات العمومية والأفقية وصندوق رقم الطالب وصندوق نموذج الامتحان.
- تصحيح الميلان الهندسي بدقة عالية من خلال خوارزمية تعتمد على 15 معلمة موزعين عمودياً وأفقياً من بين 69 معلمة داخلية، مع استخدام تحويل هندسي (Affine +) (Perspective).
- استخلاص رقم الطالب والنموذج باستخدام تحليل فقاعات مخصص يعتمد على التظليل.

- إعداد نظام معالجة مؤتمت يبدأ من الكشف عن العناصر وتحديد موقعها بدقة وصولاً إلى الحصول على ملفات نتائج بالمعلومات المستخرجة (CSV) قابلة للتكامل مع أنظمة إدارة التعليم.
- تسريع عملية إعداد بيانات التدريب باستخدام التسمية بمساعدة نموذج تعلم آلي (ML-Assisted Labeling) من خلال استخدام نموذج YOLO ضمن أداة Label Studio. تأتي أهمية هذا البحث من عدة اعتبارات علمية وتقنية، أبرزها [1]:
 - تعتمد بنية أنظمة OMR الحالية على افتراضات هندسية مثالية تفشل في التعاطي مع أوراق ممسوحة أو مصورة بزوايا ميل مختلفة أو جودة منخفضة.
 - تفتقر الحلول المنتشرة إلى المرونة في التعامل مع تصاميم ورقية غير موحدة.
 - تظهر عوائق إعدادمجموعات التدريب بشكل يدوي أمام تعليم هذه الأنظمة، حيث يُقدم هذا البحث حل متكامل لتسريع عملية تسمية البيانات بجودة عالية.

تلخص مساهمات البحث بالنقاط التالية:

- 1) جمع مجموعة بيانات كبيرة ومخصصة تشمل 9400 ورقة امتحانية ممسوحة ومصورة في بيئة تعليمية فعلية وبأنماط طباعة متعددة.
- 2) تصميم خوارزميات كشف وتصحيح هندي مزدوج أثبتت كفاءتها على كل من الصور الممسوحة والصور الملقطة بكاميرا رقمية مثل كاميرات الهاتف المحمول.
- 3) تقديم خوارزمية لاستخلاص رقم الطالب من شبكة الفقاعات.
- 4) دمج التسمية اليدوية مع التسمية شبه الآلية (Semi-auto Labeling) باستخدام نموذج YOLO عند تجهيز البيانات، لتسريع تسمية البيانات وتحسين التعامل معها.
- 5) تحقيق أتمتة شاملة للعملية، حيث يتم إعداد نتائج مخزنة قابلة للتكامل بشكل آلي.

3. دراسة مرجعية

نستعرض في هذه الفقرة العديد من الدراسات المرجعية المرتبطة بمشكلة البحث في اكتشاف العناصر البنوية الأساسية في أوراق الإجابة وتصحيح الميلان الهندسي واستخلاص المعلومات الخاصة برقم الطالب ونموذج الامتحان. وركزت على مناقشة الأساليب التقليدية والحديثة مع إبراز

الثغرات التقنية التي تمت معالجتها عبر دمج تقنيات حديثة مثل YOLOv11 وخوارزميات تصحيح الميلان، والاعتماد على بيانات واقعية ونظام تسمية بيانات شبه آلي.

1.3. مرونة التخطيط وتصحيح الميلان (Robust Alignment)

ظهرت أهمية مرونة أنظمة التصحيح الآلية في التعامل مع الاختلافات الهندسية لأوراق الامتحانات، حيث تكون الصور الواقعية للأوراق مصورة بزاوية ميل، أو تتضمن انتباحاً عن المسح غير المحوري أو الالتفات اليدوي باستخدام الكاميرات الرقمية مثل الهواف المحمولة. ولهذا فإن الخطوة الأولى الحاسمة في النظام هي تصحيح الميل والانحراف الهندسي وإعادة تسوية الورقة قبل البدء في تحليلها [8].

أظهرت الدراسات المرجعية هذه المشكلة بطرق متعددة، مثل دراسة Jain وزملاؤه [8] التي اعتمدت على استخلاص النقاط الطرفية من حواف الورقة باستخدام مرشح Canny، ثم اكتشاف الحدود المستطيلة الخاصة بصناديق الإجابة باستخدام خوارزمية Ramer–Douglas–Peucker، ليتم بعد ذلك تطبيق تحويل Homography لتسوية الورقة، حقق هذا الأسلوب دقة تصحيح زاوية بلغت 99.33% بزمن معالجة لم يتجاوز 0.5 ثانية لكل صورة دون الحاجة لأي معلمات مرجعية مطبوعة على الورقة ما يجعله ملائم للأوراق القياسية المطبوعة.

وقدم Hafeez وآخرون [9] نهجاً اعتمد على البنية الداخلية للورقة دون استخدام أي معلمات مرجعية أيضاً، بدلاً من ذلك قاموا باستخدام Hough Circle Transform لاكتشاف الفقاعات ثم تحليل توزّعها لتحديد اتجاه الورقة وتصحيح ميلها، أظهر هذا النهج مرونة في الأوراق التي لا تحتوي على معلمات واضحة لكنه بالمقابل يتطلب انتظاماً عالياً في توزيع الفقاعات حتى لا تؤثّر الانحرافات أو الفقاعات المفقودة على النتيجة النهائية، قالت هذه الاستراتيجية معدل الخطأ في تصحيح الأوراق المعرضة لطباعة أو مسح ضوئي مشوش أو تالف حيث انخفض من 13.41% و 7.22% في الأنظمة التقليدية إلى 0.44% باستخدام النظام المقترن.

وأستخدمت تقنية الكشف عن الأشياء في دراسات أخرى حديثة لتحديد المعلمات واستخدامها في تصحيح الميل، حيث طور Tinh وآخرون [10] نموذج YOLOv8 تعرفت من خلاله على أقسام الصورة بما فيها معلومات الطالب وصناديق الإجابة. عمل النموذج على اكتشاف ثلاثة معلمات داخلية في (الزاوية العليا اليسرى، الزاوية العليا اليمنى، والزاوية السفلى اليسرى) من ورقة

الامتحان وتم تحديد الميل وتصحيحه، كذلك تم التعرف على رقم الطالب بواسطة 10 تسميات ينعرف عليها النموذج وهي أرقام صحيحة من 0 إلى 9، وحقق النموذج دقة في اكتشاف رقم الطالب وإجاباته المختارة بلغت 99.5%.

2.3 التشغيل باستخدام الهواتف المحمولة ودعم ظروف التصوير المتنوعة

ظهرت الحاجة إلى تطوير أنظمة OMR قادرة على التعامل مع صور ملقطة باستخدام كاميرات الهاتف المحمولة في ظل ظروف إضاءة غير متجانسة وزوايا ميل متقاوتة. وفرضت تحديات إضافية على الأنظمة التقليدية التي غالباً ما تكون مصممة لمعالجة صور ممسوحة ضوئياً على خلفية موحدة ويزوايا مثالية [11].

ركز Largo وآخرون [11] في دراستهم على تطوير تطبيق محمول يسمح بتصوير الورقة وتحليلها في الزمن الحقيقي، وقد تضمنت المنظومة وحدة تصحيح تلقائي للميل، ثم مرحلة تعزيز بصري تعتمد على Contrast Limited Adaptive Histogram (CLAHE) لتسوية التباين داخل الصورة مما يسمح بتحسين نتائج التقطيل حتى في الصور منخفضة الجودة، ثم اعتمدت الدراسة على تحليل كثافة التقطيل داخل فقاعات الإجابة.

في دراسة أخرى للباحث Hafeez وآخرون [12] اختبروا نموذجهم على صور مأخوذة بكاميرات هاتف محمولة في ظروف مختلفة من الإضاءة والميلان (ميل بزاوية 10 إلى 20 درجة). واعتمدت الدراسة على خوارزمية Hough Circle لاكتشاف دوائر نصف قطرها 15 بكسل داخل الورقة ثم رسم خطوط تمر بمراتز هذه الدوائر؛ حيث تُحسب الزاوية بينها وبين المحور الأفقي لتحديد زاوية الميل وتصحيحها واستخدم (Pixel Projection) لتحديد الإجابات المختارة. أظهرت النتائج أن النموذج حقق معدل خطأ بقدر 3.33%.

3.3 استخدام تسمية البيانات (Labeling)

ظهرت أهمية تجهيز مجموعة البيانات وتسميتها خطوة أساسية في تطوير أنظمة OMR المبنية على التعلم العميق، وإن نجاح النماذج مثل YOLO وTransformer على الرغم من التقدم الكبير في بنيتها لا يتحقق إلا إذا دُرِّبت على بيانات مسماة بشكل جيد ومتناقض. لكن في معظم الدراسات السابقة تم التعامل مع عملية التسمية إما يدوياً بالكامل أو بأساليب محدودة، مما جعل المرحلة مكلفة من حيث الوقت والجهد [10].

اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية

تظهر ندرة ملحوظة في الدراسات التي اعتمدت على دمج التعلم الآلي مع التسمية، حيث اعتمدت غالبية الأعمال مثل تلك التي قدمها Tinh وآخرون [10] على التسمية اليدوية باستخدام أدوات مثل LabelImg.

كذلك في دراسة Hafeez وآخرون [9]، تم إجراء التسمية يدوياً بالكامل على أكثر من 86 ألف حالة ضمن الأوراق المتعددة الخيارات، ولم يدمج أي من هذه الأعمال آليات مدعومة بنماذج التعلم الآلي في عملية التسمية مما يمثل جهد بشري هائل ويطرح تساؤلات حول الاتساق الداخلي والموثوقية بين المسميات المختلفة، خصوصاً عند وجود حالات معقدة مثل الفقاعات المشطوبة أو الجزئية.

4. مواد وطرائق البحث

يستعرض هذا القسم المنهجية المفصلة التي تم اتباعها في تطوير نظام يتعرف على العناصر البنوية (المعلمات والنموذج الإمتحاني ورقم الطالب) ضمن أوراق الإجابة متعددة الخيارات باستخدام الذكاء الاصطناعي بدءاً من عملية جمع البيانات وتسميتها لاستخدامها في تدريب نموذج الكشف واستخلاص المعلومات الخاصة برقم الطالب والنموذج الإمتحاني وتصحيح الانحراف الهندسي للورقة.

1.4. توصيف مجموعة البيانات Dataset

بدأ العمل بجمع أوراق امتحانية حقيقة من جامعة حمص، شملت نماذج متعددة لمقررات واختصاصات من السنوات الأولى وحتى الخامسة. كان الهدف من تنويع البيانات هو محاكاة سيناريوهات حقيقة متنوعة تعكس الفوارق في التسويق الظاهري وعدد الأسئلة وتصميم شبكة الفقاعات وحالات التقطيل الجزئي.

تم جمع الأوراق باستخدام طريقتين:

- المسح الضوئي باستخدام جهاز Ricoh 5400 Auto Feeder Scanner : تصوير 6400 صورة لأوراق إجابة متعددة الخيارات.
- التصوير اليدوي باستخدام هواتف محمولة: تصوير 3000 صورة لأوراق إجابة متعددة الخيارات بزوايا ميل مختلف وفي ظروف إضاءة متنوعة.

تضمنت الأوراق على أربع عناصر بنوية رئيسية:

- (1) معلمات عمودية (v_markers): عددها 63 معلمة مصطفة على الحافة اليسرى لورقة الإجابة متعددة الخيارات، واستخدمت كمرجع هندي لتقدير الميل العمودي.
 - (2) معلمات أفقية (h_markers): عددها 6 معلمة (من ضمنها أول معلمة عمودية)، مصطفة أعلى ورقة الإجابة متعددة الخيارات واستخدمت لضبط المعاذة الأفقية.
 - (3) صندوق رقم الطالب: يحتوي شبكة من 10 أعمدة \times 10 صفوف لتمثيل الرقم الجامعي.
 - (4) صندوق نموذج الامتحان: يتضمن أربع فقاعات (A, B, C, D) لتحديد نسخة النموذج.
- صنفت الأوراق وخزنت ضمن مجلدات وفق نوع الجهاز المستخدم (ماش أو كاميرا)، لاستخدامها لاحقاً في عينات التدريب والتحقق.

2.4. تسمية البيانات (Labeling)

استخدمت أداة Label Studio لتسمية البيانات على مرحلتين كما يلي:

1.2.4 التسمية اليدوية (Manual Labeling)

جهزنا مجموعة البيانات الأولية بشكل يدوي باستخدام أداة Label Studio واخترنا 200 صورة امتحانية متنوعة من حيث ظروف التصوير وتصاميم الطباعة.

شملت عملية التسمية 4 فئات بنوية رئيسية داخل ورقة الإجابة:

- h_marker: المعلمات الأفقية (6 معلمات في أعلى الورقة)
- v_marker: المعلمات العمودية (9 معلمات متوزعة على يسار الورقة)
- student_info_box: صندوق رقم الطالب
- test_form_box: صندوق نموذج الامتحان

2.2.4 التسمية شبه الآلية (Semi-Auto Labeling) باستخدام YOLO

اعتمدت منهجية التسمية بمساعدة التعلم الآلي (ML-Assisted Labeling) ضمن أداة Label Studio، تسمح باستخدام نموذج مدرب مسبقاً لتسريع عملية التسمية ضمن الخطوات التالية:

- (1) تدريب نموذج YOLOv8-L على مجموعة الصور التي تم تسميتها يدوياً في الخطوة السابقة (200 صورة) وقسمت (160 صورة للتدريب، 20 للتحقق، و20 للاختبار).

اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية

(2) إجراء عملية تسمية تلقائية لـ 100 صورة جديدة من خلال كشف العناصر باستخدام نموذج YOLO المدرب.

(3) إعادة ضبط النتائج يدوياً ضمن Label Studio في حال حدوث أخطاء بسيطة في الموضع أو التسمية لضمان دقة النتائج.

3.4. النماذج والخوارزميات المستخدمة

يعتمد النظام المقترن على بنية متكاملة من النماذج العميقه والخوارزميات الهندسية، تهدف إلى اكتشاف العناصر البنوية الأساسية داخل ورقة الإجابة ومعالجة المشكلات الناتجة عن الميل أو التشوه البصري أثناء التصوير.

1.3.4. نموذج YOLOv11-m لاكتشاف العناصر البنوية

تم اختيار نموذج YOLOv11-m في هذه الدراسة بناءً على التحسينات المتميزة التي يقدمها من حيث الهيكل والأداء، مما يجعله مناسباً جداً لاكتشاف العناصر الأساسية الدقيقة داخل أوراق الإجابة، مثل المعلمات وصناديق المعلومات. يعتمد هذا النموذج على تحسينات تم إدخالها على الإصدارات السابقة من YOLO، حيث تم استبدال الوحدة التقليدية المسماة C2f بوحدة جديدة تُعرف بـ C3K2، وهي عبارة عن طبقة مدمجة تحتوي على تلافيف أصغر حجماً وأكثر كفاءة. هذا التعديل يساعد النموذج على تنفيذ العمليات بسرعة أكبر، مع الحفاظ على جودة التمثيل البصري للمعلومات المهمة في الصورة [13,14].

كذلك، يحتفظ YOLOv11 بوحدة معروفة من الإصدار السابق YOLOv8 يُدعى SPPF (Spatial Pyramid Pooling Fast) والتي تعمل على توسيع مدى الرؤية للنموذج (receptive field) وتعزيز قدرته على اكتشاف الكائنات الصغيرة داخل الصورة، مثل المعلمات الصغيرة والحدود الدقيقة لصناديق البيانات [13,14].

من أبرز الإضافات الجديدة في YOLOv11 هي وحدة C2PSA، والتي تستخدم تقنية تُعرف باسم "الانتباه متعدد الرؤوس" (Multi-Head Attention) لدمج المعلومات من مختلف القنوات والمناطق داخل الصورة. هذا يتبع للنموذج التركيز على المواقع المهمة بدقة أعلى، خاصة تلك التي تحتوي على العناصر البنوية داخل الورقة [13,14].

كما تم تحسين آلية حساب المربعات التوجيهية (anchor boxes) بشكل يجعلها تتكيف تلقائياً مع حجم وشكل الكائنات المختلفة الموجودة في الصور، وهو أمر مهم لأن عناصر مثل صندوق رقم الطالب وصندوق نموذج الامتحان قد تختلف أبعادها من ورقة لأخرى [13,14].

اعتماداً على هذه التحسينات، قمنا بتدريب نموذج YOLOv11-m على صور عالية الدقة (1920×1920 بكسل)، وباستخدام حجم دفعه تدريبية (batch size) يساوي 8. صُمم التدريب لاكتشاف أربع فئات رئيسية هي: المعلمات العمودية، المعلمات الأفقية، صندوق رقم الطالب، وصندوق نموذج الامتحان. وأظهرت نتائج الاختبار أن النموذج يعمل بكفاءة عالية في ظروف تصوير متعددة تشمل صور مأخوذة من ماسحات ضوئية أو كاميرات هواتف، مما يجعله جزءاً أساسياً في مرحلة معالجة الورقة وتصحيح انحرافها.

2.3.4. خوارزمية تصحيح الميل والتسوية الهندسية

يُعد تصحيح الميل والانحراف (Skew & Perspective Correction) من أهم مراحل معالجة أوراق الإجابة، حيث إن الانحرافات الكبيرة في الصورة يمكن أن تؤدي إلى فشل تام في استخراج المعلومات من الحقول الثابتة. تم تطوير خوارزمية تعتمد على تحليل مواضع المعلمات المرجعية h_marker و v_marker لتقدير الزاويتين الهندسيتين وتصحيحهما بتسلاسل دقيق.

تعتمد الطريقة المقترحة على ركيزتين:

- المحاذة العمودية (Vertical Alignment) : باستخدام 9 معلمات عمودية.
- المحاذة الأفقية (Horizontal Alignment) : باستخدام 6 معلمات أفقية.

أولاً يتم حساب الزاويتين باستخدام تحليل الانحدار الخطي، ثم تطبيق تحويلين هندسيين: (1) لتصحيح الدوران Affine Transform (2) لتسوية الصورة الهندسية Perspective Transform.

معاملات خوارزمية تصحيح الميل والتسوية الهندسية (Deskew & Dewarp) كما هو موضح في الجدول 1:

اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية

الجدول 1 - معاملات خوارزمية تصحيح الميل والتسوية الهندسية

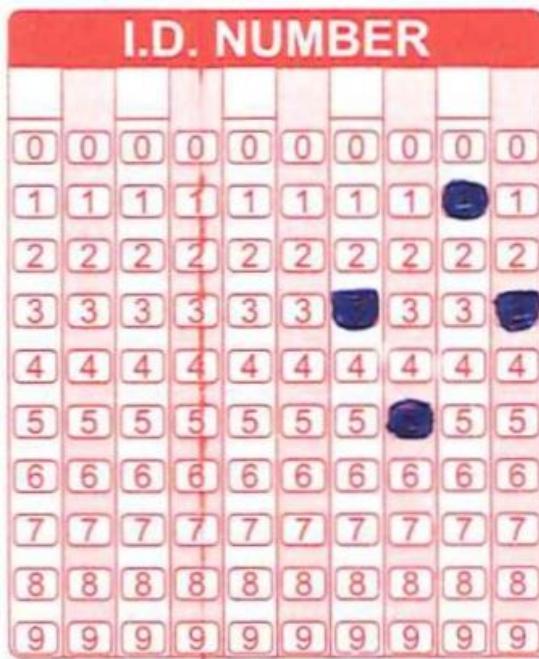
الشرح	الاستخدام	القيمة	الباراميتر
لضمان عدم اقتطاع الزاوية اليسرى من الورقة المصححة	حدود اليسار لصندوق التصحيح	45	pad_left
ضمان تغطية كل منطقة في الزاوية اليمنى	حدود اليمين لصندوق التصحيح	40	pad_right
حشوة علوية لمنع القطع من الأعلى	أعلى الورقة	45	pad_top
حشوة سفلية	أسفل الورقة	45	pad_bottom
-	غير مستخدم في deskew مباشرة	—	blockSize
-	—	—	threshold
نقاط الزوايا تُحسب حسب موقع الماركرات وصناديق الكشف، ثم تضاف الحشوات المذكورة			ملاحظات
حول "margin" معظم القيم هنا لضمان الصورة المصححة وعدم فقد بيانات مهمة			

تعتمد خوارزمية المعالجة الهندسية على سلسلة خطوات دقيقة لضمان تسوية الورقة مبينة في الخطوة 1:

- المدخلات: صورة ورقة الإجابة كما تم التقاطها بالماسح الضوئي أو الكاميرا المحمولة.
 - المخرجات: صورة مصححة ومستوية هندسياً جاهزة للتحليل.
1. بداية الخوارزمية
 2. تطبيق نموذج YOLOv11 للكشف عن العناصر البنوية التالية:
 - العلامات العمودية (h_markers) و العلامات الأفقية (v_markers)
 - صندوق رقم الطالب (student_info_box)
 - صندوق نموذج الامتحان (test_form_box)
 3. استخراج إحداثيات مراكز v_markers وتخزينها في القائمة: $V = [(x_1, y_1), \dots, (x_n, y_n)]$
 4. حساب زاوية الميل θ باستخدام خوارزمية linear regression على النقاط V
 5. تدوير الصورة بمقدار θ حول مركز الصورة للحصول على image_rot1
 6. إعداد قائمة نقاط أفقية للتدوير تضم أول h_markers + v_marker لتعزيز الثبات في أعلى الورقة $H_{extended} = h_markers + v_markers$
 7. حساب زاوية الميل الأفقي θ عبر linear fit على نقاط $H_{extended}$
 8. تدوير image_rot1 بمقدار θ حول مركز الصورة للحصول على image_rot2
 9. ترتيب v_markers تصاعدياً حسب إحداثيات y من الأعلى إلى الأسفل
 10. ترتيب h_markers تصاعدياً حسب إحداثيات x من اليسار إلى اليمين
 11. حساب النقاط الأربع لزوايا الورقة:
$$((Padding + v_marker) + (v_marker + Padding)) \text{ لأعلى } y \text{ ، } (v_marker + (v_marker + Padding)) \text{ لأدنى } y \text{ ، } (h_marker + (h_marker + Padding)) \text{ لأول } x \text{ ، } (h_marker + (h_marker + Padding)) \text{ لآخر } x$$
 12. تعريف النقاط: $src_pts = [top_left, top_right, bottom_right, bottom_left]$
 13. حساب أبعاد الصورة المصححة (العرض والارتفاع) بشكل ديناميكي من src_pts
 14. تحديد نقاط الهدف: $dst_pts = [(0, 0), (width, 0), (width, height), (0, height)]$
 15. حساب التحويل المنظوري: $P = getPerspectiveTransform(src_pts, dst_pts)$
 16. تطبيق التحويل على image_rot2:
$$image_corrected = warpPerspective(image_rot2, P)$$
 17. إرجاع النتيجة: $Return image_corrected$
 18. نهاية الخوارزمية

3.3.4. خوارزمية استخراج رقم الطالب

بعد تطبيق خوارزمية تصحيح الميل الهندسي وتسوية منظور ورقة الإجابة باستخدام التحويلات Perspective وAffine، يصبح صندوق رقم الطالب (student_info_box) المخصص لتعبئة رقم الطالب باستخدام الفقاعات مهيأً بشكل دقيق للاقتصاص والمعالجة. تحتوي هذه المنطقة على شبكة منتظمة تتألف من 10 أعمدة عمودية، يمثل كل عمود خانة عددية من الرقم الجامعي (من الأحاد إلى عشرات الملايين)، وتحتوي كل عمود على 10 فقاعات مرقمة من 0 إلى 9. تبدأ العملية بقص المنطقة المكتشفة اعتماداً على إحداثيات الكشف من YOLO، يلحق ذلك سلسلة من المعالجات لتحسين وضوح الفقاعات وتحديدها حتى الحصول على الرقم الجامعي النهائي. يبين الشكل 1 صورة صندوق رقم الطالب في أحد الأوراق الامتحانية:



الشكل 1 - صندوق رقم الطالب في أحد الأوراق الامتحانية

وكما يبين الجدول 2 معاملات خوارزمية استخراج رقم الطالب

الجدول 2 – معاملات خوارزمية استخراج رقم الطالب

الشرح	الاستخدام	القيمة	الباراميتر
حجم النافذة لعزل الخطوط/الخلفية	adaptiveThreshold	13	blockSize
تعويض للعتبة التكيفية (يخفض/يرفع الحساسية)	adaptiveThreshold	5	C
تعزيز التباين قبل العتبة	CLAHE	2.7	clipLimit
حجم شبكة تعزيز التباين	CLAHE	(8, 8)	tileGridSize
تستخدم لتحديد أماكن الفواصل الأفقية	عتبة في إسقاط الصفوف	من 0.6 max	threshold
المسافة الدنيا المقبولة بين الفواصل	لاكتشاف الفواصل بين الأسطر	20, 10	gaps
ليأخذ القليل قبل الخط	عند تحديد بداية منطقة الإجابة	-5	padding
التخلص من حواف جانبية غير مهمة	إزالة هامش جانبي	من 1% العرض	x_trim
مجال متوقع بين الفقاعات	في كشف قمم الفقاعات	(0.040, 0.050)	spacing_range
زيادة عرض القص 20%	لتوصيع صورة الفقاعة	0.2	widen_ratio
إغلاق فراغات بسيطة بعد threshold	closing (morphologyEx)	(2, 2)	kernel
تستخدم هذه القيم لجعل استخراج كل عامود/سطر دقيق ومستقر في ظروف طباعة مختلفة			ملاحظات

تبين الخوارزمية 2 تسلسل خطوات خوارزمية استخراج رقم الطالب:

اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية

- المدخلات: الصورة المصححة ومستوية هندسياً (image_corrected)
 - المخرجات: سلسلة مكونة من 10 أرقام تمثل الرقم الجامعي الكامل للطالب.
1. بداية الخوارزمية
 2. تطبيق نموذج YOLO 11 لاكتشاف إحداثيات صندوق رقم الطالب (student_info_box) من الصورة المصححة (image_corrected).
 3. قص صندوق رقم الطالب المكتشف.
 4. تطبيق المعالجة المسبقة:
 - تحويل الصورة إلى تدرج رمادي (grayscale).
 - تطبيق Gaussian Blur لتنقيل الضجيج.
 - تطبيق CLAHE لتعزيز التباين.
 - تطبيق adaptive thresholding للحصول على صورة ثنائية تُظهر الفقاعات بوضوح.
 5. استخراج الجزء السفلي من صندوق رقم الطالب الذي يحتوي الشبكة الفعلية للرقم، ويتم ذلك بتحليل الإسقاط الأفقي (horizontal projection) للبكلسات السوداء.
 6. قص المنطقة السفلية، ثم تقسيمها أفقياً إلى 10 أعمدة متساوية:
 - كل عمود يمثل خانة رقمية من رقم الطالب (الأحاد، العشرات...).
 7. لكل عمود من الأعمدة العشرة:
 - حساب الإسقاط العمودي (vertical projection) لاكتشاف موقع الفقاعات.
 - استخدام خوارزمية كشف القمم (peak detection) لاستخراج موقع 10 فقاعات مرشحة.
 - في حال الفشل، تطبيق تقسيم ثابت للعمود إلى 10 مناطق أفقية متساوية.
 8. لكل فقاعة (bubble):
 - حساب العتبة (threshold) باستخدام طريقة Otsu من صورة صندوق الإجابة كاملاً.
 - حساب نسبة البكلسات السوداء داخل الفقاعة.
 - اختيار الفقاعة ذات أعلى كثافة مظللة كممثل للرقم في هذا العمود.
 9. جمع نتائج الفقاعات المظللة لكل عمود من الأعمدة وتكوين رقم الطالب النهائي من اليمين إلى اليسار.
 10. حفظ نسخة من الصورة مع إعادة التسمية بالرقم المستخرج exam_id في مجلد النتائج
 11. نهاية الخوارزمية

الخوارزمية 2 - خوارزمية استخراج رقم الطالب

4.3.4 استخراج نموذج الامتحان (Test Form Identification)

بعد معالجة الصورة وتصحيح انحرافها وتسويتها، يصبح بالإمكان تحليل صندوق نموذج الامتحان (test_form_box) بشكل آلي وبدقة عالية. يحتوي هذا الصندوق عادةً على أربع فقاعات عمودية مرتبة من الأعلى إلى الأسفل، كل منها تمثل أحد النماذج المتاحة (A، B، C، D) ويتطلب من الطالب تظليل واحدة فقط لتحديد النموذج الخاص بورقه.

تهدف هذه المرحلة إلى تحديد الفقاعة التي تم تظليلها من قبل الطالب من خلال تحليل دقيق للكثافة داخل كل فقاعة بعد قصها واستخراجها.

يبين الشكل 2 صندوق نموذج امتحان لأحد أوراق الإجابة:



الشكل 2 – صندوق نموذج امتحان لأحد أوراق الإجابة

ويبين الجدول 3 معاملات خوارزمية استخراج نموذج الامتحان:

الجدول 3 – معاملات خوارزمية استخراج نموذج الامتحان

الشرح	الاستخدام	القيمة	الباراميتر
عتبة تكيفية في مرحلة line detection	adaptiveThreshold	21	blockSize
تعويض العتبة التكيفية	adaptiveThreshold	19	C
تعزيز التباين قبل العتبة	CLAHE	2.9	clipLimit
حجم شبكة CLAHE	CLAHE	(8, 8)	tileGridSize

اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية

الشرح	الاستخدام	القيمة	الباراميتر
كشف الفاصل أسفل رأس الصندوق	في إسقاط الصوفوف	من max	threshold
Fallback gap = 10 للصور الصعبة	لاكتشاف الفواصل	20, 10	gaps
حدود اقتصاص الصندوق	نسبة اقتصاص الجزء العلوي	من h 0.17	top
حدود اقتصاص الصندوق	نسبة اقتصاص الجزء السفلي	من h 0.83	bottom
إزالة حواف جانبية	اقتصاص جانبي	من w 0.30	left
إزالة حواف جانبية	اقتصاص جانبي	من w 0.70	right
له كل خيار bubble رباع المسافة	تقسيم لمناطق الفقاعات	h // 4	region_height
يُحسب مباشرة، ويستخدم كحد لتمييز الفقاعة	يتم حسابه لكل صورة فرعية	متغير (0.40+)	otsu_threshold
الاقتصاص النهائي/المعالجة مُعتمدة على اكتشاف خطوط أو اقتصاص نسبي يدوي			ملاحظات

وتبين الخوارزمية 3 آلية استخراج نموذج الامتحان

— المدخلات: الصورة المصححة والمستوية هندسياً (image_corrected) وتحتوي على صندوق نموذج الامتحان.

— المخرجات: المحرف الذي يمثل نموذج الامتحان المختار (A, B, C, D).
الخطوات:

1. بداية الخوارزمية
2. تطبيق نموذج YOLO 11 لاكتشاف إحداثيات صندوق نموذج الامتحان (test_form_box) من الصورة المصححة (image_corrected) واقتتصاص الصندوق.
3. تحسين الصورة للكشف عن الخطوط والفاصل:
 - تحويل المنطقة إلى تدرج رمادي (grayscale).
 - تطبيق Gaussian Blur لتنقيل الضجيج وتحسين وضوح الفقاعات.
 - تطبيق CLAHE لتعزيز التباين في ظروف الإضاءة المختلفة.
 - تطبيق adaptive thresholding لتحويل الصورة إلى صورة ثنائية تُمكِّن من تمييز الفقاعات المطللة.
4. حساب الإسقاط الأفقي لمحتوى الصورة الثانية (المناطق السوداء)، وتحديد الفواصل الأفقيَّة ذات الفجوات الكبيرة في الإسقاط لتحديد بداية منطقة الفقاعات (bubble_area).
5. قص منطقة الفقاعات بدقة، مع تقليم نسي من الحواف لضمان عزل الخيارات فقط.
6. تطبيق Otsu threshold على الصورة الأصلية لاستخراج عتبة مناسبة لفصل الفقاعات عن الخلفية.
7. تقسيم المنطقة إلى 4 أقسام أفقيَّة متساوية، كل منها يمثل خياراً من خيارات النموذج A و B و C و D.
8. يتم تطبيق كل مما يلي على كل قسم (خيار):
 - تطبيق threshold ثابت يتم حسابه على كامل الجزء الخاص بالنموذج الامتحاني.
 - حساب نسبة التظليل (البكسلات السوداء) داخل الفقاعة.
9. اختيار الخيار الذي تجاوز العتبة وكان الأعلى في نسبة التظليل كـ form_version
10. إذا لم يتجاوز أي خيار العتبة، يتم اختيار الفقاعة الأعلى كثافة كمحولة بديلة.
11. إرجاع النتيجة: إعادة نموذج الاختبار form_version
12. نهاية الخوارزمية

5. التجارب والنتائج ومناقشتها

1.5. بيئة التجريب

تم تنفيذ عملية تدريب النماذج واختبارها ضمن بيئة Google Colab Pro، والذي بدوره وفر إمكانيات مناسبة للتعامل مع النماذج العميقه والبيانات بدقة عالية، تضمنت البيئة التقنية المستخدمة وحدة معالجة مركبة Intel Xeon سداسية النوى (cores 6) مع 12 خيط معالجة (threads) بسرعة 2.2 جيجاهرتز، مما أتاح معالجة متوازية فعالة خلال تحميل البيانات والتدريب. كما تم تخصيص 83.5 جيجابايت من الذاكرة العشوائية (RAM) مما ساهم في تسريع عمليات التخزين المؤقت وتخفيض معدلات التباطؤ الناتجة عن نقص الذاكرة أثناء التدريب.

أما فيما يخص وحدة المعالجة الرسومية (GPU) فقد تم استخدام بطاقة NVIDIA A100 ذات ذاكرة من نوع HBM2 容量 40 جيجابايت، والتي تعتبر من أقوى بطاقات المعالجة المخصصة لتدريب نماذج التعلم العميق، وخاصة مع الصور الكبيرة ومعالجة المهام المتعددة. ساهمت هذه البيئة في تسريع التدريب بشكل كبير مما سمح بتحقيق كفاءة حسابية عالية أثناء تنفيذ جميع مراحل ضبط النموذج وتقديره.

2.5. مقاييس تقييم الأداء

لقياس فعالية النماذج المدرية على كشف عناصر بنوية في الورقة الامتحانية، تم اعتماد مجموعة من المعايير المعروفة في مجال الرؤية الحاسوبية والتي تمثل الأساس في تقييم نماذج التعرف على الأشياء. شملت هذه المقاييس:

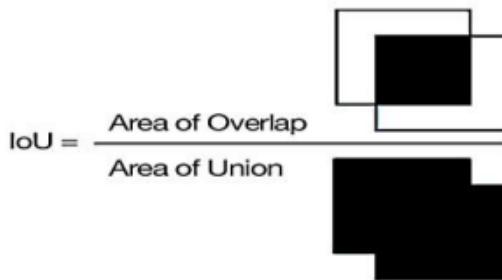
1) مصفوفة الالتباس (Confusion Matrix)

تُستخدم مصفوفة الالتباس لعرض نتائج اكتشاف النموذج ضمن أربع حالات رئيسية [15]:

- TP (True Positive): اكتشاف النموذج العنصر بشكل صحيح.
- TN (True Negative): لم يكتشف النموذج العنصر وكان فعلاً غير موجود.
- FP (False Positive): اكتشف النموذج عنصر لم يكن موجود.
- FN (False Negative): لم يكتشف النموذج العنصر رغم أنه كان موجود بالفعل.

2) التقاطع على الاجتماع (IoU - Intersection over Union)

ظهر تحديد موقع العناصر بأفضل دقة ممكنة ك مهمة رئيسية عند استخدام نماذج التعرف على الأشياء. يوضح الشكل 3 وجود صندوقين محيطين أحدهما باللون الأسود يمثل المربع المحيط بقاعدة الحقائق الأساسية للعنصر، والآخر باللون الأبيض يمثل ناتج تبؤ نموذج الكشف، نلاحظ أنهما لا يتدانان بشكل تام لذلك تحتاج إلى قياس مدى تطابق أو سوء تطابقهما كنتيجة دقة اكتشاف حدود العنصر، لذا يتم حساب التقاطع على الاجتماع IoU بقسمة مساحة تقاطع الصندوقين على مساحة اجتماعهما كما هو موضح في الشكل، ويتم اعتبار التعرف أفضل كلما اقتربت قيمة IoU من 1، ويتم اعتبار التوقع جيد عندما تتجاوز قيمة IoU عتبة محددة (مثل 0.5) [15]:



الشكل 3 – علاقة التقاطع على الاجتماع (IoU) [15]

(3) متوسط الدقة (mean Average Precision – mAP)

عند تقييم جودة نماذج الكشف عن الأشياء يُعد مقياس mAP من أكثر مؤشرات الأداء استخداماً، يقيس متوسط دقة التوقعات عبر جميع الفئات بناءً على مدى تطابق الصناديق المتوقعة مع الصناديق الحقيقية ضمن مستويات مختلفة من التداخل الهندسي. تم في هذه الدراسة استخدام معيارين رئيسيين [15]:

- mAP@0.5: يتم اعتبار التوقع صحيح عندما يكون تقاطع المساحة بين الصندوق المتوقع وال حقيقي (IoU) ≥ 0.5 .
- mAP@0.5: وهو متوسط الدقة عند عشر عتبات مختلفة من IoU (من 0.5 إلى 0.95 بفواصل 0.05)، ويُعتبر المقياس الأكثر صرامة لقياس دقة التحديد.
- يمكن حساب متوسط الدقة الكلي mAP من أجل n فئة وفق العلاقة (1) [15]:

$$mAP = \frac{\sum_{i=0}^n AP_i}{n}$$

حيث أن AP متوسط الدقة للفئة i ، وتقيس قدرة النموذج على اكتشاف الإيجابيات الحقيقية (TP) من بين جميع التوقعات الإيجابية (TP+FP) لها وتحسب وفق العلاقة (2):

$$AP = \frac{TP}{TP + FP}$$

(4) مقاييس أخرى

يستخدم البحث أيضاً عدد من المقاييس وهي: الدقة Accuracy والإحكام Precision والاستدعاء Recall ومقياس إف1 F1-Score. تُحسب هذه المقاييس وفق المعادلات التالية [16]:

$$\frac{TP + TN}{TP + TN + FP + FN} = \text{الدقة}$$

$$\frac{TP}{TP + FP} = \text{الإحكام}$$

$$\frac{TP}{TP + FN} = \text{الاستدعاء}$$

$$2 \times \frac{Precision \times Recall}{Precision + Recall} = \text{مقياس إف1}$$

3.5. تدريب نموذج YOLOv8 لتسمية البيانات (Semi-auto Labeling)

تم تطوير نموذج أولي باستخدام YOLOv8-1 في المرحلة الأولى من تجهيز مجموعة البيانات بهدف تسريع عملية التسمية اليدوية من خلال التسمية شبه الآلية التي تتيحها أداة Label Studio حيث يتم دمج النموذج المدرب داخلها لتوليد اقتراحات أولية. تم تدريب النموذج على مجموعة بيانات تحتوي على 200 صورة تم تسميتها يدوياً، تم تقسيم هذه المجموعة إلى 160 صورة

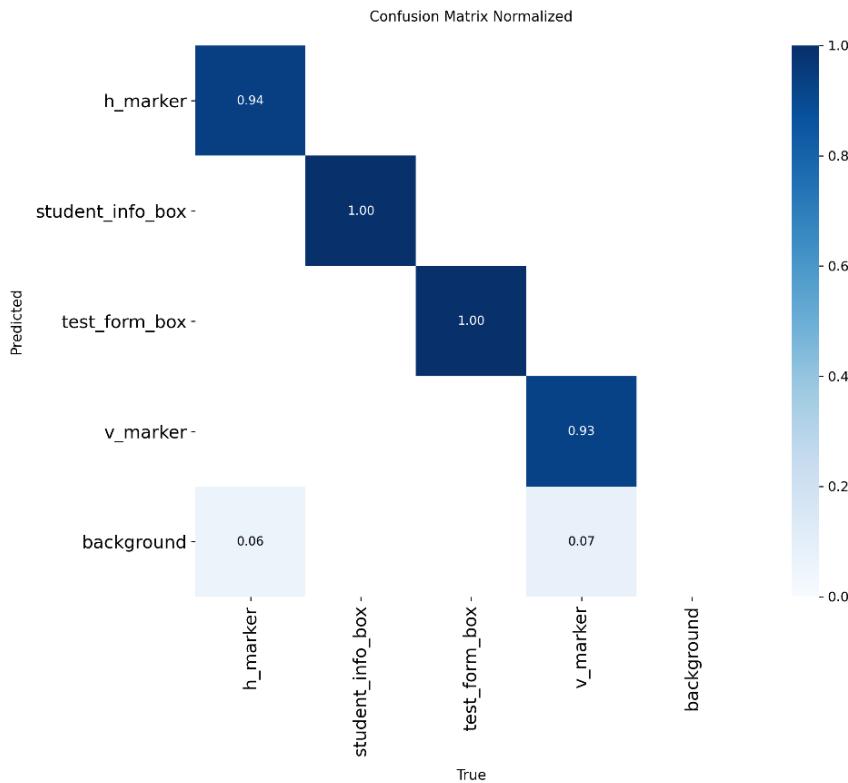
للتدريب و 20 صورة للتحقق و 20 صورة للاختبار، وشملت عملية التسمية أربع فئات بنوية أساسية: **.test_form_box** و **student_info_box** و **v_marker** و **h_marker**

أدت عملية التسمية شبه الآلية إلى إعداد 100 صورة جديدة خلال وقت قياسي، وساهمت في تقليل الجهد اليدوي بنسبة ملحوظة. وقد أظهرت نتائج النموذج التجريبي دقة عالية بلغت 98.10 %، كما أظهرت مصفوفة الالتباس ومؤشرات الأداء لكل فئة على حدة تميز النموذج في الكشف عن **test_form_box** و **student_info_box** بنسبة **mAP@0.5** بلغت 0.995، بينما كانت النتائج أقل نسبياً لفئة **h_marker** بسبب صغر حجمها النسبي. تظهر هذه النتائج في الشكل 4 وفي الجدول 1 كما يلي:

الجدول 1 -نتائج اختبار تدريب نموذج yolo-v8 على 200 ورقة

mAP@ 0.5: 0.95	mAP @ 0.5	الاستدعاء Recall	الإحكام Precision	عدد الحالات	عدد الصور	الفئات
0.762	0.970	0.940	1.000	100	20	h_marker
0.853	0.964	0.928	1.000	181	20	v_marker
0.991	0.995	1.000	1.000	20	20	student_ info_box
0.988	0.995	1.000	1.000	20	20	test_form_ box
0.898	0.981	0.967	1.000	321	20	إجمالي الفئات

اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية



الشكل 4 - مصفوفة الالتباس لنتائج اختبار تدريب نموذج yolo-v8 على 200 ورقة

يظهر الشكل 5 صور توضيحية لعملية الكشف في الكشف شبه الآلي باستخدام YOLO داخل الأداة:



الشكل 5 – صور توضيحية لاقتراحات التسمية في الكشف شبه الآلي باستخدام YOLO

4.5. تقسيم مجموعة البيانات

تم تقسيم مجموعة البيانات المكونة من 300 صورة تضم أوراق الإجابة التي سميت في مرحلة تحضير البيانات (200 منها بشكل يدوي و100 صورة تم تسميتها بشكل شبه آلي) إلى ثلاثة مجموعات وهي مجموعة التدريب Training (70% من مجموعة البيانات الكلية) ومجموعة التحقق Validation (15%) ومجموعة الاختبار Testing (15%).

5.5. الضبط الدقيق لنموذج YOLO v11m

لتحقيق أفضل أداء لنموذج YOLOv11-m استخدمنا الضبط التلقائي للمعاملات (Hyperparameter Tuning) باستخدام مكتبة Optuna، وهي مكتبة مفتوحة المصدر تعتمد على تحسين بيزاني (Bayesian Optimization) حيث تعمل على البحث العشوائي التتابعي وتوجه التجارب السابقة لتسريع الوصول إلى أفضل القيم [17]. عملت Optuna على اختبار مئات القيم المختلفة للمعاملات، واختيرت الإعدادات التي تحقق أعلى قيمة لقياس mAP@[0.5:0.95] على مجموعة التحقق.

شملت عملية الضبط معاملات متعددة، منها ما يتعلق بخصائص التعلم مثل معدل التعلم الأساسي (lr0) والرخم (momentum)، ومنها ما يتعلق بتعزيز البيانات مثل Mosaic، ودرجة الميل (shear)، والتغيرات اللونية (hsv)، بالإضافة إلى معاملات الثقة (confidence) ولا IoU (confidence) المطلوب لقبول التوقعات. يبين الجدول 2 أفضل إعدادات ضبط نتجت عن Optuna:

اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية

الجدول 2 -أفضل قيم لتدريب نموذج YOLO v11m Hyperparameter Tuning

المعامل	الوصف	القيمة
Lr0	معدل التعلم الأساسي (Learning rate)	0.000236
Momentum	الزخم في تحديث الأوزان (Momentum)	0.9336
Dropout	نسبة الإسقاط في الشبكة لتجنب overfitting	0.0638
Scale	حجم التغيرات العشوائية في الصورة أثناء التدريب	0.0878
Mosaic	نسبة استخدام تقنية Mosaic augmentation	0.567
Shear	درجة الانحراف الزاوي للصورة	17
IoU	العتبة المطلوبة لاحتساب تقاطع المربعات	0.5829
Conf	الحد الأدنى للثقة لقبول التنبؤ	0.473
hsv_h/s/v	التغيرات اللونية العشوائية في الصور للتعزيز	0.04 / 0.79 / 0.31
Patience	عدد الدورات بدون تحسن قبل التوقف المبكر	50
Batch	حجم العينة في كل خطوة تدريبية	8

تم استخدام الإعدادات المبينة في الجدول السابق لتدريب النموذج النهائي وفقاً للمعلمات التالية المبينة في الجدول 3 التالي:

الجدول 3 -إعدادات تدريب نموذج YOLO v11m

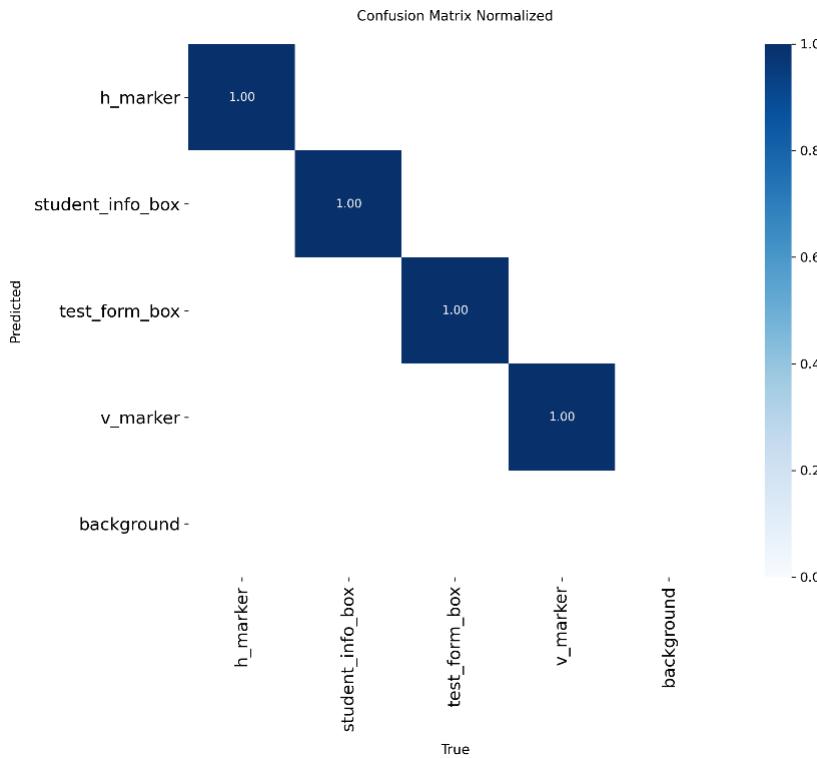
المكون	القيمة / التهيئة
عدد الدورات (Epochs)	500 دورة (301 بسبب التوقف المبكر)
حجم الصورة (imgsz)	1920 × 1920 بكسل
المحسن (Optimizer)	AdamW
التسريع المختلط (AMP)	(Automatic Mixed Precision)
عدد العمليات المتوازية (workers)	8
فاصل الحفظ الدوري (save_period)	حفظ النموذج كل 25 دورة

6.5. نتائج أداء نموذج YOLO v11m المدرب

تم تقييم أداء النموذج النهائي YOLOv11-m المدرب على مجموعة اختبار تتألف من 30 صورة تتوزع من حيث جودة التصوير وجود الميلان، اعتمد التقييم على إعدادات صارمة بعتبة (IoU ≥ 0.75) وحد أدنى للثقة (Confidence = 0.3)، وذلك لضمان الموثوقية العالية في الكشف.

1.6.5. مصفوفة الالتباس الخاصة بأداء نموذج الكشف YOLO v11m

أظهرت مصفوفة الالتباس أداء ممتاز لجميع الفئات كما هو مبين في الشكل 6:



الشكل 6 – مصفوفة الالتباس الخاصة بنتائج اختبار YOLO v11m على 300 ورقة

2.6.5. نتائج اختبار أداء نموذج YOLO v11m حسب الفئات

اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية

حقق النموذج معدلات دقة واسترجاع كاملة (100%) لمعظم الفئات كما هو مبين في الجدول

:4

الجدول 4 -نتائج اختبار أداء نموذج YOLO v11m حسب الفئات على 300 ورقة

mAP@ 0.5: 0.95	mAP @ 0.5	الاستدعاء Recall	الإحكام Precision	عدد الحالات	عدد الصور	الفئات
0.852	0.995	1	1	150	30	h_marker
0.914	0.995	0.996	1	271	30	v_marker
0.995	0.995	1	1	30	30	student_ info_box
0.984	0.995	1	1	30	30	test_form_ box
0.936	0.995	0.999	1	481	30	إجمالي الفئات

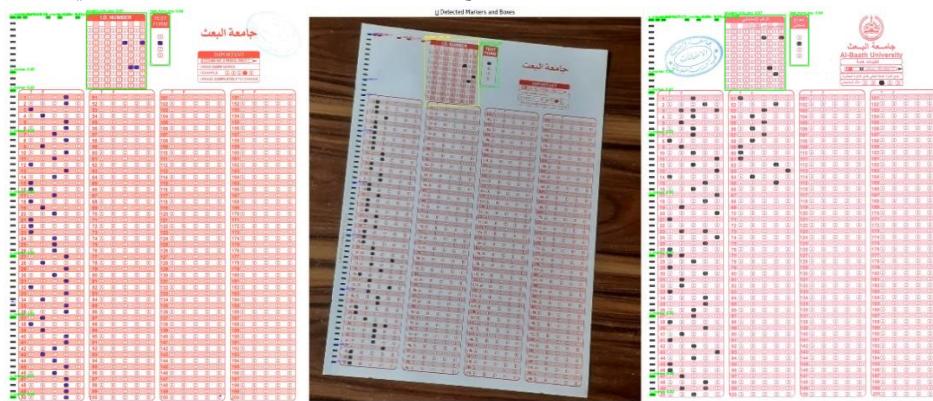
3.6.5. الأداء الإجمالي للنموذج

أظهرت النتائج أداء مميز في جميع المقاييس كما هو موضح في الجدول 5 أدناه:

الجدول 5 -نتائج اختبار الأداء الإجمالي YOLO v11m على 300 ورقة

القيمة	المقياس
0.9950	Precision (mAP@0.5)
0.9950	Recall (mAP@0.75)
0.9364	mAP@0.5:0.95
0.9950	F1 Score

توضح الصور التالية المبنية في الشكل 7 نتائج الكشف على ثلاثة أوراق من مجموعة الاختبار، وأظهرت الصور المخرجة دقة تحديد ممتازة لجميع العناصر البنوية الأربع المستهدفة في النظام:



الشكل 7 - صور توضيحية لنتائج اختبار أداء نموذج الكشف YOLO v11m

7.5. نتائج تطبيق خوارزمية تصحيح الميل والتسوية الهندسية

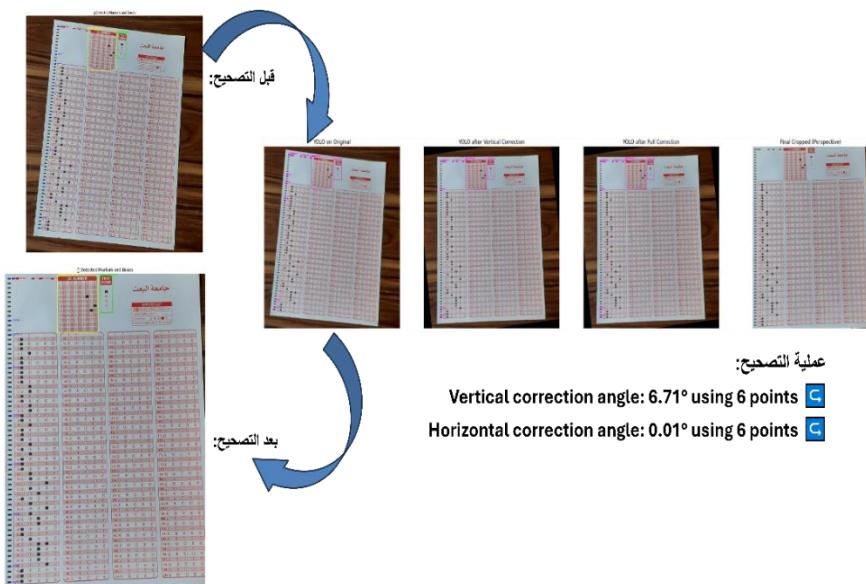
تم تنفيذ خوارزمية تصحيح الميل والتسوية الهندسية لأوراق إجابة تم اختيارها من مجموعة البيانات بالاعتماد على نتائج نموذج YOLOv11 في اكتشاف العناصر البنوية للورقة يُظهر المثال الموضح في الشكل 8 كيف تبدأ العملية بصورة ملقطة بواسطة كاميرا الهاتف المحمول بزاوية ميل واضحة، حيث يتم تطبيق مرحلتين من التصحيح الهندسي:

- المرحلة الأولى: تدوير عمودي يتم حساب زاويته عبر تحليل محاور المعلمات العمودية.
- المرحلة الثانية: تدوير أفقى دقيق اعتماداً على اصطفاف المعلمات الأفقية وأول معلمة .v_marker

بعد التصحيح السابق يتم تطبيق عملية تسوية منظور (Perspective Flattening) للحصول على صورة مستوية بالكامل جاهزة لعمليات القص والتحليل.

توضح المخرجات المعروضة في الشكل 8 انتقال الورقة من حالتها الأصلية التي تحوي على ميلان كبير إلى صورة مستوية بشكل أفضل وجاهزة لاستخراج المعلومات منها بدقة أفضل.

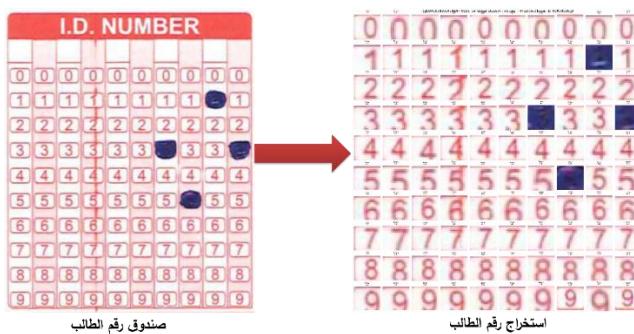
اكتشاف وتحليل العناصر البنوية في أوراق الإجابة متعددة الخيارات باستخدام YOLO v11 - دراسة مطبقة على بيانات واقعية



الشكل 8 - صور توضيحية لنتائج تطبيق خوارزمية تصحيح الميل والتسوية الهندسية

8.5. نتائج تطبيق خوارزمية استخراج رقم الطالب

أظهرت نتائج تطبيق الخوارزمية على مجموعة من الصور الواقعية عن قدرة جيدة في استخلاص الرقم الجامعي الكامل من الفقاعات المخصصة لذلك، وأظهرت الصور الناتجة انتظام واضح في تحديد الخانات الرقمية مع تمييز دقيق للفقاعات المختارة. وقد تبين أن الخوارزمية تتکيف مع تفاوت الإضاءة وجودة الطباعة، كما هو مبين في الصورة المرفقة في الشكل 9:

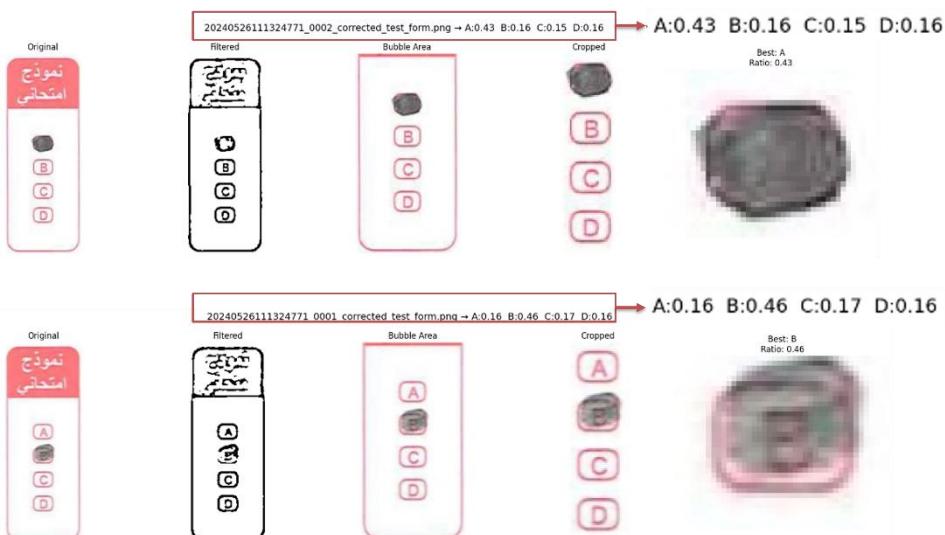


20240528103424147_0021_corrected_student_info.png → Predicted Exam ID: 000003513

الشكل 9 – صور توضيحية لنتائج تطبيق خوارزمية استخراج رقم الطالب

9.5. نتائج تطبيق خوارزمية استخراج نموذج الامتحان

أظهرت الخوارزمية المخصصة لاستخراج نموذج الامتحان (A/B/C/D) فاعلية جيدة في تحديد الفقاعة المملوقة فعلياً، رغم الاختلافات الطفيفة في طرق تعبيئة الطلاب لها والفرق بين اللونية بين الصور المختلفة. توضح النتائج المصورة المبينة في الشكل 10 قدرة النموذج على تجاهل العلامات غير المكتملة أو التظليل الجزئي أو التشويه في الطباعة، وتحديد النسخة الصحيحة بشكل مناسب عبر مختلف الصور.



الشكل 10 - صور توضيحية لنتائج تطبيق خوارزمية استخراج نموذج الامتحان

6. مقارنة مع الدراسات المرجعية السابقة:

أن معظم الدراسات المرجعية التي تم استعراضها في هذا البحث ([12-8]) انحصرت في تقييم أداء أنظمة التصحيح الآلي لأوراق الإجابة بناءً على معيار دقة تصنيف الفقاعات (Bubble Classification Accuracy)، أي قدرة النظام على التمييز بين الفقاعات المطللة وغير المطللة، دون أن تطرق إلى تقديم نتائج تفصيلية أو مؤشرات أداء رقمية لاكتشاف العناصر البنوية داخل الورقة الامتحانية مثل المعلومات العمودية والأفقية أو صناديق المعلومات (student_info_box و test_form_box)، كما لم تقدم هذه الدراسات تقييمات مستقلة لمراحل التصحيح الهندسي (Skew & Perspective Correction) باستخدام الخوارزميات الثلاث (YOLO، تصحيح الميل الهندسي، وخوارزميات استخراج صناديق المعلومات).

بالتالي، لم تتوفر في الأدبيات المنشورة نتائج كمية مباشرة حول كفاءة اكتشاف وتحديد هذه العناصر البنوية أو دقة التصحيح الهندي كمهام منفصلة، بل انحصرت معظم التقارير على عرض معدل الدقة النهائي للنظام في تصنيف إجابات الطلاب، وهو ما لا يتيح إجراء مقارنة مباشرة مع النظام المقترن في هذه الدراسة.

لذا، جاء هذا البحث ليس هذه الفجوة المنهجية، إذ يقدم ولأول مرة تقارير رقمية دقيقة لمقاييس اكتشاف العناصر البنوية وتصحيح الانحراف الهندي، ويوثق نتائج مفصلة باستخدام نموذج YOLOv11 ومعالجة هندسية متقدمة ضمن بيئة بيانات واقعية متباينة، مما يعزز من أهمية نتائجه العلمية.

7. الاستنتاجات والتوصيات

يوضح البحث الحالي أن النظام يعالج أوراق الإجابة كما هي، ويُخرج النتائج بناءً على ما هو مطلل فعلياً من قبل الطالب دون التدخل في صحة البيانات أو اكتشاف الأخطاء المنطقية الناتجة عن سوء التظليل أو السهو (مثل ترك النموذج الامتحاني دون تظليل، أو وجود أكثر من فقاعة مطللة في خانة واحدة من الرقم الجامعي). في حال عدم تمكن النظام من استخراج النتائج بشكل

كامل (نتيجة لخطأ في الإدخال أو فشل في الاستخراج)، يتم ترك الحقل فارغاً أو تعيين قيمة افتراضية (كإرجاع "0" في الرقم الجامعي)، مع تمييز هذه الحالات في سجل النتائج.

لا يتضمن النظام في نسخته الحالية آلية تلقائية للتحقق من صحة أو منطقة القيم المستخرجة أو إصدار تنبية حول أخطاء الإدخال البشري، وإنما يعتمد على توفير معلومات كافية (صور، نتائج فارغة، علامات فشل) ليتمكن المستخدم من مراجعتها يدوياً عند الحاجة.

من الممكن مستقبلاً دمج وحدات تحقق تلقائي (validation) تقوم بفحص صحة الرقم الجامعي (عدد الخانات، تميز خانة واحدة فقط في كل عمود)، والتحقق من وجود تظليل لنموذج الامتحان، مع إصدار تقارير عن الحالات الشاذة أو الناقصة.

أثبتت هذا البحث فعاليته في معالجة العديد من الإشكاليات المنهجية والتطبيقية المرتبطة بأنظمة تصحيح أوراق الإجابة متعددة الخيارات (OMR)، وذلك من خلال تقديم إطار متكامل يدمج بين الرؤية الحاسوبية والتعلم العميق ضمن بنية قابل للتطبيق العملي.

أولاً، يمثل توفير مجموعة بيانات ضخمة وحقيقة إحدى أهم المساهمات العلمية في هذا البحث، إذ إن غالبية الدراسات السابقة كانت تعتمد علىمجموعات بيانات صغيرة، غالباً ما تكون صناعية أو غير منشورة. إن طبيعة هذه البيانات - التي جُمعت من أوراق ممسوحة أو مصورة بكاميرا رقمية مثل كاميرات الهاتف المحمولة، وتحتوي على تباين حاد في جودة التظليل والانحراف الزاوي - مكنت من اختبار النموذج في ظروف قريبة من الواقع العملي، وساهمت في تحسين قدرة النماذج على التعميم.

ثانياً، تم تطبيق منهج التسمية شبه الآلية (Semi-Auto Labeling) باستخدام نموذج YOLO أولي للمساعدة في عملية تجهيز مجموعة التدريب، مما ساعد في تسريع الإعداد وتخفيض نسبة الخطأ البشري، مع الحفاظ على تناسق البيانات المسممة. تمثل هذه المنهجية حل عملي لتوسيعمجموعات البيانات بدون كلفة بشرية مرتفعة، وهي قابلة للتكرار في مجالات متعددة من تطبيقات الرؤية الحاسوبية.

ثالثاً، أثبت نموذج YOLOv11 تفوق ملحوظ مقارنةً بنموذج YOLOv8 وهو النموذج المستخدم في معظم الدراسات السابقة سواء من حيث القدرة على اكتشاف العناصر الدقيقة (مثل المعلومات المرجعية وصناديق البيانات) أو من حيث الدقة الإجمالية. وقد بيّنت اختبارات النموذج

تحسن فعلي في مقاييس mAP عند استخدامه في بيئة بيانات واقعية، وهو ما يضعه ضمن الخيارات الواحدة لنماذج OMR الحديثة.

رابعاً، اعتمد البحث على منهجية تصحيح مزدوجة الميلان باستخدام المعلم الهيكلي المكتشفة (v_markers، h_markers)، وهو ما تجاوز ضعف معالجة الانحراف الزاوي في العديد من الدراسات السابقة التي اعتمدت على حواف الورقة فقط. حسّنت هذه المنهجية بشكل مباشر من جودة تسوية الورقة وبالتالي من دقة جميع المهام اللاحقة مثل الاستخراج والتصنيف.

خامساً، تم تطوير خوارزميات دقيقة لاستخراج كل من رقم الطالب ونموذج الامتحان عبر تحليل الفقاعات مباشرة دون الاعتماد على OCR أو إدخال يدوي، مما جعل النظام أكثر توافقاً مع أوراق لا تتضمن أرقاماً مكتوبة أو تحتوي على تشويش في الطباعة أو المسح، وقد حققت الخوارزميات نتائج قوية في الاكتشاف.

توصيات بحثية مستقبلية:

- إعادة تصميم المعمارية على هيئة نموذج واحد متعدد المهام (Multi-task Learning) لنقليل عدد المراحل المنفصلة وتحقيق تعلم مشترك بين المهام (مثل الكشف، التصنيف، التصحيح).
- التوسيع في تصميم النموذج ليشمل تصاميم أوراق امتحانية مختلفة (غير موحدة)، بما في ذلك تعدد الواقع المحتملة لصندوق المعلومات أو أنماط تضليل مختلفة للفقاعات.
- دمج النظام داخل واجهة إلكترونية كاملة تشمل رفع الصور واكتشاف النماذج وتصحيح ميلان الورقة واستخراج بيانات الطالب والورقة وتوليد النتائج النهائية، مما يمهد لاعتماد النظام في بيئة مؤسسية حقيقة (جامعة أو مدرسة).

8. المراجع

- [1] de Elias, E. M., Tasinazzo, P. M., & Hirata Jr, R. 2021 Optical mark recognition: Advances, difficulties, and limitations, SN Computer Science, Vol. 2(5), 367.
- [2] Mahmud, S., Biswas, K., Alam, A., Rudro, R. A. M., Anannya, N. J., Mouri, I. J., & Nur, K. 2024 Automatic multiple choice question

- evaluation using tesseract ocr and yolov8, 2024 IEEE Conference on Artificial Intelligence (CAI), pp. 246–252, IEEE.
- [3] Huang, Q., & Zhao, T. 2024 Data collection and labeling techniques for machine learning, arXiv preprint arXiv:2407.12793.
 - [4] Khanam, R., & Hussain, M. 2024 Yolov11: An overview of the key architectural enhancements, arXiv preprint arXiv:2410.17725.
 - [5] Erkelens, C. J. 2015 The perspective structure of visual space, *i-Perception*, Vol. 6(5), 2041669515613672.
 - [6] Wagner, M., & Gambino, A. J. 2016 Variations in the anisotropy and affine structure of visual space: A geometry of visibles with a third dimension, *Topoi*, Vol. 35(2), 583–598.
 - [7] Label Studio Documentation. (n.d.) <https://labelstud.io/guide/>
 - [8] Jain, V., Malik, S., & Bhatia, V. 2022 Robust image processing based real-time optical mark recognition system, 2022 IEEE 6th Conference on Information and Communication Technology (CICT), pp. 1–5, IEEE.
 - [9] Hafeez, Q., Aslam, W., Lali, M. I., Ahmad, S., Alqahtani, M., & Shafiq, M. 2023 Fault Tolerant Optical Mark Recognition, *Computers, Materials & Continua*, Vol. 74(2).
 - [10] Tinh, P. D., & Minh, T. Q. 2024 Automated Paper-based Multiple Choice Scoring Framework using Fast Object Detection Algorithm, *International Journal of Advanced Computer Science & Applications*, Vol. 15(1).
 - [11] Largo, L. D., Guillermo, J., Jancinal, A. R., & Wata, M. 2022 Bubble sheet multiple choice mobile checker with test grader using optical mark recognition (omr) algorithm, 2022 5th International Conference on Electronics and Electrical Engineering Technology (EEET), pp. 27–33, IEEE.
 - [12] Hafeez, Q., Aslam, W., Aziz, R., & Aldehim, G. 2024 An Enhanced Fault Tolerance Algorithm for Optical Mark Recognition Using Smartphone Cameras, *IEEE Access*.
 - [13] Mao, M., & Hong, M. 2025 YOLO Object Detection for Real-Time Fabric Defect Inspection in the Textile Industry: A Review of YOLOv1 to YOLOv11, *Sensors* (Basel, Switzerland), Vol. 25(7), 2270.
 - [14] He, Z., Wang, K., Fang, T., Su, L., Chen, R., & Fei, X. 2024 Comprehensive Performance Evaluation of YOLOv11, YOLOv10, YOLOv9, YOLOv8 and YOLOv5 on Object Detection of Power Equipment, arXiv preprint arXiv:2411.18871.

- [15] Imran, A.I., Imran, B. 2023 Improved YOLOv3–Tiny model for human detection, Al-Baath University Journal – Series of Mechanical, Electrical and Informatics Engineering Sciences, Vol. 45(82), 211–222.
- [16] Obi, J.C. 2023 A comparative study of several classification metrics and their performances on data, World Journal of Advanced Engineering Technology and Sciences, Vol. 8(1), 308–314.