

استخراج فقاعات الإجابة من أوراق الإجابة متعددة الخيارات باستخدام

نموذج SegFormer وتقنيات تحليل الصورة

المهندس محمد أحمد علي¹

الدكتور سهيل الحمود³

إشراف: الدكتورة اليدا إسبر²

الملخص

تهدف هذه الدراسة إلى تطوير نظام فعال ودقيق لاستخراج فقاعات الإجابة من أوراق الإجابة متعددة الخيارات باستخدام تقنيات التقسيم الدلالي (Semantic Segmentation) وخصوصاً نموذج التعلم العميق SegFormer-B0. تم تدريب النموذج على صور بحجم ثابت (1024×1024 بكسل)، باستخدام خوارزمية التحسين AdamW ودالة الخسارة CrossEntropyLoss. أظهرت النتائج أداءً متميزاً، حيث وصلت الدقة (Accuracy) إلى 0.9977، والاستدعاء (Recall) إلى 0.9961، والدقة التنبؤية (Precision) إلى 0.9970، ومعامل F1 Score إلى 0.9965، ومؤشر التقاطع على الاتحاد (IoU) إلى 0.9931، ما يدل على كفاءة النموذج العالية في التحديد الدقيق لحدود صندوق الإجابة.

اعتمدت الدراسة أيضاً على خوارزمية متكيفة لاكتشاف حدود صفوف الأسئلة، باستخدام تقنية الإسقاط الأفقي وتحليل القناة الحمراء (Red Channel Projection) لتحديد وفصل الأسطر بدقة. كما تم تطبيق آلية احتياطية (Fallback) لتجنب فشل عملية التقسيم في الحالات التي تفتقر فيها الصورة لمؤشرات واضحة. تضمن البحث في مرحلته الأخيرة استخدام خوارزمية التجميع DBSCAN لاكتشاف وتحديد مواضع الفقاعات بدقة عالية، مما يعزز موثوقية النتائج وملاءمة النظام للاستخدام العملي في تطبيقات التصحيح الآلي داخل البيئات التعليمية.

الكلمات المفتاحية: التعرف الآلي على العلامات (OMR)، المحولات (Transformers)،

SegFormer، تقنيات الإسقاط (Projection Techniques)، فقاعات الإجابة (Bubbles).

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

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

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

Extracting Answer Bubbles from Multiple-Choice Answer Sheets Using SegFormer and Image Analysis Techniques

Eng. Mohammed Ahmad Ali

Supervision: Dr. Alida Isber

Dr. Suhel Hammoud

ABSTRACT

This study aims to develop an efficient and accurate system for extracting answer bubbles from multiple-choice answer sheets using semantic segmentation techniques. especially, the deep learning model SegFormer-B0. The model was trained on images with a fixed size (1024×1024 pixels) utilizing the AdamW optimizer and the CrossEntropyLoss function. The results demonstrated outstanding performance, with accuracy reaching 0.9977, recall 0.9961, precision 0.9970, F1 score 0.9965, and intersection over union (IoU) 0.9931, indicating the model's high capability in accurately delineating the boundaries of the answer box.

The study also employed an adaptive algorithm to detect question row boundaries, utilizing horizontal projection and red channel analysis (Red Channel Projection) to accurately segment and separate the rows. A fallback mechanism was implemented to ensure the robustness of the segmentation process in cases where clear boundary indicators were absent. In the final stage, the DBSCAN clustering algorithm was used to accurately detect and locate the answer bubbles, further enhancing the reliability of the results and ensuring the system's suitability for practical use in automated grading applications within educational environments.

Keywords: Optical Mark Recognition (OMR), Transformers, SegFormer, Semantic Segmentation, Projection Techniques, Answer Bubbles, Automated Grading.

1. مقدمة

توفر عملية تصحيح أوراق الإجابة متعددة الخيارات السرعة والموضوعية في رصد أداء عمليات التقييم الامتحانية مما جعلها من المهام الأساسية في البيئات التعليمية [7]. ومع ذلك، تواجه الأنظمة التقليدية لتصحيح هذه الأوراق تحديات تقنية عند التعامل مع تباين جودة الأوراق الممسوحة ضوئياً أو المطبوعة بأنماط مختلفة [7]. من بين الأساليب الأكثر شيوعاً في الأبحاث الحديثة نماذج كشف الكائنات (Object Detection) مثل سلسلة YOLO، والتي حققت نجاحاً ملحوظاً في تحديد مواضع العناصر الصغيرة والمتوسطة الحجم بدقة عالية داخل الصور واكتشاف العناصر المرجعية في الورقة الامتحانية. [3,6,7]

تواجه عملية اكتشاف العناصر الكبيرة مثل صناديق الإجابة (Answer Box) صعوبة كبيرة بالرغم من كفاءة YOLO في رصد العناصر الدقيقة كالعلامات والرموز، خاصة في وجود اختلاف كبير في أبعاد الكائنات ضمن الصورة نفسها [3,7,19]. يعود ذلك إلى طبيعة بنية YOLO المعتمدة على تقسيم الصورة إلى شبكات صغيرة وتخصيص مرشحات (Anchors) للأحجام الأكثر شيوعاً، مما يجعلها أكثر ملاءمة لاكتشاف الكائنات الصغيرة أو المتوسطة، وأقل دقة أو استقراراً عند محاولة اكتشاف الكائنات الكبيرة التي تحتل جزءاً كبيراً من الصورة أو تتداخل مع الحواف، مما قد يؤدي إلى قصور في تحديد حدودها بدقة أو حتى فقدانها تماماً في بعض الحالات. [3,6,7]

تتيح نماذج التقسيم الدلالي المعتمدة على المحولات (Transformers) مثل SegFormer معالجة الصورة كاملة بشكل تسلسلي مع فهم السياق المكاني لكل جزء من أجزاء الصورة بغض النظر عن حجم أو موقع العنصر [6]. يمتاز نموذج SegFormer بقدرته على استخلاص مناطق الاهتمام الكبيرة والدقيقة بنفس الكفاءة، مستفيداً من بنية الانتباه الذاتي (Self-Attention) التي تدعم الربط بين جميع أجزاء الصورة بشكل ديناميكي وفعال. [6]

تهدف هذه الورقة إلى تقديم إطار منهجي متكامل لاستخراج صناديق الإجابة من أوراق الإجابة متعددة الخيارات، ثم تقطيعها إلى صفوف أسئلة واستخراج الفئات بدقة عالية. يعتمد النظام المقترح على تدريب نموذج SegFormer-B0 لتقسيم صورة الورقة إلى صندوق الإجابة والخلفية، متبوعاً بخوارزميات إسقاط وتجزئة متكيفة تعتمد على التحليل الإحصائي والإسقاط الأفقي لتحليل الحدود الفاصلة بين الأسطر. وفي المرحلة الأخيرة، يتم استخراج الفئات باستخدام تقنيات معالجة

مسبقة متقدمة وخوارزمية التجميع DBSCAN لضمان فصل الفقاعات حتى في حالات التداخل أو التشويش.

تستعرض الدراسة نتائج تجريبية دقيقة على بيانات واقعية، وتقرن أداء النظام مع الأنظمة التقليدية، مؤكدة تفوق النظام المقترح في دقة الاستخلاص وموثوقية الأداء، مما يعزز فرص اعتماده في التطبيقات التعليمية الفعالية ويساهم في تقليل الوقت والجهد المطلوبين لعملية التصحيح اليدوي.

2. أهداف وأهمية البحث

يهدف هذا البحث إلى تطوير نظام مؤتمت ودقيق لاستخراج فقاعات الإجابة من أوراق الإجابة متعددة الخيارات مع التركيز على معالجة التحديات المرتبطة باختلاف جودة الصور وتعدد أنماط الطباعة. يسعى النظام المقترح إلى الاستفادة من إمكانيات نماذج التقسيم الدلالي المعتمدة على المحولات (مثل SegFormer) لتجاوز تحديات نماذج كشف الكائنات التقليدية، خاصة فيما يتعلق باكتشاف العناصر البنيوية الكبيرة مثل صناديق الإجابة، والتي غالباً ما تتعرض للفقدان أو عدم التحديد الدقيق بسبب قيود بنية YOLO [3,6,7].

تتجلى أهمية هذا البحث في تقديم إطار عمل متكامل وقابل للتعميم على مختلف أنماط وتصاميم أوراق الإجابة، بما يعزز دقة التصحيح الآلي ويحد من تدخل العامل البشري، وبالتالي يقلل من احتمالية الخطأ ويزيد من كفاءة العملية التعليمية. كما يساهم النظام المقترح في تسهيل إعداد بيانات تدريبية منظمة وحديثة لأنظمة الذكاء الاصطناعي الأخرى، ويضع الأسس التقنية اللازمة لتطوير تطبيقات تصحيح آلي قابلة للاستخدام في البيئات التعليمية الواقعية بمختلف الظروف والإمكانيات التقنية [6,7]، حيث يمكن تلخيص أهمية البحث الحالي في النقاط التالية:

1. **تحسين دقة التصحيح الآلي** يساهم النظام المقترح في رفع مستوى الدقة عند استخراج مناطق الإجابة والفقاعات، مقارنةً بالأساليب التقليدية، مما يقلل احتمالية وقوع الأخطاء البشرية.

2. **معالجة اختلافات جودة الصور** يوفر النظام مرونة عالية في التعامل مع التباين الكبير في جودة الأوراق الممسوحة ضوئياً أو المطبوعة، سواء من حيث الدقة أو تباين الألوان وأنماط الطباعة. [1,6]

3. تجاوز تحديات نموذج YOLO يتغلب النموذج المقترح على قصور نماذج كشف الكائنات التقليدية مثل YOLO في اكتشاف العناصر الكبيرة، مثل صناديق الإجابة، من خلال توظيف نماذج التقسيم الدلالي. [3,6,7]
4. إمكانية التعميم يتمتع النظام بقابلية التكيف مع تصاميم وأشكال متعددة لأوراق الإجابة، مما يجعله ملائماً لمختلف المؤسسات التعليمية.
5. توفير إطار بيانات حديث يتيح النظام إنشاء بيانات تدريبية منظمة وحديثة، يمكن الاستفادة منها في تطوير وتدريب نماذج ذكاء اصطناعي متقدمة أخرى. [6]
6. تقليل الزمن والجهد يسرع النظام من عمليات التصحيح ويقلل الحاجة للتدخل اليدوي، ما ينعكس إيجاباً على كفاءة العملية التعليمية ككل.

3. دراسة مرجعية

تُعتبر تقنيات التعرف الضوئي على العلامات (OMR) واحدة من أبرز التقنيات في مجال التعليم والتقييم الآلي، حيث تتيح إمكانية التصحيح التلقائي والسريع لأوراق الاختبارات التي تعتمد على الإجابات ذات الخيارات المتعددة. فعلى مدار العقود الماضية، شهدت هذه التقنيات تحولات كبيرة، من الأنظمة التقليدية التي اعتمدت على أجهزة مخصصة ومكلفة [18]، إلى الأنظمة القائمة على تقنيات التصوير الرقمي والمعالجة الحاسوبية الحديثة [16]. ومنذ اعتماد التقنيات الرقمية في هذا المجال، ظهرت تحديات عديدة، أبرزها اختلاف جودة الصور وتباين أساليب الطباعة، ما أدى لضرورة تطوير خوارزميات أكثر ذكاءً وقابلية للتكيف مع هذه الظروف [7، 16].

3.1 نماذج تقليدية Traditional Models

في البداية، ركزت الأبحاث على أساليب تقليدية مثل استخدام خوارزميات كشف الحواف ومعالجة الصور باستخدام مرشحات بسيطة، والتي أظهرت تحديات كبيرة في التعامل مع الضوضاء والتشويش الموجود في الصور الناتجة عن المساحات الضوئية أو الكاميرات الرقمية منخفضة الجودة [12، 17، 13]. ومع زيادة الطلب على حلول أقل تكلفة وأكثر فعالية، برزت تقنيات جديدة تعتمد على المعالجة المسبقة المتقدمة، كاستخدام تقنيات تحسين التباين مثل (CLAHE) [1]، وطرق الفلتر مثل (Gaussian Blur) وتقنيات عتبة التكيف (Adaptive Thresholding)، التي وفرت قدرة أكبر على معالجة الصور وتحسين دقة الكشف [1، 13].

3.2. نماذج تقنيات التعلم الآلي Artificial Intelligence Models

ومع ظهور تقنيات التعلم الآلي والشبكات العصبية الاصطناعية، انتقلت الأبحاث إلى استخدام شبكات مثل Neocognitron لتحقيق أداء أفضل في كشف وتحديد العلامات الضوئية [11]. وعلى الرغم من أن هذه النماذج وفرت تحسينات ملحوظة مقارنة بالطرق التقليدية، إلا أنها ظلت محدودة من حيث القدرة على التعميم والتعامل مع البيئات والظروف المتنوعة.

3.3. نماذج شبكات عصبية عميقة Deep Neural Networks Models

في السنوات الأخيرة، شهد مجال OMR تقدماً كبيراً نتيجة إدخال نماذج كشف الكائنات المتقدمة القائمة على التعلم العميق، مثل عائلة نماذج YOLO التي برزت كحلول قوية لقدرتها على كشف العناصر بدقة عالية وبسرعة كبيرة [3]. ومع ذلك، أوضحت دراسات حديثة مثل دراسة Mondal وآخرين (2024) [3]، أن هذه النماذج تواجه تحديات واضحة عند محاولة اكتشاف العناصر الكبيرة مثل صناديق الإجابة، وذلك بسبب التقسيم الشبكي للصورة واعتماد هذه النماذج بشكل أساسي على التنبؤات المحلية لكل شبكة على حدة، مما يجعلها أقل فعالية في الحالات التي تتطلب معالجة سياق بصري واسع النطاق [3، 8].

لهذا السبب، برزت نماذج التقسيم الدلالي (Semantic Segmentation) المستندة إلى المحولات (Transformers)، كبديل فعال وواعد في هذا المجال [6، 9]. تستفيد هذه النماذج من آلية الانتباه الذاتي (Self-Attention)، التي تتيح لها فهم العلاقات السياقية بين الأجزاء المختلفة للصورة، مما يزيد من قدرتها على تحديد المناطق ذات الاهتمام مهما كانت كبيرة أو صغيرة، مع الحفاظ على كفاءة حسابية عالية [6]. ومن أبرز هذه النماذج نموذج SegFormer، الذي أظهر نتائج متقدمة في مهام التقسيم الدلالي في تطبيقات متنوعة كتصنيف الآفات الجلدية [5]، وتصنيف التعبيرات الوجهية [2]، وكذلك في تطبيقات رؤية الحاسوب العامة [6، 8].

كما بيّنت دراسة Khan وآخرين (2022) [6] أن النماذج المبنية على المحولات توفر قدرات فائقة على استيعاب العلاقات بعيدة المدى بين عناصر الصورة، ما يمنحها تفوقاً واضحاً مقارنة بالشبكات الترشحية التقليدية (CNN)، التي تعتمد بشكل رئيسي على العمليات المحلية والمحدودة [8]. من ناحية أخرى، أثبت نموذج Vision Transformer (ViT) الذي قدمه Dosovitskiy

وآخرون (2020) [9]، قدرة متميزة في مهام التصنيف البصري المعقدة، ما شجع على توظيف تقنيات مشابهة في مجال OMR لتعزيز قدرة الأنظمة على التعامل مع البيئات التعليمية المعقدة. إضافة إلى تقنيات الكشف والتقسيم، اهتم الباحثون أيضاً بأساليب متقدمة في معالجة النتائج وتقييم دقتها، وذلك باستخدام مصفوفات الارتباك (Confusion Matrix) ومقاييس تقييم متقدمة مثل Precision، Recall، F1-Score و IoU [4، 15]. وقد أكدت هذه الدراسات أهمية اختيار مقاييس التقييم الصحيحة لتحديد مدى قدرة الأنظمة على الاستجابة لمتطلبات الواقع التعليمي بشكل دقيق [4، 15].

بناءً على ما سبق، تتضح أهمية الجمع بين تقنيات التعلم العميق المتقدمة مثل نماذج SegFormer، وتقنيات المعالجة الإحصائية للصورة والإسقاط الأفقي، من أجل تطوير أنظمة قادرة على تحقيق دقة عالية جداً في استخراج صناديق الإجابة وفقاعات الخيارات من أوراق الإجابة متعددة الخيارات، مع تحقيق أقصى قدر ممكن من المرونة والتكيف مع الظروف الواقعية التي تواجه هذه التطبيقات.

4. مواد وطرائق البحث

يستعرض هذا القسم توصيف مجموعة البيانات والنموذج والخوارزميات المستخدمة ومقاييس التقييم المعتمدة وتفاصيل بيئة التجربة.

1.4. توصيف مجموعة بيانات البحث Dataset

بدأ العمل بجمع أوراق امتحانية حقيقية من جامعة حمص، شملت نماذج متعددة لمقررات واختصاصات من السنوات الأولى وحتى الخامسة. كان الهدف من تنويع البيانات هو محاكاة سيناريوهات حقيقية متنوعة تعكس الفوارق في التنسيق الطباعي وعدد الأسئلة وتصميم شبكة الفقاعات وحالات التظليل 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 عدة عناصر بنيوية بالإضافة إلى رئيسي وهو صندوق الإجابة (answer_box) حيث يمثل صناديق الإجابة من كل ورقة، يحوي كل صندوق على 50 مستطيل يمثل فقاعات الأسئلة الامتحانية لاستخدامها لاحقاً في عينات التدريب والتحقق.

2.4. النماذج والخوارزميات المستخدمة

تضمنت منهجية البحث اعتماد مجموعة من النماذج والخوارزميات الحديثة والمتكاملة بهدف تحقيق دقة عالية في عملية استخراج وتحليل صناديق الإجابة وفقاعات الخيارات من أوراق الإجابة متعددة الخيارات. تم تقسيم العمل إلى مراحل متسلسلة، حيث يؤدي كل نموذج أو خوارزمية دوراً أساسياً ضمن إطار النظام الكلي. وفيما يلي توضيح للنماذج والخوارزميات المعتمدة:

1.2.4 نموذج SegFormer للتقسيم الدلالي

اعتمدت الدراسة على نموذج SegFormer-B0، وهو أحد نماذج التقسيم الدلالي (Semantic Segmentation) القائمة على بنية المحولات (Transformers). يتميز النموذج بقدرته على استخلاص مناطق الاهتمام في الصور، وتحديد صناديق الإجابة بدقة عالية. تم تدريب النموذج على صور بحجم 1024×1024 بكسل باستخدام خوارزمية AdamW ودالة خسارة CrossEntropyLoss، واعتمد إخراج النموذج على أقنعة ثنائية الفئة (صندوق الإجابة/خلفية). يعد نموذج SegFormer-B0 من النماذج الحديثة والفعالة في مجال التقسيم الدلالي للصور (Semantic Segmentation)، حيث يعتمد على بنية المحولات (Transformers) بدلاً من الشبكات العصبية الترشيحية التقليدية (CNN). يمتاز SegFormer بكونه يجمع بين الدقة العالية في تحديد حدود الكائنات داخل الصور وبين كفاءته الحسابية، ما يجعله مناسباً للتطبيقات العملية واسعة النطاق مثل استخراج مناطق الإجابة من أوراق الامتحانات.

1.1.2.4 بنية النموذج

يعتمد SegFormer على تصميم هرمي متعدد المستويات، حيث يتم استخراج التمثيلات المميزة من الصورة عبر مراحل متعددة. تستخدم كل مرحلة آلية Attention مخصصة، ما يسمح للنموذج بفهم العلاقات السياقية بين الأجزاء المختلفة من الصورة على المدى البعيد والقصير معاً. على عكس شبكات CNN التي تعتمد على نطاق استقبال محلي محدود

(receptive field)، تتيح بنية Transformers للنموذج استيعاب بنية الصورة بالكامل وفهم حدود الكائنات الكبيرة والصغيرة بنفس الكفاءة [6، 7].

2.1.2.4 مميزات SegFormer

دقة عالية في التقسيم: يحقق النموذج أداءً متفوقاً في تحديد مناطق الاهتمام حتى في ظل اختلاف الإضاءة أو جودة الطباعة.

كفاءة حسابية: تم تصميم النموذج ليكون خفيفاً من حيث عدد المعاملات الحسابية (Parameters)، مما يسمح بتدريبه وتشغيله بكفاءة حتى على موارد عتادية محدودة.

مرونة مع اختلاف أبعاد الكائنات: قادر على معالجة واكتشاف العناصر البنيوية الكبيرة مثل answer box والعناصر الصغيرة بنفس الدقة، وهو ما يمثل نقطة قوة مقارنة بنماذج YOLO أو نماذج CNN التقليدية [6، 7].

قابلية التعميم: أظهر النموذج قدرة على التكيف مع صور متنوعة المصدر، سواء كانت ممسوحة ضوئياً أو ملتقطة بكاميرا محمولة، دون الحاجة لتعديلات جوهرية على البنية.

3.1.2.4 آلية الاستخدام في البحث:

تم تدريب نموذج SegFormer-B0 على مجموعة بيانات موسومة يدوياً، حيث اعتبر كل بكسل في الصورة إما ينتمي إلى فئة "صندوق الإجابة" أو "الخلفية". تم توحيد حجم جميع الصور إلى 1024×1024 بكسل، واستخدمت خوارزمية AdamW كخوارزمية تحسين، بينما تم اعتماد دالة CrossEntropyLoss لاحتساب الخسارة.

بعد التدريب، تمكن النموذج من توليد قناع دلالي ثنائي الفئة بدقة شبه تامة، بحيث تمثل المنطقة المصنفة كصندوق إجابة أساساً لعمليات التقسيم والتحليل اللاحقة.

4.1.2.4 سبب اختيار النموذج

تم اختيار SegFormer تحديداً لما يوفره من توازن بين الأداء والكفاءة الحسابية، ولأن بنيته المبنية على المحولات تمنحه ميزة استثنائية في استيعاب السياق العام للصورة، الأمر الذي يتفوق فيه على نماذج مثل YOLO التي غالباً ما تقشل في اكتشاف أو تحديد العناصر الكبيرة التي تغطي مساحات واسعة من الصورة. وبذلك شكل SegFormer خياراً مثالياً

كمرحلة ثانية بعد استخراج العناصر البنيوية في منظومة معالجة وتصحيح أوراق الإجابة متعددة الخيارات.

2.2.4 خوارزمية الإسقاط الأفقي والتقسيم الصفي

تم تطوير خوارزمية متكيفة تعتمد على الإسقاط الأفقي (Horizontal Projection) وتحليل القناة الحمراء (Red Channel Analysis) لتحديد وفصل صفوف الأسئلة داخل صندوق الإجابة. تعتمد هذه الخوارزمية على حساب إسقاط كثافة البكسلات عمودياً، وتطبيق عتبات ديناميكية لتحديد مواضع الصفوف بدقة، مع دعم آلية احتياطية (Fallback) لضمان استمرارية الأداء في حال عدم وضوح المؤشرات.

تعتمد خوارزمية الإسقاط الأفقي والتقسيم الصفي على سلسلة خطوات دقيقة لضمان استخراج المستطيلات على شكل صناديق يمثل كل صندوق منها سؤال يحوي على فقاعات الإجابة مبينة في الخوارزمية 1:

الخوارزمية 1 - خوارزمية الإسقاط الأفقي والتقسيم الصقي

— المدخلات: صورة صندوق الإجابة (answer box) بعد القص والمعالجة المبدئية.
— المخرجات: صور منفصلة لكل سطر أسئلة Q01.png إلى Q50.png

1. بداية الخوارزمية
2. تحسين الصورة (Preprocessing)
 - تحويل الصورة إلى تدرج رمادي.
 - تطبيق CLAHE لتحسين التباين.
 - تطبيق Gaussian Blur لإزالة التشويش.
 - تطبيق Adaptive Thresholding لاستخراج حدود الفقاعات وتوضيح التباين.
3. إسقاط الخطوط (Horizontal Projection)
 - جمع القيم عبر كل صف (axis=1) لإسقاط الصورة إلى بُعد واحد.
 - توليد ملف إسقاط أفقي يوضح توزيع الكثافة السوداء (intensity) في الصفوف.
4. اكتشاف الخطوط الأفقية الفاصلة (Dynamic Thresholding)
 - حساب العتبة الديناميكية:
 - $\text{threshold} = \text{mean} + 0.8 * \text{std}$ للإسقاط.
 - تم أخذ العتبة 0.8 تجريبياً لأنه يجب أن يحقق الحد الفاصل بين كل سؤال والذي يليه أكثر من 80% من العرض الكلي الخاص بكل answer box.
 - تحديد المؤشرات ذات الكثافة العالية، والتي تمثل الفواصل بين الأسطر.
5. استخراج مستطيلات الأسطر (Row Region Detection)
 - دمج المؤشرات القريبة (بفارق أقل من 5 بيكسل) لتكوين حدود كل سطر. (y1, y2)
 - اعتبار كل زوج (y1, y2) يمثل بداية ونهاية سطر أسئلة.
6. معالجة الحالات الخاصة (Fallback Row Splitting)
 - إذا لم يتم اكتشاف 50 أو 51 سطراً بوضوح:
 - تفعيل الوضع الاحتياطي: (fallback) تقسيم الصورة إلى ارتفاعات متساوية بناءً على عدد الأسطر المتوقع.
7. تحسين تقطيع fallback باستخدام الحواف الحمراء (Red Edge Refinement)
 - في الوضع الاحتياطي فقط، استخدام Red Projection (فرق القناة الحمراء عن المتوسط) لتعديل y2 بدقة لكل سطر.
 - تحسين دقة نهاية كل سطر وتجنب التداخل.
8. تقطيع الأسطر وحفظها (Row Cropping and Saving)
 - قص كل سطر من الصورة الأصلية باستخدام الإحداثيات. (y1, y2)
 - حفظ كل سطر كصورة منفصلة بالترتيب. (Q01.png, Q02.png, ... Q50.png)
9. التحقق البصري (Debug Visualization)
 - عرض مراحل التقطيع (الصورة الأصلية، الرمادية، المعالجة، الإسقاط الأحمر) باستخدام matplotlib.
 - حفظ هذه النتائج للمرجعة اليدوية.
10. إرجاع النتائج
 - صور الأسطر المقسمة جاهزة للمعالجة اللاحقة.

3.2.4 خوارزمية اكتشاف الفقاعات DBSCAN

تبعد تقسيم الصندوق إلى صفوف، استُخدمت خوارزمية التجميع المكاني DBSCAN (Density-Based Spatial Clustering of Applications with Noise) لاكتشاف مواقع الفقاعات الفردية ضمن كل صف. تتيح هذه الخوارزمية تجميع النقاط القريبة وفصل الفقاعات المظلمة فعلياً عن الضوضاء أو العلامات غير المقصودة، مما يرفع من موثوقية النتائج في ظروف الطباعة أو التصوير غير المثالية.

تعتمد خوارزمية اكتشاف الفقاعات على سلسلة خطوات دقيقة لضمان استخراج فقاعات الإجابة على شكل صناديق يمثل كل صندوق منها فقاعة إجابة مبيّنة في الخوارزمية 2:

الخوارزمية 2 - خوارزمية اكتشاف الفقاعات DBSCAN

— المدخلات: صورة سطر سطر (بعد التقسيم الصف) —

○ تحديد المجموعة ذات أكبر مساحة محاطة كمجموعة مرشحة تمثل الفقاعة الفعلية في هذا العمود.

6. استخراج الفقاعة بدقة: (Precise Bubble Extraction)

○ حساب المستطيل المحيط (Bounding Box) حول المجموعة المختارة من نقاط DBSCAN.

○ قص صورة الفقاعة بدقة من العمود، وحفظها في ملف منفصل مع تسمية واضحة (A, B, C, D, E).

7. التحقق البصري: (Debug Visualization)

○ رسم مستطيل حول كل فقاعة مكتشفة على صورة السطر الأصلية مع وضع الحرف المقابل (A-E).

○ حفظ الصورة الناتجة ضمن مجلد Debug للمراجعة اليدوية والتدقيق البصري.

8. إرجاع النتائج:

○ صور الفقاعات الخمسة (A-E) جاهزة للمعالجة أو التصنيف اللاحق.

9. نهاية الخوارزمية

الصغيرة في الفقاعات.

4. اكتشاف الفقاعات (Contours Detection)

○ استخراج جميع الكنتورات (Contours) داخل كل عمود.

○ تم استخراج الكنتورات باستخدام خوارزمية DBSCAN الموضحة في الخطوة 5.

○ فلتر الكنتورات بناءً على معايير الحجم (Area) والنسبة البعدية (Aspect Ratio)، لاستبعاد الضجيج والعناصر غير المنطقية.

5. تجميع باستخدام DBSCAN (Bubble Clustering)

○ حساب مركز كل فقاعة (Center Point) في العمود.

○ تطبيق خوارزمية DBSCAN لتجميع المراكز القريبة ضمن نفس الفقاعة، دون الحاجة لتحديد عدد الفقاعات مسبقاً.

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 العميقة والبيانات بدقة عالية، تضمنت البيئة التقنية المستخدمة وحدة معالجة مركزية Intel Xeon سداسية النوى (6 cores) 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 بسعة 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 وتقييمه.

2.5. مقاييس تقييم الأداء

لقياس فعالية النماذج المدربة على كشف عناصر بنيوية في الورقة الامتحانية، تم اعتماد مجموعة من المعايير المعروفة في مجال الرؤية الحاسوبية والتي تُمثل 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 [19]:

- TP (True Positive): اكتشف 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 كمهمة رئيسية عند استخدام نماذج التعرف على الأشياء. يوضح الشكل 3 وجود صندوقين محاطين أحدهما باللون الأسود يمثل المربع المحيط بقاعدة الحقائق الأساسية للعنصر، والآخر باللون الأبيض يمثل ناتج تنبؤ نموذج الكشف، نلاحظ أنهما لا يتداخلان بشكل تام لذلك نحتاج 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) [19]:

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

الشكل 1 - علاقة التقاطع على الاجتماع (IoU) [19]

(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 [19,15]:

- $mAP@0.5$: يتم اعتبار التوقع صحيح عندما يكون تقاطع المساحة بين الصندوق المتوقع والحقيقي $(IoU) \geq 0.5$.
- $mAP@0.5:0.95$: 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) [19,15]:

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

حيث أن AP متوسط الدقة للفئة i، ونقيس 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 [15]:

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

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

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

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

3.5. 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 باستخدام نموذج مدرب مسبقاً) إلى ثلاث مجموعات وهي مجموعة التدريب Training (70%) من مجموعة البيانات الكلية) ومجموعة التحقق Validation (15%) ومجموعة الاختبار Testing (15%).

4.5. الضبط الدقيق لنموذج SegFormer

لتحقيق تم ضبط نموذج SegFormer-B0 بعناية لضمان أداء عالٍ في مهمة التقسيم الدلالي لصناديق الإجابة. يوضح الجدول التالي أهم إعدادات ومعايير التدريب، مع شرح موجز لكل منها:

الجدول 1 - إعدادات تدريب نموذج SegFormer

| المكوّن / المعامل | القيمة | الشرح |
|-------------------------|--------------------------------|---|
| النموذج المستخدم | SegFormer-B0 (Transformers) | نموذج تقسيم دلالي حديث يعتمد على بنية المحولات (Transformers) |
| نوع التعلم | Semantic Segmentation | تقسيم الصورة إلى مناطق دلالية (صندوق إجابة/خلفية) |
| عدد الفئات | 2 | 0 خلفية، 1 صندوق إجابة |
| حجم الصورة | 1024 × 1024 | توحيد أبعاد جميع الصور لتسهيل التدريب والمعالجة |
| نوع الإخراج | قناع ثنائي الفئة | صورة لكل بكسل تحدد ما إذا كان ينتمي للصندوق أم الخلفية |
| المحسن | AdamW | إصدار محسن من Adam مناسب لنماذج Transformers |
| دالة الخسارة | CrossEntropyLoss | تقيس الخطأ في تصنيف كل بكسل داخل القناع |
| عدد الدورات (Epochs) | 50 | عدد مرات مرور النموذج على مجموعة البيانات كاملة |
| حجم الدفعة (Batch Size) | 8 | عدد الصور التي تُستخدم في كل خطوة تدريبية واحدة |

5.5. نتائج أداء نموذج SegFormer المدرب

تم تقييم أداء النموذج النهائي SegFormer المدرب على مجموعة اختبار تتألف من 30 صورة تنوعت من حيث جودة التصوير ووجود الميلاق.

1.5.5 الأداء الإجمالي للنموذج

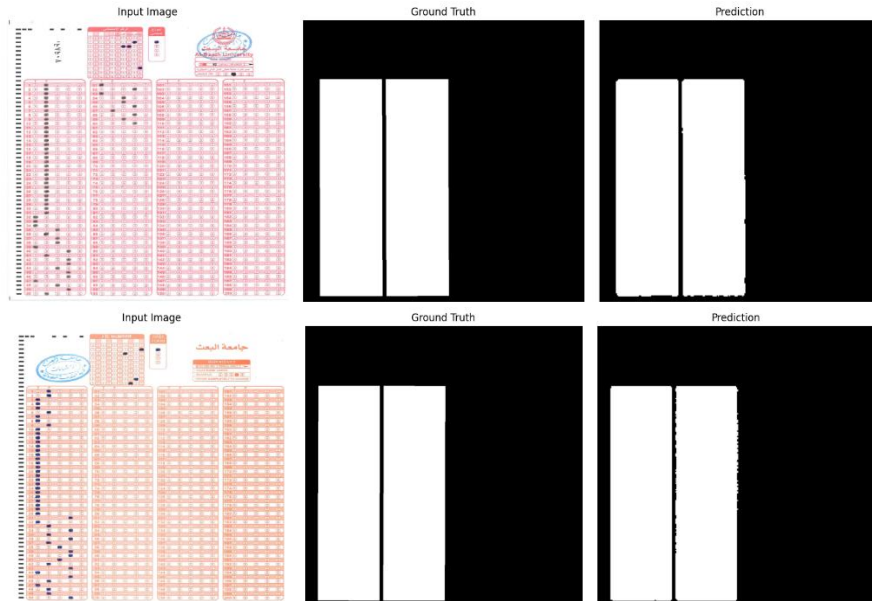
أظهرت النتائج أداء مميز في جميع المقاييس كما هو موضح في الجدول 5 أدناه:
الجدول 2 - نتائج اختبار الأداء الإجمالي SegFormer على 300 ورقة

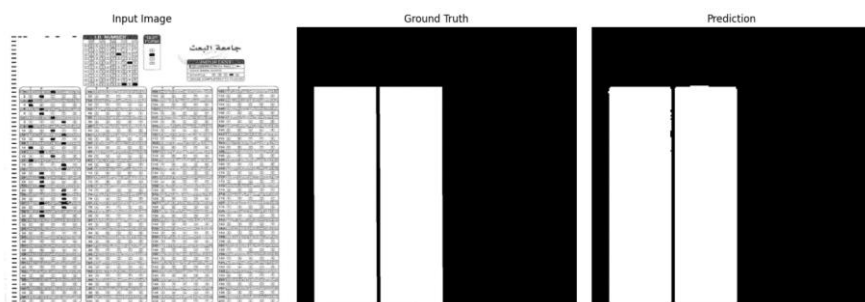
| المقياس | القيمة |
|-----------|--------|
| Accuracy | 0.9977 |
| Precision | 0.9970 |

استخراج فقاعات الإجابة من أوراق الإجابة متعددة الخيارات باستخدام نموذج SegFormer وتقنيات تحليل الصورة

| المقياس | القيمة |
|---------------|--------|
| Recall | 0.9961 |
| F1 Score | 0.9965 |
| IoU (Jaccard) | 0.9931 |

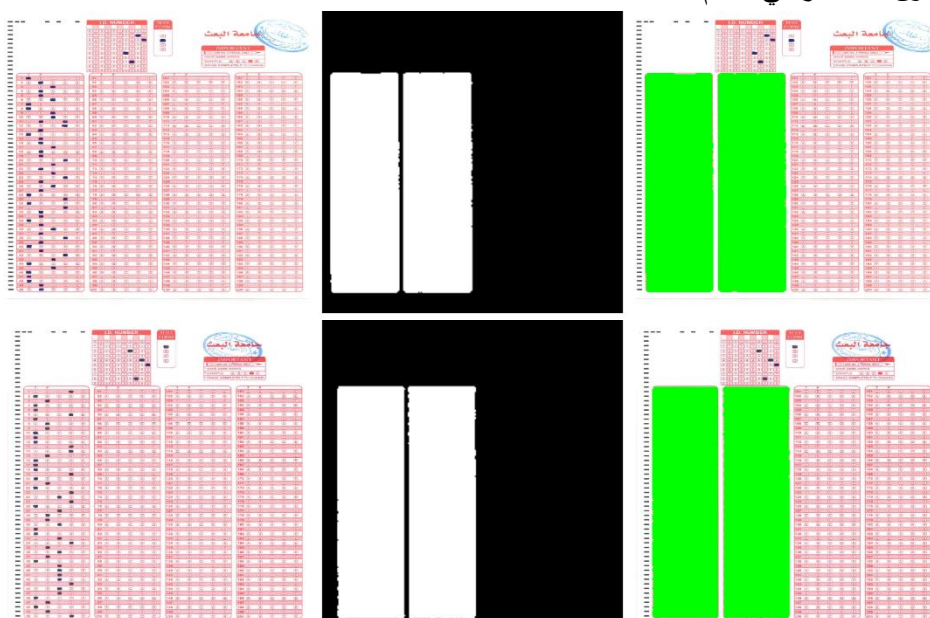
توضح الصور التالية المبينة في الشكل 2 نتائج الكشف على ثلاث صور بحيث يبين الشكل الصورة الاصلية وبجانبها القناع الخاص اضمن تسمية البيانات وبجانبه القناع الناتج من توقع النموذج SegFormer، وأظهرت الصور المخرجة دقة تحديد ممتازة لصندوق الإجابة في كل ورقة:



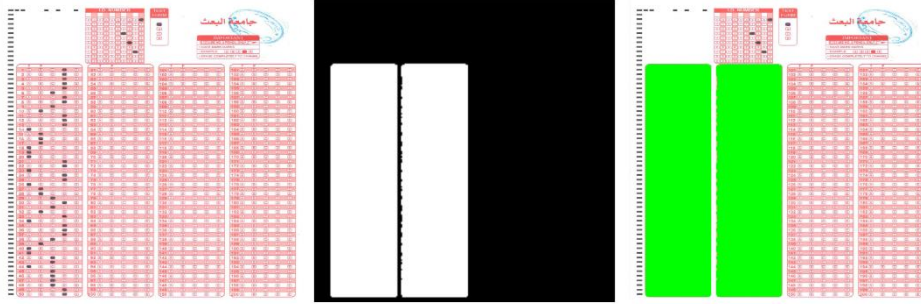


الشكل 2 - صور توضيحية لنتائج اختبار أداء نموذج الكشف SegFormer

وتوضح الصور التالية المبينة في الشكل 3 نتائج الكشف على ثلاث صور بحيث يبين الشكل الصورة الاصلية وبجانبها القناع الذي نتج عن النموذج بالإضافة الى الجزء المقطع من الصورة باللون الأخضر في القسم الثالث:



استخراج فقاعات الإجابة من أوراق الإجابة متعددة الخيارات باستخدام نموذج SegFormer وتقنيات تحليل الصورة



الشكل 3 - صور توضيحية لنتائج اختبار أداء النموذج على الصور ويظهر القسم المقطع بالأخضر

وتوضح الصور التالية المبينة في الشكل 4 والذي يظهر أربع صناديق إجابة مقطوعة من اثنتين من الصور الأصلية:

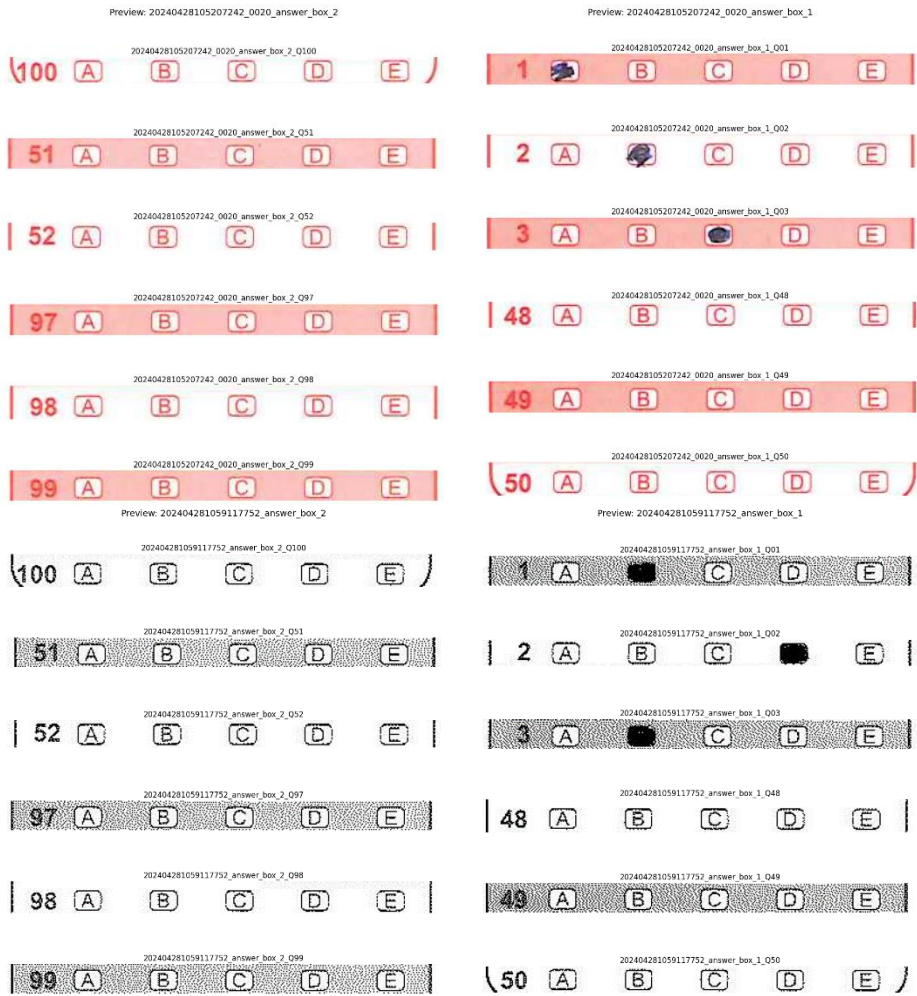
الشكل 4 - صور توضيحية لنتائج صناديق الإجابة المقطوعة من الصور الأصلية

6.5. نتائج تطبيق خوارزمية الإسقاط الأفقي والتقسيم الصفّي

أظهرت خوارزمية الإسقاط الأفقي والتقسيم الصفّي فعالية عالية في فصل صندوق الإجابة إلى صفوف أسئلة بشكل دقيق وموثوق. نجحت الخوارزمية في تحديد معظم الحدود الفاصلة بين الصفوف حتى في حالات اختلاف الطباعة أو وجود ضجيج بصري، كما أن آلية fallback تضمنت استمرار عملية التقطيع بشكل صحيح في الحالات الصعبة. أدى ذلك إلى تحقيق معدل دقة مرتفع في استخراج صفوف الأسئلة بشكل مستقل، مما سهل لاحقاً عمليات اكتشاف الفقاعات وتصنيفها، وساهم في تحسين جودة البيانات المدخلة للمراحل اللاحقة من النظام.

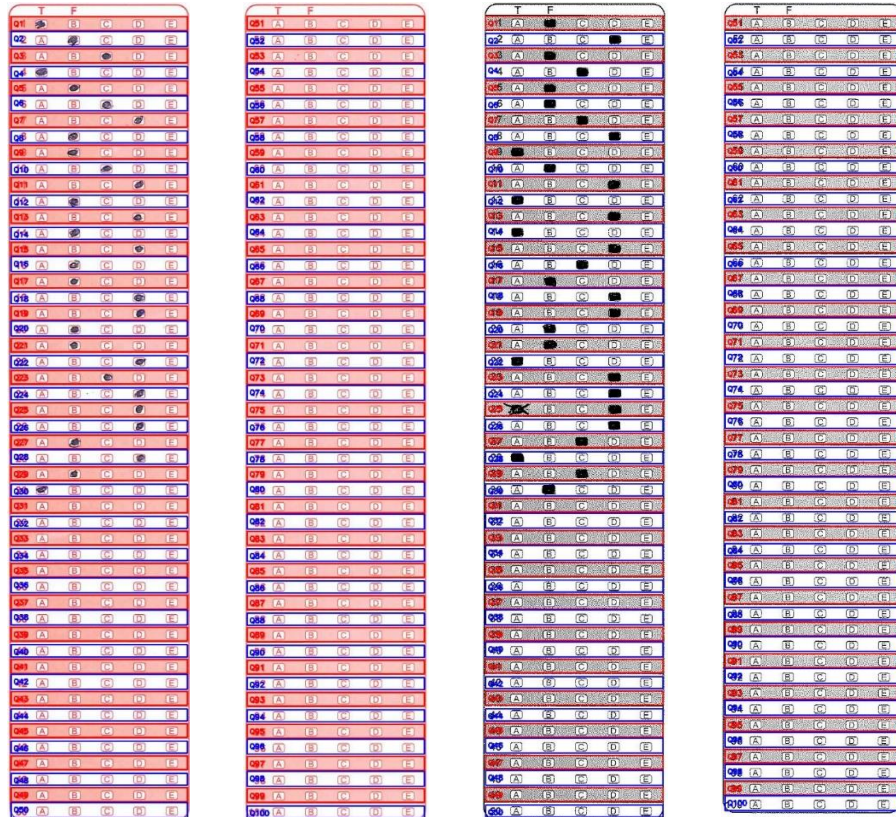
توضح المخرجات المعروضة في الشكل 5 إظهار اقتطاع أول 3 صناديق أسئلة وآخر 3 صناديق أسئلة من صور صناديق الإجابة الخاصة بعدة أوراق امتحانية

استخراج فقاعات الإجابة من أوراق الإجابة متعددة الخيارات باستخدام نموذج SegFormer وتقنيات تحليل الصورة



الشكل 5 - صور توضيحية لنتائج تطبيق خوارزمية الإسقاط الأفقي والتقسيم الصفّي

توضح المخرجات المعروضة في الشكل 6 إظهار الصناديق الخاصة بكل الأسئلة مرسومة بشكل كامل على صور صناديق الإجابة الخاصة بعدة أوراق امتحانية



الشكل 6 - صور توضيحية لنتائج تطبيق خوارزمية الإسقاط الأفقي والتقسيم الصفّي على الصور الخاصة بصناديق الإجابة كاملة

7.5. نتائج تطبيق خوارزمية اكتشاف الفقاعات DBSCAN

أثبتت خوارزمية DBSCAN فعالية كبيرة في اكتشاف وتحديد مواقع الفقاعات داخل كل صف من صفوف الإجابة، حتى في ظل وجود تداخل أو تشويش ناتج عن اختلاف الطباعة أو جودة المسح الضوئي. نجحت الخوارزمية في تجميع المراكز الحقيقية للفقاعات وفصلها عن العلامات الوهمية أو الضجيج، ما أدى إلى رفع دقة النقطيع واستخلاص الفقاعات بشكل مستقل وواضح. كما ساعدت في ضبط حدود الفقاعات بدقة عالية بغض النظر عن شكلها أو تباين الإضاءة داخل

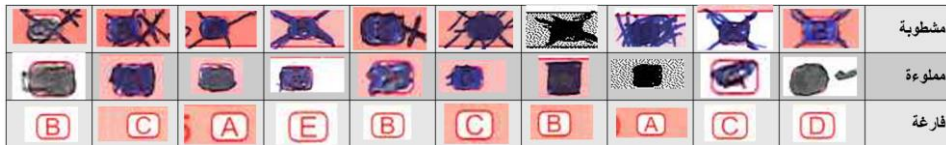
استخراج فقاعات الإجابة من أوراق الإجابة متعددة الخيارات باستخدام نموذج SegFormer وتقنيات تحليل الصورة

العمود، وساهمت في إنتاج صور فقاعات جاهزة لمرحلة التصنيف الآلي بدقة وموثوقية عالية، كما هو مبين في الصورة المرفقة في الشكل 7.



الشكل 7 - صور توضيحية لنتائج تطبيق خوارزمية اكتشاف الفقاعات DBSCAN

وبين الشكل 8 صور الفقاعات الناتجة التي سيتم تمريرها لاحقاً إلى المصنف الذي سوف يصنف كل صورة فقاعة ضمن 3 صفوف (مشطوبة - مملوءة - فارغة)



الشكل 8 - صور للفقاعات الناتجة ضمن 3 صفوف

6. الاستنتاجات والتوصيات

أظهرت نتائج هذه الدراسة أن توظيف النماذج الحديثة للتقسيم الدلالي، وتحديدًا نموذج SegFormer-B0، يشكل نقلة نوعية في معالجة أوراق الإجابة متعددة الخيارات، خصوصاً في المهام المتعلقة باستخراج صناديق الإجابة بدقة عالية بغض النظر عن اختلاف جودة الطباعة أو ظروف التصوير. لقد بينت التجارب أن استخدام SegFormer يمكن النظام من تحقيق عزل شبه تام لصندوق الإجابة، مما وفر أساساً قوياً ودقيقاً لبقية مراحل التحليل والتقطيع اللاحقة.

كما أثبتت خوارزمية الإسقاط الأفقي والتقسيم الصفّي فاعليتها في تقسيم صندوق الإجابة إلى صفوف متجانسة حتى في حالات وجود ضجيج أو عدم وضوح الحدود الطباعية، مستفيدة من

آليات التكيف والحسابات الإحصائية الديناميكية. وعززت آلية fallback من مرونة النظام، حيث ضمنت استمرارية التقطيع وجودته في الحالات الاستثنائية، مما ساعد على توليد بيانات منظمة وموثوقة لمرحلة اكتشاف الفقاعات.

أما على مستوى اكتشاف الفقاعات، فقد أظهرت خوارزمية DBSCAN قدرة كبيرة على تجميع مراكز الفقاعات الحقيقية وفصلها عن الضجيج أو العلامات غير المقصودة، حتى في الصور منخفضة الجودة أو ذات التباين الضوئي الحاد. وكنتيجة لذلك، أمكن استخراج الفقاعات الفردية بشكل دقيق، مما مهد لمرحلة التصنيف الآلي دون الحاجة لتدخل يدوي أو تصحيحات لاحقة.

انطلاقاً من هذه النتائج، توصي الدراسة بما يلي:

- اعتماد نهج التقسيم الدلالي المدعوم بالمحولات مثل SegFormer كخيار أولي في مهام فصل العناصر البنيوية الكبيرة داخل أوراق الإجابة، نظراً لمرونته وقابليته العالية للتعميم على بيانات جديدة.
- الاستمرار في تطوير خوارزميات الإسقاط والتقسيم التكيفي مع دمج مؤشرات إضافية من القنوات اللونية أو أساليب تعزيز التباين لضمان دقة أكبر في أصعب السيناريوهات.
- توظيف خوارزميات التجميع الذكي مثل DBSCAN كأداة موثوقة في فصل الفقاعات، خصوصاً في الأوراق التي تعاني من ضجيج بصري أو تظليل غير منتظم.
- توسيع نطاق التقييم ليشمل مجموعات بيانات أكبر ومتنوعة تغطي مختلف ظروف التصوير والطباعة، من أجل تعزيز ثقة المجتمع البحثي بتعميم النظام واقتراحه كحل قياسي في المؤسسات التعليمية.
- دعم مخرجات النظام بواجهات تحقق بصري (Debug Visualization) لتتيح للمستخدم النهائي مراجعة النتائج وتدقيقها بسهولة، ما يرفع من موثوقية وأمان النظام في بيئة العمل الفعلي.

7. المراجع:

- [1] Mai, J., Gao, C., & Bao, J. (2025). Domain generalization through data augmentation: A survey of methods, applications, and challenges. *Mathematics*, 13(5), 824.
- [2] Shen, Z. (2024). A comparative study of hybrid CNN and Vision Transformer models for facial emotion recognition. *IEEE International Conference on Dependable Systems and Their Applications (DSA)*, 401–408.
- [3] Mondal, S., De, P., Malakar, S., & Sarkar, R. (2024). OMRNet: A lightweight deep learning model for optical mark recognition. *Multimedia Tools and Applications*, 83(5), 14011–14045.
- [4] Obi, J. C. (2023). A comparative study of several classification metrics and their performances on data. *World Journal of Advanced Engineering Technology and Sciences*, 8(1), 308–314.
- [5] Nie, Y., Sommella, P., Carratù, M., O'Nils, M., & Lundgren, J. (2022). A deep CNN transformer hybrid model for skin lesion classification of dermoscopic images using focal loss. *Diagnostics*, 13(1), 72.
- [6] Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. (2022). Transformers in vision: A survey. *ACM Computing Surveys (CSUR)*, 54(10s), 1–41.
- [7] DE Elias, E., Tasinaffo, P., & Hirata, R. (2021). Optical mark recognition: Advances, difficulties, and limitations. *SN Computer Science*, 2(5), 367.
- [8] Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., & Ghayvat, H. (2021). CNN variants for computer vision: History, architecture, application, challenges and future scope. *Electronics*, 10(20), 2470.
- [9] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., & Houshy, N. (2020). An image is worth 16×16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

- [10] Afifi, M., & Hussain, K. F. (2019). The achievement of higher flexibility in multiple-choice-based tests using image classification techniques. *International Journal on Document Analysis and Recognition (IJDAR)*, 22, 127–142.
- [11] Gorokhovatskyi, O. (2016). Neocognitron as a tool for optical marks recognition. *IEEE International Conference on Data Stream Mining & Processing (DSMP)*, 169–172.
- [12] Patel, R., Sanghavi, S., Gupta, D., & Raval, M. S. (2015). CheckIt–A low cost mobile OMR system. *IEEE TENCON Region 10 Conference*, 1–5.
- [13] Sanguansat, P. (2015). Robust and low-cost Optical Mark Recognition for automated data entry. *IEEE International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 1–5.
- [14] Parul, H. M., & Kaur, M. (2012). A novel optical mark recognition technique based on biogeography based optimization. *International Journal of Information Technology and Knowledge Management*, 5(2), 331–333.
- [15] Ting, K. M. (2011). Confusion Matrix. In Sammut, C., & Webb, G. I. (Eds.), *Encyclopedia of Machine Learning* (pp. 209–209). Springer, Boston, USA.
- [16] Zampirolli, F., Gonzalez, J. A. Q., & Neves, R. D. O. (2010). Automatic correction of multiple-choice tests using digital cameras and image processing. *Universidade Federal do ABC, Brazil*.
- [17] Deng, H., Wang, F., & Liang, B. (2008). A low-cost OMR solution for educational applications. *IEEE International Symposium on Parallel and Distributed Processing with Applications*, 967–970.
- [18] Hussmann, S., & Deng, P. W. (2005). A high-speed optical mark reader hardware implementation at low cost using programmable logic. *Real-Time Imaging*, 11(1), 19–30.
- [19] 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.

