

تحليل وتحسين أداء نموذج YOLOv8 للكشف عن الكائنات باستخدام مجموعة بيانات BDD100K مخصصة

أ.د. يعرب ديوب*

د. جعفر سلمان**

م. كلودا إبراهيم***

(تاريخ الإيداع ٢٠٢٥/١١/٢٣ . قُبل للنشر في ٢٠٢٦/١/٧)

□ ملخص □

تهدف هذه الدراسة إلى تقييم أداء نموذج YOLOv8 في مهمة الكشف عن المركبات والعناصر المرورية ضمن بيانات حضرية معقدة، وذلك باستخدام نسخة مخصصة من قاعدة بيانات BDD100K تضم تسع فئات مرورية رئيسية. تم تدريب كل من YOLOv8n و YOLOv8m أولاً باستخدام الإعدادات القياسية، وقدمنا منهجية مقترحة لتحسين نموذج YOLOv8m عبر دمج تقنيات تعزيز بيانات متقدمة ورفع دقة الصور وتغيير خوارزمية التحسين إلى AdamW.

أظهرت النتائج تفوق نموذج YOLOv8m على YOLOv8n في جميع مؤشرات التقييم، حيث حقق $mAP50=0.5905$ مقابل 0.4219 ، و $Precision=0.6801$ مقابل 0.6288 . وبعد إجراء التعديلات وفق المنهجية المقترحة، حقق YOLOv8m تحسناً إضافياً في الأداء ليصل إلى $mAP50=0.6116$ و $Precision=0.7026$ ، مع انخفاض في قيم الخسارة على بيانات التحقق، مما يشير إلى زيادة ثبات النموذج وتحسن قدرته على التعميم. تعكس هذه النتائج أهمية تخصيص البيانات وتحسين استراتيجيات التدريب لرفع دقة نماذج الكشف في أنظمة النقل الذكية.

الكلمات المفتاحية: YOLOv8، الرؤية الحاسوبية، BDD100K، التعلم العميق، النقل الذكي، التعرف على النماذج.

* أستاذ في جامعة طرطوس - كلية هندسة تكنولوجيا المعلومات والاتصالات - قسم هندسة تكنولوجيا المعلومات.

** أستاذ مساعد في جامعة طرطوس - كلية هندسة تكنولوجيا المعلومات والاتصالات - قسم هندسة تكنولوجيا المعلومات.

*** طالبة دكتوراه في جامعة طرطوس - كلية هندسة تكنولوجيا المعلومات والاتصالات - قسم هندسة تكنولوجيا المعلومات.

Analysis and improve of the YOLOv8 object detection model performance using a custom BDD100K dataset

Prof. Yaroub Dayoub *

Dr. Jafar slman **

Eng. Kloda Ibrahim ***

(Received 23/11/2025 . Accepted 7/1/2026)

□ ABSTRACT □

This study aimed to evaluate the performance of the YOLOv8 model in detecting vehicles and traffic elements within complex urban environments, using a customized version of the BDD100K database containing nine major traffic categories. Both YOLOv8n and YOLOv8m were initially trained using standard settings. We then proposed a methodology to improve the YOLOv8m model by incorporating advanced data augmentation techniques, increasing image resolution, and changing the optimization algorithm to AdamW.

The results showed that the YOLOv8m model outperformed YOLOv8n in all evaluation indicators, achieving $mAP50 = 0.5905$ versus 0.4219 , and Precision = 0.6801 versus 0.6288 . After implementing the proposed methodology, YOLOv8m achieved further performance improvements, reaching $mAP50 = 0.6116$ and Precision = 0.7026 , along with a decrease in loss values on the validation data. This indicates increased model stability and improved generalizability. These results reflect the importance of data allocation and improving training strategies to increase the accuracy of detection models in intelligent transportation systems.

Keywords: YOLOv8, Computer Vision, BDD100K, Deep Learning, Intelligent Transportation, Pattern Recognition.

* Professor in Information and Communication Technology Engineering Faculty, Tartous University, Syria.

** Assistant Professor in Information and Communication Technology Engineering Faculty, Tartous University, Syria.

*** Phd student in Information and Communication Technology Engineering Faculty, Tartous University, Syria.

مقدمة

يشهد مجال أنظمة النقل الذكية تطورًا متسارعًا مدفوعًا بالحاجة إلى تحسين إدارة المرور، وتعزيز السلامة المرورية، وتسهيل التنقل في البيئات الحضرية المعقدة. ويُعد الكشف الدقيق عن المركبات والعناصر المرورية عنصرًا أساسيًا في العديد من التطبيقات الحساسة مثل المراقبة المرورية، التنبؤ بحالات الازدحام، القيادة الذاتية، وتحذير الاصطدامات [1]. وقد دفعت هذه الحاجة المتنامية إلى الاعتماد على تقنيات الرؤية الحاسوبية والتعلم العميق، لما تقدمه من قدرات عالية في تحليل الصور والتعامل مع التباين الكبير في الإضاءة والزوايا والأحجام [2].

تطورت خوارزميات الكشف عن الكائنات بشكل ملحوظ خلال العقد الأخير مدفوعة بالتقدم في بني الشبكات العصبية الالتفافية (CNNs)، وظهور وحدات معالجة رسومية عالية القدرة (GPUs) مكّنت من تدريب نماذج أعمق وأكثر تعقيدًا. يمكن تصنيف خوارزميات الكشف عن الكائنات تقليديًا إلى نماذج ذات طورين (Two-Stage Detectors) مثل Faster R-CNN، والتي تعتمد على توليد مقترحات أولية للكائنات قبل تصنيفها؛ ونماذج ذات طور واحد (One-Stage Detectors) مثل SSD وYOLO، والتي تتعامل مع الكشف والتصنيف ضمن عملية واحدة، ما يجعلها أسرع وأكثر ملاءمة للتطبيقات ذات الزمن الحقيقي [3]. وتشير الدراسات الحديثة إلى أن نماذج الطور الواحد، وبالأخص سلسلة YOLO، أصبحت الخيار الأكثر شيوعًا في تطبيقات النقل الذكي، نظرًا لقدرتها على الموازنة بين السرعة والدقة بكفاءة أعلى مقارنة بالطرائق ذات الطورين [4].

وفي هذا السياق، تمثل سلسلة YOLOv8 من أهم ما توصلت إليه نماذج الكشف ذات الطور الواحد، إذ توفر تحسينات في البنية الشبكية (Backbone) ورؤوس الكشف (Heads)، وتدعم آليات تدريب متقدمة تحقق أداءً أعلى على مختلف مجموعات البيانات القياسية. وبالنظر إلى التحديات الفريدة لبيانات المرور — مثل تعدد الكائنات، اختلاف الأحجام، الاحتجاب، والتنوع في ظروف الإضاءة — فإن تقييم قدرات هذه النماذج وتخصيصها لهذه المهمة يُعد ذا أهمية كبيرة [5].

ومن الدراسات المرجعية التي أُجريت في هذا المجال نذكر ما يلي:

١- دراسة في عام ٢٠٢٥ [6]، قدمت هذه الدراسة مقارنة بين أداء نموذج yolo8 ونموذج yolo5 على قاعدة البيانات BDD100K بعد تخصيصها بحيث تم استخدام محتوى الصور ضمن فئة "المركبات" (أي خمس فئات فقط هي السيارات والشاحنات والحافلات والدراجات النارية والدراجات الهوائية). اعتمدت الدراسة على رأس كشف خالٍ من المراسي Anchor-Free Detection Head وتصميم مُحسّن في العمود الفقري للنموذج (Backbone)، مما يساهم في زيادة جودة التمثيل البصري وتقليل التعقيد الحسابي. وأكدت الدراسة تفوق YOLOv8 على الإصدار السابق YOLOv5 عبر مقارنة مباشرة بينهما، إذ حقق النموذج الجديد زيادة ملحوظة في F1-score (٠,٨٧ مقابل ٠,٦٧)، وفي mAP50 (٠,٩١ مقابل ٠,٦٨٢)، وكذلك في mAP50-95 (٠,٧٤ مقابل ٠,٤٤٩).

٢- دراسة في عام ٢٠٢٤ [7]، قدمت هذه الدراسة خوارزمية YOLOv8m خفيفة الوزن للكشف عن كائنات الطرق، وذلك من خلال تصميم وحدة الالتفاف خفيفة الوزن C2fGSConv، ودمج رأس كشف خفيف الوزن بهدف تبسيط الخوارزمية وتحسين سرعة الكشف، متبوعًا بدمج آلية الانتباه الفعّالة للتعويض عن فقدان الدقة الناتج عن التصميم خفيف الوزن. تم تقييم النموذج المُحسّن على مجموعة بيانات BDD100K الخاصة

بكائنات الطرق، وأظهرت النتائج تحسناً واضحاً مقارنة بالإصدار الأصلي من YOLOv8m. حقق النموذج المحسن زيادة بمقدار ٥,٥% في الدقة (mAP50) عن الإصدار الأصلي من YOLOv8m حيث بلغت قيمته $mAP50=48.8$ ، وانخفاض في العمليات الحسابية بنسبة ٣١,٣٩% وتحسناً بنسبة ٢٧,٤% في سرعة الكشف. رغم التحسن البسيط في دقة الكشف عن الكائنات وتحسن سرعة الكشف تبقى الدقة منخفضة وتحتاج إلى تحسين للموازنة بين سرعة الكشف عن الكائنات ودقة اكتشافها.

٣- دراسة في عام ٢٠٢٣ [8]، ركزت هذه الدراسة على تطوير نموذج يستهدف التعرف الدقيق والسريع على إشارات المرور فقط. اعتمد الباحثون على نموذج YOLOv8 الذي تم تدريبه على مجموعة بيانات محلية مخصصة لإشارات المرور في الهند، وتم تصميم النموذج ليعمل في الزمن الحقيقي ضمن ظروف مرور معقدة. حقق النموذج أداءً مرتفعاً تمثل في دقة بلغت ٩٨,٣٨% و $mAP0.50$ وصل إلى ٩٨,٥٠، ما يُبرز فعالية YOLOv8 في مهام الكشف المروري ذات الحساسية العالية.

بالرغم من التقدم الملحوظ الذي كشفته هذه الدراسات في تطوير نماذج YOLO وتعزيز قدراتها، إلا أن كلاً منها ركز على جانب محدد من المشهد المروري دون تقديم تقييم شامل يجمع بين تعدد الكائنات، اختلاف أحجامها، وتنوع ظروف التصوير. فالدراسة الأولى تخصصت في المركبات فقط وتجاهلت الفئات الدقيقة والحساسة مثل المشاة والإشارات والعلامات الطرقية، في حين ركزت الدراسة الثانية على تحسين سرعة وكفاءة YOLOv8 دون معالجة كافية لمشكلة انخفاض الدقة. أما الدراسة الثالثة فقد اقتصر على إشارات المرور دون التعامل مع كامل تعقيد المشهد الحضري.

أهمية البحث وأهدافه

تبرز أهمية هذا البحث في كونه يجمع بين تخصيص قاعدة بيانات ضخمة مثل BDD100K لتتناسب مع تطبيقات المرور الحضري من جهة، وتقييم أداء نسختين مختلفتين من نموذج YOLOv8m و YOLOv8n من جهة أخرى، مما يتيح فهماً أعمق لعلاقة حجم النموذج ودقته مقابل متطلباته الحسابية. كما يتناول البحث دور عمليات تعزيز البيانات (Data Augmentation) في تحسين أداء النماذج وتقليل ظاهرة فرط التخصيص (Overfitting)، وهي من الجوانب الجوهرية في نجاح أنظمة الكشف في البيئات الحقيقية. إن النتائج المستخلصة من هذه الدراسة تساهم في توجيه الجهود نحو تطوير تطبيقات عملية أكثر كفاءة في مراقبة المرور في الزمن الحقيقي وتطوير أنظمة اتخاذ القرار الذكية ضمن بنى المدن الذكية.

وبناءً على ذلك، تهدف هذه الدراسة إلى تقييم أداء نموذجي YOLOv8m و YOLOv8n في مهمة الكشف عن المركبات والعناصر المرورية باستخدام نسخة مخصصة من قاعدة بيانات BDD100K، وتقديم منهجية مقترحة لتحسين أداء النموذج الأفضل (YOLOv8m) عبر تقنيات تدريب متقدمة. كما تهدف الدراسة إلى تحليل تأثير تحسينات البيانات والبنية على مؤشرات الأداء المختلفة، وتحديد مدى قابلية النموذج للتطوير والتخصيص بما يتناسب مع متطلبات أنظمة النقل الذكية.

طرائق البحث ومواده

1. خوارزمية الكشف بنظرة واحدة (YOLO): You Only Look Once

1.1 عائلة نماذج YOLO وتطورها

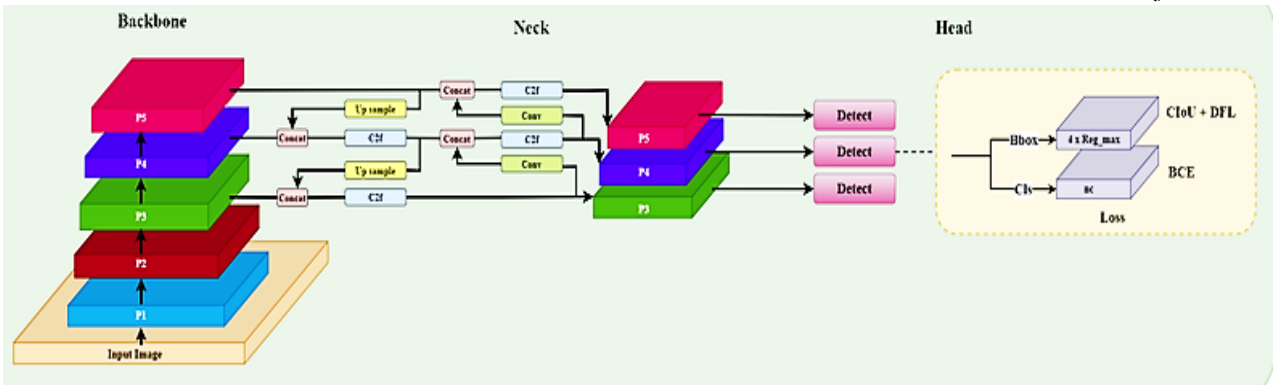
تُعدّ YOLO واحدة من أكثر خوارزميات الكشف عن الكائنات أحادية المرحلة شهرةً نظراً لقدرتها الفائقة على العمل في الزمن الحقيقي، حيث تعتمد على الشبكات العصبونية العميقة لاكتشاف وتحديد نوع الكائنات داخل الصور من خلال مرور واحد فقط بدلاً من تحليل المناطق المحتملة للكائنات كما هو الحال في Faster R-CNN، مما يوفر سرعة عالية جداً في معالجة الصور، حتى لو لم تكن الأكثر دقة من بين جميع خوارزميات الكشف. ظهرت أول نسخة من نموذج YOLO عام 2016، حيث قدمت فكرة ثورية تتمثل في تنفيذ الكشف عن الأجسام كمسألة انحدار واحدة (Single Regression Problem) بدلاً من مراحل متعددة كما في النماذج السابقة [4].

تطور النموذج عبر عدة إصدارات، بدءاً من YOLOv2 وYOLOv3 اللذين حسّنا الدقة باستخدام ميزات متعددة الطبقات، مروراً بـ YOLOv4 الذي أدخل تحسينات في التدريب والتعزيز البصري [10]، وصولاً إلى YOLOv5 الذي أصبح مرجعاً أساسياً لتطبيقات الرؤية في الوقت الحقيقي [3]. أما الإصدارات الأحدث مثل YOLOv7 [11] وYOLOv8 [8] فقد اعتمدت على تصاميم أكثر كفاءة واستقلالية عن المراسي (Anchor-free Design)، مع رؤوس كشف مفصولة (Decoupled Heads) لزيادة دقة التنبؤات.

يمثل YOLOv8 خطوة متقدمة في هذا التطور بفضل إعادة بناء بنيته المعمارية لتكون أكثر مرونة وقدرة على التعامل مع البيئات الديناميكية [8].

1.2 البنية المعمارية المتقدمة لنموذج YOLOv8

تمثل البنية المعمارية لـ YOLOv8 تطوراً نوعياً في مجال الرؤية الحاسوبية، حيث تجمع بين المبادئ الأساسية للكشف أحادي المرحلة وتقنيات التعلم العميق المتطورة. يعتمد النموذج على فلسفة تصميمية تعطي أولوية للكفاءة Computational Efficiency دون التضحية بالدقة، مما يجعله خياراً مثالياً للتطبيقات التي تتطلب معالجة في الوقت الفعلي مثل أنظمة المرور الذكية [8].



الشكل (1): بنية نموذج yolo8 [8]

يوضح الشكل (1) بنية نموذج yolo8 حيث تتكون بنية YOLOv8 من ثلاث وحدات أساسية مترابطة وهي:

1- الشبكة الأساسية أو العمود الفقري (Backbone)

تمثل الشبكة الأساسية قلب النموذج النابض، حيث تقوم باستخلاص الميزات Features من الصور المدخلة عبر سلسلة من الطبقات التلافيفية المتخصصة. في YOLOv8، تم استخدام تحسين متقدم على بنية CSPDarkNet المعروفة، والتي تعتمد على مبدأ التقسيم الجزئي عبر المراحل Cross Stage Partial. هذا المبدأ يعمل على تقسيم تدفق البيانات إلى مسارين متوازيين: أحدهما للمعالجة العميقة عبر سلسلة من التحويلات، والآخر للحفاظ على المعلومات الأصلية من خلال اتصال مباشر. هذه الآلية لا تقلل من التعقيد الحسابي فحسب، بل تحسن أيضاً من تدفق التدرجات خلال التدريب، مما يسهل عملية التعلم لميزات معقدة [8].

2- الرقبة (Neck)

بعد استخراج الميزات على مستويات مختلفة من التجريد، تأتي مرحلة دمج هذه الميزات فيما يعرف برقبة الهرم. يستخدم YOLOv8 بنية PANet المعززة (Path Aggregation Network)، التي تعمل على الجمع بين أفضل ما في العالمين: الدقة المكانية العالية من المستويات المبكرة والمحتوى الدلالي الغني من المستويات المتأخرة. تعتمد هذه البنية على مسارين متكاملين: مسار هابط ينقل الميزات الدلالية من المستويات العليا إلى المستويات الدنيا، ومسار صاعد ينقل الميزات المكانية في الاتجاه المعاكس. هذا التكامل يخلق تمثيلاً غنياً للميزات يحسن بشكل ملحوظ من قدرة النموذج على كشف الكائنات بمختلف أحجامها [8].

3- رأس الكشف (Head)

يمثل رأس الكشف المرحلة النهائية في عملية الكشف، حيث يتحول التمثيل المجرد للميزات إلى تنبؤات ملموسة للمربعات المحيطة بالكائنات. يتميز YOLOv8 باستخدام رأس كشف غير مرتبط بالمراسي Anchor-Free، وهو تحول جذري عن التصميمات التقليدية. أي أنه بدلاً من الاعتماد على مجموعة محددة مسبقاً من المراسي ذات الأحجام والنسب الثابتة، يتنبأ النموذج مباشرة بإحداثيات المربعات المحيطة بالنسبة لمراكز الخلايا. هذه الآلية تقلل من التعقيد الحسابي وتحسن الأداء مع الكائنات غير الاعتيادية، كما تزيل الحاجة لضبط معاملات المراسي التي كانت تشكل تحدياً في النماذج السابقة [8].

1.3 بنية نموذج YOLOv8m

نموذج YOLOv8m يمثل النسخة المتوسطة في عائلة YOLOv8، ويتميز بتوازن ممتاز بين الدقة والكفاءة الحسابية. يتكون النموذج من 92 طبقة و 25,8 مليون معلمة ويُنفذ نحو 78,7 مليار عملية عائمة (GFLOPs) لكل صورة، وهو ما يتيح له تحقيق دقة مرتفعة مع سرعة استجابة مناسبة للتطبيقات شبه الفورية [8].

يستفيد YOLOv8m من بنية CSPDarknet-C2f في العمود لتحسين استخراج السمات العميقة، ومن تكامل FPN-PAN في العنق لتعزيز دمج المعلومات عبر المستويات. أما رأس الكشف فيتميز بقدرته على التعامل مع الكائنات الصغيرة (مثل إشارات المرور) والكبيرة (مثل الشاحنات والحافلات) بكفاءة مقاربة. تُعد هذه البنية خياراً مناسباً للتطبيقات التي تتطلب أداءً دقيقاً ومستقرًا دون قيود صارمة على الطاقة الحسابية [12].

1.4 بنية نموذج YOLOv8n

نموذج YOLOv8n (Nano) هو النسخة الأخف وزناً في عائلة YOLOv8، صُمم ليعمل على الأجهزة محدودة الموارد مثل الأنظمة المدمجة أو وحدات الحوسبة الطرفية. يحتوي هذا النموذج على عدد أقل من الطبقات والمعاملات مقارنة بـ YOLOv8m، مما يقلل من التعقيد الحسابي واستهلاك الذاكرة بشكل كبير، على حساب انخفاض طفيف في الدقة. يُستخدم YOLOv8n عادةً في تطبيقات الوقت الحقيقي عالية السرعة، مثل كاميرات المراقبة الذكية أو الأنظمة المحمولة، حيث يُفضل زمن الاستجابة القصير على أقصى دقة ممكنة. ورغم خفة وزنه، فإنه يحتفظ بالبنية الأساسية لـ YOLOv8 من حيث التصميم CSP-C2f والرأس المفصول Anchor-free، مما يجعله نموذجًا فعالاً وقابلًا للتطبيق في سيناريوهات متعددة [12].

2. قاعدة البيانات BDD100K

تُعَدّ قاعدة بيانات BDD100K (Berkeley DeepDrive) واحدة من أكبر وأشمل مجموعات البيانات المستخدمة في أبحاث القيادة الذاتية والرؤية الحاسوبية. تم تطويرها في جامعة كاليفورنيا - بيركلي، وتضم ١٠٠,٠٠٠ صورة وفيديو جُمعت من بيانات قيادة حقيقية في مدن متعددة، تحت ظروف مختلفة من الطقس والإضاءة والتوقيت [9].

تتميز هذه القاعدة بتنوعها الكبير من حيث المشاهد والظروف البيئية، إذ تشمل لقطات نهائية ولييلية، ومشاهد تحت المطر أو الغيوم، مما يجعلها موردًا مثاليًا لتدريب النماذج التي يُفترض أن تعمل في البيئات الواقعية المعقدة. تحتوي قاعدة البيانات على مجموعة من الفئات التي تغطي معظم العناصر المرورية، مثل المركبات، المشاة، الدراجات، إشارات المرور، والعلامات الطرقية [9]. غير أن هذا التنوع الكبير قد يُشكل تحديًا في تطبيقات معينة مثل إدارة المرور الحضرية، حيث يُفضّل التركيز على الفئات الأكثر ارتباطًا بالمشهد المروري اليومي.

لذلك، تم في هذا البحث تخصيص القاعدة لتشمل فقط تسع فئات أساسية هي: (bike, bus, car, motor, person, rider, traffic light, traffic sign, truck) وذلك لتقليل التعقيد وتحسين أداء النموذج في الكشف عن الكائنات الأكثر أهمية في بيانات الطرق الحضرية. هذا التخصيص سمح بتحقيق توازن أفضل بين دقة الكشف وكفاءة التدريب، إذ تم تقليص عدد الفئات إلى تلك التي تؤثر مباشرة في تدفق المرور وسلامة المركبات والمشاة. كما ساعد هذا التخصيص على رفع جودة البيانات من خلال تنقية التسميات (Labels) غير الدقيقة أو الفئات غير ذات الصلة، مما زاد من موثوقية النموذج أثناء التقييم في سيناريوهات واقعية.

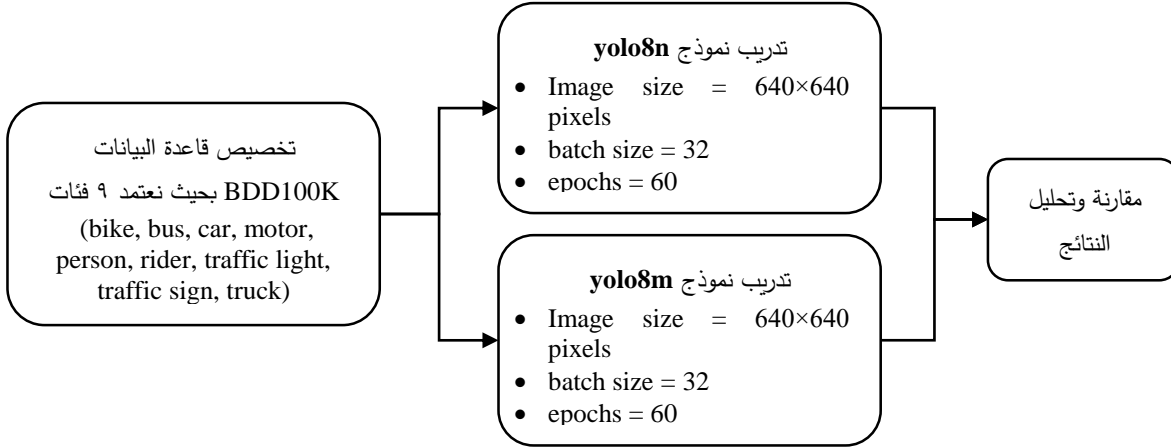
3. منهجية البحث

تم إجراء البحث وفق مجموعة من السيناريوهات:

١- السيناريو الأول:

يوضح الشكل (٢) الخطوات المتبعة في السيناريو الأول، حيث قمنا بتدريب كل من YOLOv8n و YOLOv8m (دون إجراء تعديلات) على النسخة المخصصة من قاعدة BDD100K التي تحتوي على تسع فئات مرورية، باستخدام الإعدادات الافتراضية في مكتبة Ultralytics YOLOv8 وذلك بهدف تحليل العلاقة بين حجم النموذج، الدقة، والكفاءة الحسابية في بيانات المرور الحضرية الواقعية. قمنا بتقسيم قاعدة البيانات إلى مجموعات تدريب (train) وتحقق (validation) وفق منهجية YOLO القياسية، حيث تم توحيد أبعاد الصور إلى 640×640 بكسل واختيار حجم دفعة (batch size=32) وتم تحديد عدد عصور التدريب (epochs) إلى ٦٠ حقبة.

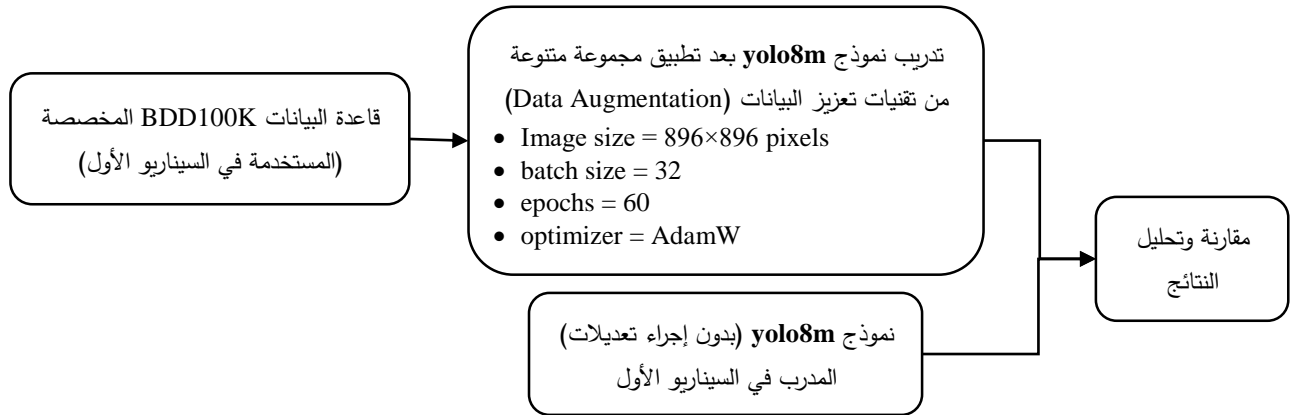
الشكل (٢): الخطوات المتبعة في السيناريو الأول



٢- السيناريو الثاني:

يوضح الشكل (٣) الخطوات المتبعة في السيناريو الثاني، حيث قدمنا منهجية مقترحة لإجراء تعديلات على نموذج yolo8m من أجل دراسة تأثيرها على أداء النموذج. تم زيادة حجم الصور ليصبح ٨٩٦×٨٩٦ بكسل، كما تم تحديد حقب التدريب ٦٠ حقة بالإضافة الى استخدام محسن (optimizer) AdamW بدلاً من SGD المستخدم بشكل افتراضي في السيناريو الأول من البحث. ومن أجل تحسين قدرة النموذج على التعميم ومواجهة التباينات في المشاهد الحضرية، تم تطبيق مجموعة متنوعة من تقنيات تعزيز البيانات (Data Augmentation). تشمل هذه العمليات كلاً من التحويلات الهندسية (Geometric) واللونية (Photometric)، بما يسمح محاكاة ظروف تصوير مختلفة مثل تغيير الإضاءة أو الحجب أو الحركة، وهي:

- (1.0) Mosaic: دمج أربع صور في صورة واحدة لتحسين كشف الأجسام الصغيرة والسياق العام.
- (0.2) MixUp و (0.2) CutMix: دمج صور متعددة جزئياً لتقليل فرط التخصص وزيادة تنوع العينات.
- RandAugment: تطبيق تحولات عشوائية تشمل السطوح والتباين والتدوير.
- Flip أفقي (fliplr = 0.5): لزيادة تنوع الاتجاهات في المشاهد.
- (0.5) Random Erasing: حذف مناطق صغيرة من الصورة لمحاكاة الحجب الجزئي.
- تحسين ألوان HSV: تعديلات في درجة اللون والتشبع والإضاءة (Saturation=0.7، Hue=0.015، Value=0.4).



الشكل (٣): الخطوات المتبعة في السيناريو الثاني

النتائج والمناقشة

تم تنفيذ عملية تدريب النماذج في بيئة Google Colab التي تتيح استخدام وحدات معالجة الرسومات NVIDIA Tesla T4 المزودة بذاكرة ١٦ GB. تم استخدام ما يقارب ٧٠٠٠٠٠ صورة للتدريب (training) و ١٠٠٠٠٠٠ صورة للتحقق (validation)، وهو ما يضمن تنوع العينات وتحقيق تمثيل متوازن لسيناريوهات الطرق المختلفة.

تم تقييم النماذج بناءً على مجموعة من المقاييس القياسية في مهام الكشف عن الكائنات والتي تعكس قدرة النموذج على تمييز الأجسام، وتحديد مواقعها بدقة، والتعامل مع التنوع في أحجام وأشكال المركبات وعناصر الطريق، وهي:

- ١- الدقة (Precision): تمثل الدقة النسبة بين عدد الكائنات الصحيحة المكتشفة إلى مجموع الكائنات التي تم التعرف عليها من قبل النموذج. بمعنى آخر، تقيس قدرة النموذج على تجنب الإيجابيات الكاذبة (False Positives).
 - ٢- الاستدعاء (Recall): يقيس الاستدعاء قدرة النموذج على اكتشاف جميع الكائنات الحقيقية في الصورة، أي تجنب الإيجابيات السالبة (False Negatives).
 - ٣- متوسط الدقة عند (mAP50): يعتبر من أهم المقاييس المعتمدة لتقييم أداء نماذج الكشف عن الأجسام، إذ يعبر عن متوسط الدقة عند تداخل بين الصندوق الحقيقي والتنبؤي بنسبة ٥٠%.
 - ٤- متوسط الدقة (mAP50-95): يُعد هذا المقياس أكثر صرامة، إذ يُحسب متوسط الأداء عبر عتبات تداخل متعددة من ٠,٥ إلى ٠,٩٥، مما يعطي صورة أكثر شمولية عن جودة التنبؤ بالمواقع والأحجام.
 - ٥- Box Loss: يشير إلى دقة النموذج في تحديد موقع الصندوق الحدي (Bounding Box) للكائن بالنسبة إلى موقعه الحقيقي.
 - ٦- CIs Loss: يقيس مدى صحة تصنيف الكائنات داخل الصناديق المكتشفة.
 - ٧- DFL Loss: دالة خسارة متقدمة تهدف إلى تحسين التموضع الدقيق لحواف الصندوق الحدي، إذ تعتمد على توزيع احتمالي لقيم الإحداثيات بدلاً من التنبؤ بقيمة واحدة.
- كما أشرنا سابقاً، تم إجراء البحث وفق مجموعة من السيناريوهات:

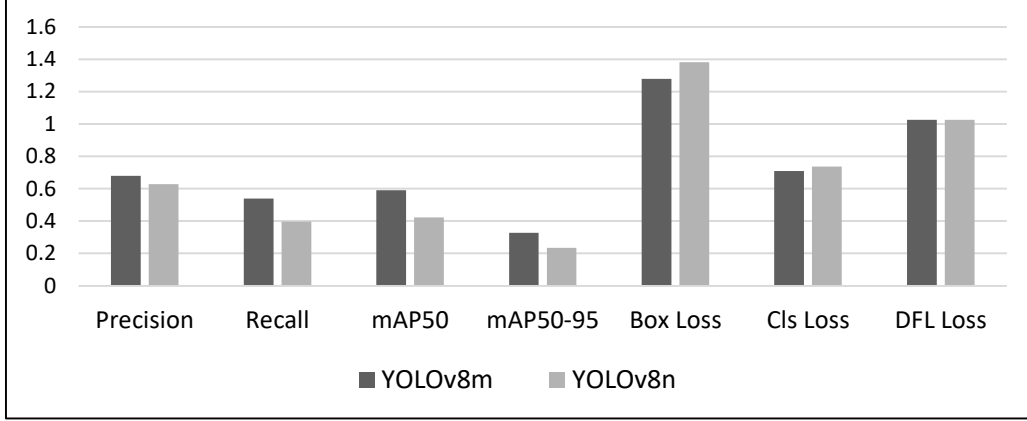
١- بالنسبة لنتائج السيناريو الأول:

يوضح الجدول (١) والشكل (٤) مقارنة بين نتائج تدريب كل من YOLOv8m و YOLOv8n على قاعدة البيانات BDD100K المخصصة وفق المقاييس القياسية المستخدمة في عملية تقييم الأداء في مهام الكشف عن الكائنات.

الجدول (١): مقارنة بين نتائج تدريب كل من نموذج yolov8m ونموذج yolov8n على قاعدة البيانات BDD100K المخصصة

المقاييس	YOLOv8m	YOLOv8n
Precision	0.6801	0.6288
Recall	0.5391	0.3972

0.4219	0.5905	mAP50
0.2340	0.3275	mAP50-95
1.3824	1.2799	Box Loss
0.7378	0.7089	Cls Loss
1.02538	1.02552	DFL Loss



الشكل (٤): مقارنة بين نتائج تدريب كل من نموذج yolov8m ونموذج yolov8n على قاعدة البيانات BDD100K المخصصة

تظهر النتائج أن نموذج YOLOv8m قد توصل إلى دقة بمقدار ٠,٦٨ مقارنة بـ ٠,٦٣ لنموذج YOLOv8n، مما يشير إلى أن YOLOv8m كان أكثر موثوقية في التعرف على الكائنات من الخلفية أو الكائنات غير المرغوبة. يُعزى هذا التحسن إلى البنية الأعمق للنموذج، التي تمكنه من استخراج ميزات أكثر تعقيداً ودقة من الصور ذات الخلفيات المزدحمة، وهي سمة جوهرية في مشاهد المرور الحضرية.

– أما بالنسبة للاستدعاء (Recall)، تحسن هذا المؤشر بشكل ملحوظ في YOLOv8m (0.54) مقارنة بـ YOLOv8n (0.40)، يعني ذلك أن YOLOv8m استطاع اكتشاف عدد أكبر من الكائنات الموجودة فعلاً في المشهد، وهو مؤشر على مرونته في التعامل مع الحالات الصعبة مثل الاحتجاب الجزئي واختلاف أحجام الأجسام داخل الصور. بالمقابل، يُظهر YOLOv8n محدودية نسبية في هذا الجانب بسبب بساطة بنيته التي تقلل من قدرته على استيعاب التنوع البصري العالي في بيانات المرور الواقعية.

– بالنسبة لمقياس متوسط الدقة (mAP50)، حقق YOLOv8m قيمة ٠,٥٩ مقابل ٠,٤٢٢ لـ YOLOv8n، هذا الفرق الكبير يعكس قدرة YOLOv8m على تحديد مواقع الأجسام بشكل أدق، حتى مع وجود تباين في الإضاءة أو الزوايا، ما يؤكد أن البنية الأعمق للنموذج مكنته من بناء تمثيل مكاني أكثر تماسكاً للكائنات.

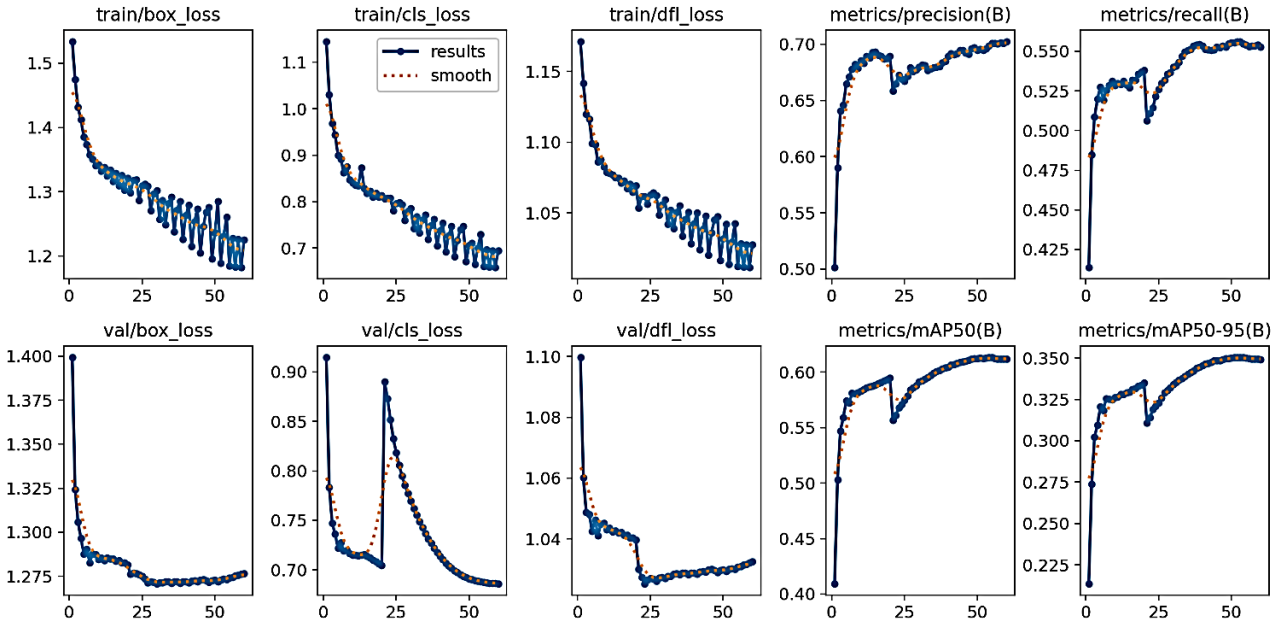
– وأخيراً، بالنسبة لمتوسط الدقة (mAP50-95)، فقد سجّل YOLOv8m قيمة ٠,٣٣ مقابل ٠,٢٣٤ لـ YOLOv8n، وهو الأعلى بين جميع المقاييس. يدل هذا الفارق الكبير على أن YOLOv8m ليس فقط أكثر دقة في تحديد وجود الكائن، بل أيضاً في رسم حدوده بدقة أعلى على امتداد نطاق واسع من المعايير. وبالتالي، يمكن اعتباره النموذج الأكثر موثوقية لتطبيقات النقل الذكي التي تتطلب تحديداً دقيقاً للموقع والشكل، مثل تقدير المسافة الآمنة بين المركبات أو اكتشاف المشاة.

– بالنسبة لتحليل قيم مقاييس الخسارة على بيانات التحقق (validation loss) لكل من النماذج، تظهر النتائج أن YOLOv8m حقق خسائر أقل في مكونات الصندوق والتصنيف، حيث كانت قيمة Box

Loss أقل في نموذج YOLOv8m مما يعني دقة تحديد أعلى للمواقع، أيضاً كانت قيمة Cls Loss في نموذج YOLOv8m مما يعني أنه يحقق تصنيف أكثر استقراراً. أما بالنسبة لتشابه قيم DFL Loss بين النموذجين فيعني أن مستوى تموضع الحدود البؤرية (Focal Localization) بقي قريباً، وهو متوقع لأن هذا المكون يعتمد أكثر على الخواص المشتركة للهيكل العام في سلسلة YOLOv8.

٢- بالنسبة لنتائج السيناريو الثاني:

يوضح الشكل (٥) نتائج تدريب نموذج YOLOv8m على قاعدة البيانات BDD100K بعد تطبيق المنهجية المقترحة التي شملت تحسينات على بيئة التدريب مثل تعزيز البيانات (Data Augmentation)، زيادة دقة الصور، وتغيير خوارزمية التحسين. حيث أظهر النموذج المعدل تحسناً إضافياً في مؤشرات الأداء، مع ثبات واستقرار أكبر في منحنيات التدريب.



الشكل (٥): نتائج نموذج YOLOv8m (بعد التعديل)

- تظهر النتائج أعلاه تطور مؤشرات التدريب والتحقق لنموذج YOLOv8m بعد التعديل عبر ٦٠ حقبة تدريب، حيث يمكن ملاحظة أن جميع المنحنيات تتبع نمط انحداري منتظم للخسارة مع استقرار تدريجي في مؤشرات الأداء، مما يعكس توازن عملية التعلم وتحسن النموذج في كل من التعميم والدقة.

- تظهر منحنيات الخسارة على بيانات التدريب (train/box_loss, train/cls_loss, train/df_l_loss) انخفاضاً واضحاً ومنتزحاً عبر حقب التدريب، مع ثبات نسبي بعد الحقبة ٤٥ تقريباً. حيث يشير انخفاض Box Loss من ١,٥ إلى نحو ١,٢ إلى تحسن دقة تحديد مواقع الصناديق الإطارية. كما يشير تراجع Cls Loss من ١,١ إلى أقل من ٠,٧ إلى أن النموذج أصبح أكثر تمييزاً بين الفئات التسع (مثل المركبات، الأشخاص، إشارات المرور). أما فقد DFL Loss انخفض بشكل ثابت إلى نحو ١,٠٣، مما يعكس تحسن تمثيل الحدود الدقيقة للأجسام داخل الإطارات.

- بالنسبة لمنحنيات الخسارة على بيانات التحقق (val/box_loss, val/cls_loss, val/df_l_loss)، فيمكن ملاحظة تطابق شبه مثالي بين منحنيات التدريب والتحقق، وهو مؤشر على انخفاض خطر الإفراط في التعلم.

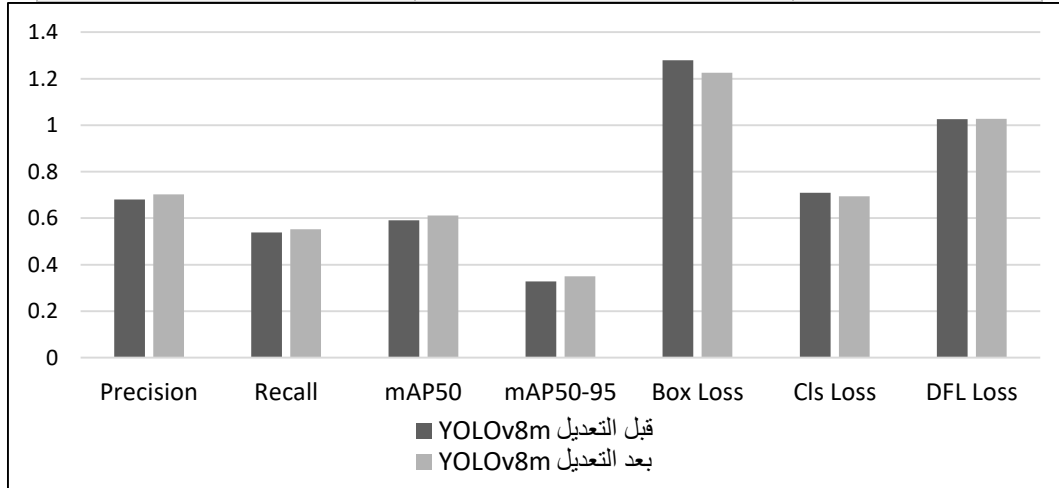
(Overfitting). فبعد مرحلة التذبذب الطفيف في الحقب الأولى، استقرت قيم الخسارة حول قيم ثابتة (\approx Box 1.27، \approx Cls 0.70، \approx DFL 1.03)، يدل هذا الاستقرار على أن النموذج يحتفظ بأداء جيد على بيانات لم يرها سابقاً، وهو ما يعزز قابلية تعميمه في بيانات مرور جديدة.

– أما بالنسبة لمقاييس الأداء (Precision, Recall, mAP50, mAP50-95)، تُظهر المنحنيات الأربعة الخاصة بها تحسناً مستمراً ولسلاً حتى نهاية التدريب. حيث نلاحظ أن الدقة Precision ارتفعت تدريجياً من نحو 0.50 إلى أكثر من 0.70، مما يدل على انخفاض مستمر في عدد الإيجابيات الكاذبة. أيضاً، نلاحظ أن Recall استقر حول 0.55 بعد ارتفاع سريع خلال أول 20 عصرًا، مما يشير إلى أن النموذج أصبح يكتشف معظم الأجسام الحقيقية باطراد. كذلك أظهرت كل من mAP50 و mAP50-95 نموًا متدرجًا حتى استقرتا عند 0.61 و 0.35 على التوالي، مما يعكس تحقيق توازن جيد بين تحديد موقع الكائن ودقة حدوده.

الجدول (٢): مقارنة بين نتائج تدريب كل من نموذج yolov8m (قبل التعديل) ونموذج yolov8m (بعد التعديل) على قاعدة البيانات

BDD100K المخصصة

المقياس	YOLOv8m (بعد التعديل)	YOLOv8m (قبل التعديل)
Precision	0.70261	0.6801
Recall	0.55278	0.5391
mAP50	0.61156	0.5905
mAP50-95	0.34909	0.3275
Box Loss	1.22531	1.2799
Cls Loss	0.69434	0.7089
DFL Loss	1.02752	1.02552



الشكل (٦): مقارنة بين نتائج تدريب كل من نموذج yolov8m قبل وبعد التعديل على قاعدة البيانات BDD100K المخصصة

– يظهر الجدول (٢) والشكل (٦) مقارنة بين نتائج تدريب نموذج YOLOv8m قبل وبعد التعديل، حيث نلاحظ أن النموذج بعد إجراء التعديلات حقق دقة (\approx Precision 0.703) مقارنة بـ 0.680 في النسخة الأصلية، وهو ما يشير إلى انخفاض نسبة الإيجابيات الكاذبة (False Positives) وتحسن موثوقية التنبؤات. يعكس هذا التحسن قدرة النموذج المعدل على التمييز بين الأجسام الفعلية والعناصر غير المرغوبة في الصور المزدحمة، خصوصاً بعد أن أتاحت عمليات تعزيز البيانات (Data Augmentation) تعرض النموذج لتوزيعات مكانية وإضاءة أكثر تنوعاً.

- كما ارتفع مؤشر الاستدعاء (Recall) بشكل طفيف من ٠,٥٣٩ إلى ٠,٥٥٣، ما يدل على تحسن قدرة النموذج على اكتشاف عدد أكبر من المركبات والعناصر الحقيقية ضمن المشاهد المعقدة. هذا الارتفاع، يُعد مهماً في تطبيقات المرور الواقعية، حيث يكون الخطأ الناتج عن عدم اكتشاف كائن موجود (False Negative) أكثر خطورة من اكتشاف كائن غير موجود بالفعل (False Positive)، لا سيما في سيناريوهات حساسة مثل اكتشاف المشاة أو الدراجات النارية.
- أما بالنسبة لمتوسط الدقة عند (mAP50)، فقد تحسّن من ٠,٥٩٠٥ إلى ٠,٦١١٦، مما يشير إلى تحسّن في جودة تموضع الصناديق التنبؤية ودقة الإطارات الناتجة. هذا الارتفاع الطفيف يعكس استقرار النموذج المعدّل في المراحل الأخيرة من التدريب.
- وبالنسبة إلى mAP50-95، ارتفع المقياس من ٠,٣٢٧٥ إلى ٠,٣٤٩١، ما يعكس تحسّن قدرة النموذج على التنبؤ الدقيق بالمواقع عبر نطاق واسع من معايير التداخل. هذا المقياس يُظهر أن النموذج المعدّل أصبح أكثر دقة في تمثيل الشكل الهندسي للكائنات واستخراج حدودها بدقة عالية، وهي سمة ضرورية في البيئات الحضرية المعقدة التي تتطلب تمييزاً واضحاً بين المركبات المتقاربة مكانياً.
- من ناحية أخرى، شهدت مقاييس الخسارة على بيانات التحقق (Validation Losses) تحسّناً عاماً، فقد انخفضت قيمة Box Loss من 1.2779 إلى 1.2255، مما يدل على تحسّن في دقة تحديد مواضع الكائنات. كما تراجعت Cls Loss من 0.7089 إلى 0.6943، وهو ما يعكس تصنيفاً أكثر ثباتاً للفئات التسع المستخدمة في مجموعة البيانات. أما DFL Loss فقد بقي قريباً من قيمته السابقة (≈ 1.03)، ما يعني أن دقة تحديد البنية المكانية داخل صندوق الحدود حافظت على استقرارها.
- بصورة عامة، تُظهر النتائج أن التحسينات المضافة على إعدادات التدريب لم تُغيّر فقط من القيم الرقمية للمقاييس، بل أثّرت أيضاً على السلوك الديناميكي للنموذج أثناء التدريب، إذ أصبحت منحنيات الخسارة أكثر سلاسة واستقراراً، ما يدل على توازن أفضل بين الدقة والتعميم. وبالتالي، يمكن اعتبار النموذج YOLOv8m (بعد التعديل) الأكثر اتزاناً بين النماذج الثلاثة، كونه يجمع بين الأداء المرتفع والاستقرار التدريبي، مما يجعله الأنسب لتطبيقات الرؤية الحاسوبية في أنظمة النقل الذكية التي تتطلب كشافاً موثوقاً ومرتزناً في الزمن الحقيقي.
- بمقارنة النتائج التي تم التوصل إليها في هذا البحث مع نتائج الدراسات السابقة يتبين أن نموذج YOLOv8m المعدّل وفق المنهجية المقترحة حقق تقدماً واضحاً على عدد من الجوانب التي واجهت قسوراً في الأعمال الأخرى كما يلي:
- **الدراسة الأولى [٦]**، التي ركّزت على خمس فئات فقط من المركبات، حققت قيم mAP50 بلغت ٠,٩١ وأداءً جيداً في دقة التصنيف، لكنها لم تختبر النموذج على الفئات الدقيقة والحساسة مثل المشاة والعلامات الطرقية وإشارات المرور، وهي فئات تُعدّ أصعب بطبيعتها بسبب صغر حجمها وارتفاع معدلات الحجب. وفي المقابل، استخدم هذا البحث تسع فئات بما فيها الفئات الصغيرة، ومع ذلك حافظ نموذج YOLOv8m المعدّل وفق المنهجية المقترحة على أداء مرتفع حيث بلغ $mAP50 \approx 0.61$ و $mAP50-95 \approx 0.35$ ، وهو إنجاز مهم بالنظر إلى توسع نطاق المهمة وتعقيدها مقارنة بالدراسة الأولى.

– أما الدراسة الثانية [٧]، فقد ركزت على تقليل حجم النموذج وتحسين سرعة الاستدلال، وحققت زيادة طفيفة في الدقة (٠,٥%) مقارنة بالنسخة الأصلية مع تقليل كبير في العمليات الحسابية. إلا أن قيمة mAP50 بقيت منخفضة نسبياً (mAP50 ≈ 48.8%) مقارنة نموذج YOLOv8m المعدل وفق المنهجية المقترحة (mAP50 ≈ 0.61)، مما يشير إلى أن خفض التعقيد الحسابي جاء على حساب القدرة على نمذجة التفاصيل الدقيقة.

– وفيما يتعلق بالدراسة الثالثة [٨]، التي ركزت على كشف إشارات المرور باستخدام YOLOv8، فقد حققت دقة مرتفعة mAP50 ≈ 0.985، إلا أن نطاق الدراسة كان محدوداً بفئة واحدة فقط وضمن ظروف محلية محدودة. وتُظهر نتائج هذا البحث أن تعميم النموذج ليعمل على تسع فئات مختلفة وأكثر صعوبة يؤدي بطبيعة الحال إلى انخفاض في قيمة mAP50، إلا أن الأداء يبقى قوياً ومتوازناً بفضل التصميم البنيوي لـ YOLOv8m وقدرته على التعامل مع تباين المشاهد.

الاستنتاجات والتوصيات

أظهرت نتائج هذه الدراسة أن اختيار النموذج المناسب، مع ضبط الإعدادات وتجهيز قاعدة البيانات بشكل منهجي، ينعكس بصورة مباشرة على أداء أنظمة الكشف المرئي في بيئات المرور الحضرية المعقدة. فقد بينت التجارب أن نموذج YOLOv8m يتفوق بوضوح على النموذج الأصغر YOLOv8n من حيث جميع المقاييس الرئيسية، بما في ذلك الدقة (Precision) الاستدعاء (Recall) ومتوسط الدقة (mAP)، وهو ما يؤكد أهمية العمق البنيوي وعدد المعاملات في تحسين جودة التمثيل البصري وتمكين النموذج من التعرف على الأنماط الأكثر تعقيداً في الصور.

كما أثبتت المرحلة الثانية من التدريب – والتي شملت تحسينات على بيئة التدريب مثل تعزيز البيانات (Data Augmentation)، زيادة دقة الصور، وتغيير خوارزمية التحسين – أنها تساهم بشكل فعال في رفع أداء النموذج، حيث ظهر التحسن واضحاً في منحنيات الخسارة وفي قيم mAP، إضافة إلى زيادة الاستقرار في مرحلة التحقق. تعكس هذه النتائج أن جودة البيانات ومنهجية التدريب لا تقل أهمية عن بنية النموذج نفسها في الوصول إلى أداء متقدم.

ولكن على الرغم من ذلك، كشفت الدراسة أيضاً عن وجود حدود واضحة لأداء النموذج، إذ حافظت بعض مقاييس الخطأ (خصوصاً المتعلقة بتموضع الصناديق) على مستويات يمكن تحسينها وهذا يشير إلى أن التحسينات على مستوى البيانات أو الإعدادات وحدها غير كافية للوصول إلى أداء مثالي.

وبناءً على هذه النتائج، يمكن الاستنتاج أن نموذج YOLOv8m يمتلك أساساً قوياً ومناسباً لتطبيقات النقل الذكي، إلا أن تحسين أدائه بشكل أكبر يتطلب الانتقال من تعديل بيئة التدريب إلى تعديل البنية المعمارية للنموذج نفسه. ويمكن أن تشمل هذه التعديلات:

– دمج آليات الانتباه Attention لرفع حساسية النموذج تجاه الأجسام الصغيرة والمتداخلة.

– إدخال وحدات Transformer لتحسين فهم العلاقات السياقية داخل الصورة.

– تطوير رؤوس متعددة للكشف (Multi-Head Detection) لتحسين التعامل مع اختلاف

أحجام الكائنات في مشاهد المرور.

وتشير نتائج هذه الدراسة إلى أن تطوير بنية YOLOv8m هو المسار الأكثر وعدًا لتحقيق أداء أعلى وموثوقية أكبر، خصوصًا في سيناريوهات المرور الواقعية التي تتطلب أعلى مستويات الدقة والزمن الحقيقي في آن واحد.

المراجع

- [1] Qureshi, K. N., & Abdullah, A. H. (2013). A survey on intelligent transportation systems. *Middle-East Journal of Scientific Research*, 15(5), 629-642.
- [2] Guo, Z., Huang, Y., Hu, X., Wei, H., & Zhao, B. (2021). A survey on deep learning based approaches for scene understanding in autonomous driving. *Electronics*, 10(4), 471.
- [3] Maity, M., Banerjee, S., & Chaudhuri, S. S. (2021, April). Faster r-cnn and yolo based vehicle detection: A survey. In *2021 5th international conference on computing methodologies and communication (ICCMC)* (pp. 1442-1447). IEEE.
- [4] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
- [5] Sohan, M., Sai Ram, T., & Rami Reddy, C. V. (2024). A review on yolov8 and its advancements. In *International Conference on Data Intelligence and Cognitive Informatics* (pp. 529-545). Springer, Singapore.
- [6] Desta, A. D., & Jian, C. (2025). Enhancing YOLOv8 for Vehicle Detection in Intelligent Traffic Management. *Metallurgical and Materials Engineering*, 31(4), 190-200.
- [7] Gao, D., Chen, T., & Miao, L. (2024, August). The YOLOv8m road object detection algorithm based on a lightweight design. In *Journal of Physics: Conference Series* (Vol. 2816, No. 1, p. 012034). IOP Publishing.
- [8] Biswas, S., Acharjee, S., Ali, A., & Chaudhari, S. S. (2023, December). YOLOv8 based Traffic Signal Detection in Indian Road. In *2023 7th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)* (pp. 1-6). IEEE.
- [9] Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F., ... & Darrell, T. (2020). Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2636-2645).
- [10] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- [11] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 7464-7475).
- [12] Raharjo¹, A. B., Dumont, F., & Thibaudeau, E. (2024, October). Comparative Analysis of YOLO-Based Object. In *Advances in Computing and Data Sciences: 8th International Conference, ICACDS 2024, Vélizy, France, May 9–10, 2024, Revised Selected Papers* (Vol. 2194, p. 93). Springer Nature.