



الجمهورية العربية السورية
وزارة التعليم العالي والبحث العلمي
جامعة دمشق
كلية الهندسة الميكانيكية والكهربائية
قسم هندسة الحواسيب والأتمتة

تحسين أداء المركبات ذاتية القيادة باستخدام تقانات الكشف العميق

رسالة قدمت لنيل درجة الماجستير في هندسة وبرمجة الروبوت

إعداد المهندس

وسيم مأمون يونس

إشراف

الدكتور المهندس

رؤوف حمدان

العام الدراسي 2024-2025

قرار الحكم مع توقيع اللجنة

الموافقة على تأليف لجنة الحكم على رسالة الماجستير في قسم الحواسيب والأتمتة التي أعدها الطالب وسيم مأمون يونس بعنوان " تحسين أداء المركبات ذاتية القيادة باستخدام تقانات الكشف العميق " بكلية الهندسة الميكانيكية والكهربائية من السادة الأساتذة :

الدكتور المهندس سمير كرمان

الدكتور المهندس محمد مازن المحاييري

الدكتور المهندس رؤوف حمدان

تصريح خطي

أنا الموقع أدناه أصرح بعلمي الكامل وقبولي:

1. أن كل ما ينتج عن البحث والأطروحة هو ملكية فكرية ومالية بالتشارك مع جامعة دمشق، وأنني ألتزم بأخذ موافقة الجامعة في حال رغبتني بنشر البحث أو الأطروحة أو جزء منها نصًّا أو مضمونًا خارج إطار الجامعة (من دور نشر أو مكنتبات أو مواقع إلكترونية وغيرها من وسائل النشر).

2. أنه لا يوجد أي جزء من هذه الأطروحة اقتبس من عمل علمي آخر أو أنجز للحصول على شهادة أخرى في جامعة دمشق أو أية جامعة أو معهد تعليمي داخل أو خارج الجمهورية العربية السورية.

الاسم والتوقيع

صفحة شكر وتقدير

يشرفني أن أتوجه بجزيل الشكر والعرفان إلى كلية الهندسة الميكانيكية والكهربائية بجامعة دمشق وإدارتها وكامل أعضاء هيئتها التدريسية وموظفيها، هذه المؤسسة العلمية العريقة التي أتاحت لي فرصة التعلم واكتساب أحدث الخبرات.

وأخص بالشكر الدكتور المشرف رؤوف حمدان عضو هيئة تدريسية في قسم هندسة الحواسيب والأتمتة لتفضله بالإشراف على هذه الرسالة، والذي لولا توجيهاته العلمية وخلقه الرفيع وصبره الإنساني لما تم هذا العمل المتواضع، مع تقديري الخالص للدكتور مأمون يونس الذي كان سنداً وعوناً لي لإنجاز هذا البحث ، وشكر خاص مملوء بالمحبة والعرفان للمهندسة دارينا يونس التي ساعدتني في تنفيذ وإنجاز هذا البحث ، ولا يسعني إلا أن أتقدم بالشكر والعرفان لرئيس قسم هندسة الحواسيب والأتمتة الدكتور سمير كرمان وكل كلمات الشكر لا تكفي للتعبير عن الامتنان لما قدمه من عطاء مستمر ورعاية، مع شكري وتقديري أيضاً لكامل الكادر التدريسي والفني والإداري في قسم هندسة الحواسيب والأتمتة لكل ما قدموه من مساعدات بكل صدق وود خلال تواجدي بينهم.

كما يسعدني أن أتوجه إلى الأساتذة الأفاضل رئيس لجنة الحكم وأعضائها بالشكر الكبير لموافقتهم على مناقشة هذه الرسالة وتحكيمها، وامتناني لكل ما سيقدمونه من ملاحظات علمية قيمة ستسهم في تصويب البحث وتطويره فشكراً لكم على عطائكم المتميز.

الفهرس

7	 فهرس الأشكال
8	 فهرس الجداول
9	 قائمة الاختصارات والمصطلحات العلمية
10	 الملخص
11	 التمهيد
12	 المقدمة
12	أ- إشكالية البحث:
13	ب- أسئلة البحث:
13	ت- الهدف من البحث:
13	ث- أهمية البحث:
14	ج- مبررات البحث:
14	خ- منهج البحث وأدواته:
14	د- محتويات الأطروحة:
15	 الفصل الأول المفاهيم النظرية
16	1.1- مقدمة
17	2.1- مفهوم تعلم الآلة ML
18	3.1- أنواع تعلم الآلة
19	4.1- خوارزميات التصنيف
19	1.4.1- التراجع اللوجستي (LR)
20	5.1- الشبكات العصبونية
24	6.1- خوارزميات الكشف باستخدام التعلم العميق
24	1.6.1- خوارزميات YOLO
27	2.6.1- خوارزميات YOLOv8n
29	 الفصل الثاني الدراسات المرجعية
30	1.2- مقدمة
30	2.2- الدراسات المرجعية لمساعدة السائق

3.2-الدراسات المرجعية لكشف الطرق.....	Error! Bookmark not defined.
4.2- الدراسات المرجعية لمساعدة ذوي الاحتياجات الخاصة.....	45
الفصل الثالث الدراسة العملية وخطوات التنفيذ.....	47
1.3- مقدمة.....	48
2.3- تفكيك الخوارزميات الفردية.....	48
1.2.3- نقطة التقاطع الفعال لخسارة الاتحاد EloU.....	48
2.2.3- الخسارة المرتبطة بعدم اليقين الهندسي GUA.....	49
3.2.3- الاتساق التنافسي أو التدريب التنافسي AC.....	50
4.2.3- التعلم العميق البايزي لتقدير عدم اليقين BU.....	51
5.2.3- آليات الانتباه.....	52
3.3- إستراتيجية الدمج.....	48
4.3- تحليل الاختبارات.....	48
1.4.3- هيكلية التحقق المنهجي.....	48
2.4.3- محاكاة معاملات الخوارزميات.....	56
5.3- الخسارة الكمومية.....	65
الفصل الرابع آلية عمل دالة الخسارة وتأثير المعاملات.....	68
1.4- مكونات دالة الخسارة.....	69
2.4- معاملات دالة الخسارة.....	70
3.4- استراتيجية محاكاة الخوارزميات.....	70
4.4- تطوير الخوارزمية.....	71
5.4- تحليل أداء النماذج.....	76
6.4- استقرار التدريب وفعالية الطريقة.....	77
الفصل الخامس الاستنتاجات والأفاق المستقبلية.....	84
1.5- الاستنتاجات.....	85
2.5- الأفاق المستقبلية.....	86
المراجع.....	87

فهرس الأشكال

الرقم	العنوان	رقم الصفحة
الشكل (1-1)	الفرق بين البرمجة التقليدية وتعلم الآلة	20
الشكل (2-1)	Sigmoid تابع	23
الشكل (3-1)	يبين عمل الشبكة العصبونية	27
الشكل (4-1)	يبين البنية الأساسية للشبكة العصبونية	32
الشكل (5-1)	يبين أنواع الكشف عن الأغراض	36
الشكل (1-2)	التصميم المقترح	38
الشكل (2-2)	يوضح تدفق المعالجة لتصميم الكشف المطبق المعتمد على YOLO	49
الشكل (3-2)	يوضح الهيكل العام لنظام YOLO إلى جانب المدخلات والمخرجات	50
الشكل (4-2)	يبين أبعاد حلقات الكشف كما تم تقديرها بواسطة النموذج	53
الشكل (5-2)	شبكة Faster R-CNN مع VGG16 المعدلة (MVGG16)	54
الشكل (6-2)	يبين أداء نماذج YOLOv5 ذات الأحجام المختلفة	54
الشكل (1-4)	يوضح مكونات دالة الخسارة	55
الشكل (2-4)	يوضح مراحل تطوير الخوارزمية النووية	56
الشكل (3-4)	يوضح المخطط العملي للنموذج المقترح	57
الشكل (4-4)	يوضح منحي التعلم الخاص بالخوارزمية المقترحة	58
الشكل (5-4)	يوضح التحسين التراكمي المتقدم للنماذج	58
الشكل (6-4)	يوضح منحنى الخسارة للنماذج الخمسة	59
الشكل (7-4)	يوضح التحسين المتقدم أثناء التدريب	60
الشكل (8-4)	يوضح موازنة الدقة والاستدعاء ومكونات	61
الشكل (9-4)	يوضح خطط الاتزان للنماذج الخمسة	62

فهرس الجداول

الرقم	العنوان	رقم الصفحة
الجدول (1-2)	يسرد أرقام مجموعة البيانات المستخدمة	67
الجدول (2-2)	يحدد سلوكيات القيادة المستخدمة	71
الجدول (3-2)	يسرد عدد الصور المستخدمة في أوضاع التدريب والتحقق والاختبار	79
الجدول (4-2)	يصف النتائج التجريبية بين YOLO-fastest وYOLO-fastest-xl	79
الجدول (5-2)	يبين مقارنة أداء F1-Score وFNR و mAP بين النماذج الأربعة المعتمد على YOL	67
الجدول (6-2)	يوضح مقارنة بين التطبيقات المختلفة لاكتشاف CNN المستندة إلى YOLO	71
الجدول (7-2)	يوضح نتائج المقارنة بين YOLOv3 وYOLOv4	67
الجدول (8-2)	يبين حلقات الكشف كما تم تقديرها بواسطة النموذج	71
الجدول (1-3)	يوضح استراتيجية محاكاة الخوارزمية	67
الجدول (2-3)	يوضح التعديلات القياسية للمعاملات	71
الجدول (1-4)	يوضح تأثير معاملات دالة الخسارة	67
الجدول (2-4)	يوضح منهجية التطوير المتدرج	71
الجدول (3-4)	يوضح أداء النموذج المقترح بعد التحسين	67
الجدول (4-4)	يوضح أفضل خمسة نماذج	71
الجدول (5-4)	يوضح أداء النماذج الخمسة	71
الجدول (6-4)	يوضح مقارنة النموذج المقترح مع الدراسات المرجعية	71

قائمة الاختصارات والمصطلحات العلمية

المصطلح	المعنى	المعنى في اللغة الانكليزية
YOLO	أنت تنظر مرة واحدة فقط	You Only Look Once
AI	الذكاء الاصطناعي	Artificial intelligence
RCNN	الشبكات العصبونية الالتفافية المعتمدة على المنطقة	Region-based Convolutional Neural Networks
ML	تعلم الآلة	Machine learning
CNN	الشبكات العصبونية الالتفافية	Convolutional Neural Networks
DPU	وحدة معالجة التعلم العميق	Deep Learning Processor Unit
EIoU	التقاطع الفعال لخسارة الاتحاد	Efficient Intersection over Union Loss
cls	خسارة التصنيف	classification Loss
dfl	خسارة البؤر التوزيعية	Distribution Focal Loss
IoU	تقاطع فوق الاتحاد	Intersection over Union
C IoU	التقاطع الشامل فوق الاتحاد	Complete Distribution Focal Loss
AC	الاتساق التنافسي	Adversarial Consistency
BU	اللايقين	Bayesian Uncertainty
GUA	الخسارة المرتبطة بعدم اليقين الهندسي	Geometric Uncertainty-Aware Loss
VM	آلة افتراضية	Virtual machine
QL	الخسارة الكمومية	Quantum Loss
QB	الكوانتم البايزي	Quantum Bayesian
mAP50	مقاييس متوسط الدقة المتوسطة	Mean Average Precision
WAN	شبكة المنطقة الواسعة	Wide Area Network

الملخص

أدت التطورات الحديثة في الذكاء الاصطناعي (AI) إلى تحسين كبير في قدرات اكتشاف الأشياء للمركبات ذاتية القيادة، وخاصة باستخدام الشبكات العصبونية الالتفافية (CNNs). ومع ذلك، فإن تحقيق مستويات عالية من الدقة والسرعة في وقت واحد في بيئات المركبات لا يزال يمثل تحديًا. لذلك، تقترح هذه الورقة منهجية "محاكاة المعاملات" كبديل عملي وفَعَالٍ للتعديل المباشر على دالة الخسارة حيث حققت أعلى دقة شاملة mAP50 بين النماذج المختبرة وأعلى تحسن في الدقة الدقيقة mAP50-95 بنسبة +9.11% ، مما يثبت تفوقها في التحديد الدقيق لمواقع الأغراض كما حققت توازنًا مثاليًا بين الدقة Precision والاستدعاء Recall . تم تحقيق تحسن ملحوظ في أداء نموذج YOLOv8n دون المساس بسرعته أو استقراره، مما يجعله أكثر ملاءمة لتطبيقات الزمن الحقيقي مثل المركبات ذاتية القيادة وروبوتات الإنقاذ. ومن الملاحظ أن النموذج المقترح يحقق زيادة كبيرة في الدقة، مع تحسينات تتراوح من 5 إلى 9 بالمائة مقارنة بنماذج YOLO المستقلة. نتائج هذا البحث لها آثار عملية على دمج تقنيات الذكاء الاصطناعي في أنظمة القيادة الذاتية.



المقدمة

شهدت السنوات الأخيرة تطوراً مذهلاً في مجال التكنولوجيا، وخاصة في قطاع النقل، حيث أصبحت السيارات ذاتية القيادة، التي تعتمد على تقنيات الذكاء الاصطناعي، والتي تُعد واحدة من أبرز الابتكارات التي تهدف إلى تحسين السلامة وكفاءة القيادة. إن السيارات ذاتية القيادة هي مركبات قادرة على التنقل والتفاعل مع محيطها بشكل مستقل دون الحاجة لتدخل بشري في القيادة، حيث تعتمد هذه السيارات على مجموعة من الأنظمة المتقدمة مثل المستشعرات، والكاميرات، والرادار، وأجهزة الاستشعار التي توفر معلومات دقيقة حول البيئة المحيطة، كما تساهم هذه الأجهزة في مراقبة حركة المرور، ومسافة الأجسام القريبة، والعوائق المحتملة، بالإضافة إلى ذلك، تعتمد هذه السيارات على خوارزميات ذكية لتحليل البيانات واتخاذ قرارات القيادة المناسبة، مثل التوجيه، والتوقف، أو تغيير المسار، بهدف توفير قيادة آمنة وفعالة.

يعد التعلم العميق أحد الحلول المحتملة لاكتشاف الأشياء ومشاكل إدراك المشهد، والتي يمكن أن تمكن السيارات التي تعتمد على الخوارزميات والبيانات. في هذا البحث، نهدف إلى سد الفجوة بين التعلم العميق والسيارات ذاتية القيادة من خلال استطلاع شامل. التعلم العميق هو نموذج حسابي متعدد الطبقات يستخدم لاستخراج الميزات والتعلم التمثيلي على مستويات مختلفة من التجريد. التعلم العميق هو فرع من تعلم الآلة يستخرج تلقائياً الميزات والأنماط من البيانات الأولية ويقوم بالتنبؤات أو يتخذ إجراءات بناءً على بعض وظائف المكافأة. وهو يتألف من تقنيات مثل الشبكات العصبونية، والأساليب الاحتمالية الهرمية، ونماذج التعلم الخاضعة للإشراف وغير الخاضعة للإشراف، والتعلم المعزز العميق (DRL). ويعتمد التعلم العميق على التطورات المبكرة في الشبكات العصبونية الاصطناعية (ANN).

أ- إشكالية البحث:

بالنسبة للمركبات ذاتية القيادة الذكية، فإنها تلعب دوراً مهماً في الاستخدام السهل للمستخدم في البيئات المزدهمة، لذلك كان التركيز مؤخراً على توفير وظائف التعرف على الأشياء وتجنب الاصطدام. فإن استخدام الشبكة العصبونية الالتفافية CNN المستندة إلى وحدة معالجة التعلم العميق في خوارزمية YOLO لمعالجة الصورة واكتشاف الكائنات وتصنيفها بسرعة ودقة عالية، وتسمح للمركبات ذاتية القيادة صغيرة الحجم التي يمكن استخدامها في المساعدة بالاستكشاف عن طريق

الابصار الحاسوبي وبالتالي نحن بحاجة لتطوير خوارزميات اكتشاف الأغراض في الصور والعامل المهم هو السرعة لاكتشاف الأغراض وتجنب الاصطدام .

ب- أسئلة البحث:

يحاول البحث الإجابة عن الأسئلة الآتية :

1. ما هي متطلبات التشغيل الأساسية التي يعزز بعضها البعض للسيارات ذاتية القيادة التي تعمل بشكل كامل؟

1. ما هي الإنجازات والتطورات التي تم تحقيقها في السيارات ذاتية القيادة في السنوات الـ 20 الماضية وما هي بعض الاتجاهات البحثية الواعدة للعقد القادم؟

2. هل من الممكن في نهاية المطاف تحقيق المعرفة والإدراك على المستوى البشري في السيارات ذاتية القيادة؟

3. لماذا يعتبر التعلم العميق تقنية واعدة لحل اكتشاف الأشياء وإدراك المشهد في السيارات ذاتية القيادة؟ ما هي أحدث نماذج التعلم العميق المستخدمة للكشف عن الأشياء وإدراك المشهد في السيارات ذاتية القيادة

4. ما هي أحدث وأنجح تقنيات الكشف عن الأجسام المطبقة على المركبات ذاتية القيادة والاتجاهات الواعدة لمزيد من البحث؟

ت- الهدف من البحث:

تحسين أداء خوارزمية Yolo وزيادة دقتها وسرعتها ودقتها للاستكشاف لتمكين المركبات ذاتية القيادة للتعرف على الأشياء وأماكن تواضع الحفر في الطرقات لتجنب وقوع الحوادث أو التعرف على أماكن وجود العالقين في الزلازل بأسرع زمن ممكن .

ث- أهمية البحث:

ويعتمد نظام الكشف عن المركبات، الذي يستخدم صوراً منخفضة الجودة تلتقطها كاميرا فيديو أحادية مثبته في الجزء الأمامي من السيارة، على الشبكات العصبونية الالتفافية "CNN" أو YOLO. يمكن لـ YOLO8n استخراج الميزات العامة للصور باستخدام الطبقات الالتفافية والحصول على ملامح أكثر دقة للمركبات. توفر بنية YOLO8n المقترحة في المقالة كشفاً عالي الدقة لصور المركبات. وتوضح التجارب التي تم إجراؤها باستخدام مجموعة بيانات GTI أن نظام الكشف عن

المركبات المعتمد على YOLO8n يحقق نتائج دقيقة للغاية ويكون أكثر قوة في التعامل مع الأشكال المختلفة من الصور.

ج- مبررات البحث:

تسبب العوائق في الطرقات العديد من الحوادث وتشكل تهديداً كبيراً على حياة المواطنين وأيضاً عند حدوث الزلازل يصعب تحديد الابنية التي تحتوي على اشخاص عالقين يسبب ضياع في الوقت والجهد وبالتالي يأتي دور أنظمة الرؤية للسيارات ذاتية القيادة، ودور التعلم العميق في تفسير الرؤية المعقدة، وتعزيز الإدراك، وتشغيل المناورات الحركية في السيارات ذاتية القيادة وبالتالي نحن بحاجة لتطوير خوارزميات للمركبات ذاتية القيادة للتعرف على الأشياء ولإنقاذ حياة الناس بأسرع وقت ممكن

ح- منهج البحث وأدواته

اعتمد الباحث على المنهج التحليلي والتجريبي ، فدرست أداء خوارزمية YOLO8n ، وحلّل أدائها ومن ثم اقترحت تطويرات على جزئية ضبط الحزمة واختبر أثرها عبر مجموعة من التجارب، تم إعداد نموذج الكشف باستخدام خوارزمية YOLO8n وتم تدريبه على Google Collab ؛ لإنجاز أفكار البحث، وإجراء تجاربه.

خ- محتويات الأطروحة:

تتألف الأطروحة من ثلاثة فصول، تتناول أدبيات المسألة المدروسة، وتبين الأعمال السابقة، ومن ثم توضح التحسينات المقترحة وتناقش نتائجها، والأفاق المستقبلية للبحث. يحتوي التمهيد على مقدمة عامة، وخطة البحث، وعرضاً موجزاً لمحتويات الأطروحة. يستعرض الفصل الأول المفاهيم النظرية الأساسية التي تساعد على بناء الإطار العام لموضوع البحث؛ فيوضح المفاهيم الأساسية في خوارزميات التعلم العميق مثل الشبكات العصبونية وخوارزميات الـ YOLO ، وتناول الفصل الثاني الدراسات المرجعية كاستخدام خوارزميات الكشف عن الحفر باستخدام الـ YOLO ، ومقارنة النتائج مع الدراسات المرجعية المختلفة وسلبيات كل دراسة ، ويتطرق الفصل الثالث إلى الدراسة العلمية وخطوات التنفيذ وقياس الأداء بين الأنواع المختلفة من الخوارزميات المدمجة المقترحة ، ويدرس الفصل الرابع تابع الخسارة وتأثيره على المعاملات ، أما الفصل الخامس يتطرق إلى الإسهامات العلمية للبحث، والاستنتاجات التي خلص إليها، فضلاً عن التوصيات المستقبلية التي قدمها.

الفصل الأول

المفاهيم النظرية

- ❖ مفهوم تعلم الآلة ML
- ❖ أنواع تعلم الآلة
- ❖ خوارزميات التصنيف

الدراسة النظرية

1.1 - مقدمة

تشهد صناعة السيارات تحولاً سريعاً مع دمج تقنيات الذكاء الاصطناعي (AI). من المتوقع أن يشهد السوق العالمي للذكاء الاصطناعي في المركبات نمواً كبيراً، حيث من المتوقع أن تصل قيمة برمجيات الذكاء الاصطناعي للسيارات إلى 6.6 مليار دولار أمريكي بحلول عام 2026، لتتوسع بمعدل سنوي متوسط قدره 36.15%.

يلعب اكتشاف الأشياء دوراً حاسماً في أنظمة القيادة الذاتية، حيث يمكن السيارات ذاتية القيادة من إدراك وتصنيف الأشياء في محيطها بدقة. يتم استخدام استراتيجيتين رئيسيتين بشكل شائع للكشف عن الأشياء، وهما النهج ذو المرحلتين والمرحلة الواحدة.

يتم اعتماد الأساليب ذات المرحلتين، مثل الشبكة العصبونية الالتفافية القائمة على المنطقة (RCNN) والشبكة الأسرع R-CNN، على نطاق واسع في كل من البحث الأكاديمي والصناعة نظراً لدقتها العالية. تستخدم هذه الأساليب شبكة مقترحات المنطقة (RPN) لإنشاء مقترحات الكائنات المحتملة، تليها مرحلة التصنيف والمربع المحيط بالانحدار. ومع ذلك، فإنها غالباً ما لا ترقى إلى مستوى تلبية متطلبات المعالجة في الوقت الفعلي للكشف عن الأشياء في بيئات المركبات.

من ناحية أخرى، فإن الأساليب أحادية المرحلة، والتي تتمثل في النظر مرة واحدة فقط (YOLO) وكاشف الطلقة الواحدة (SSD)، تتميز بسرعتها. تتنبأ هذه الطرائق بشكل مباشر بفئات الكائنات وإحداثيات المربع المحيط من تقييم واحد للشبكة. إنها تحقق أداءً في الوقت الفعلي ولكنها قد تضحي ببعض الدقة في سيناريوهات معينة. على الرغم من مزاياها، يظل اكتشاف الأجسام ذات الأحجام المختلفة وفي ظل ظروف القيادة الصعبة يمثل تحدياً كبيراً لهذه الأساليب أحادية المرحلة. وللتغلب على القيود المفروضة على كل من الأساليب ذات المرحلتين والمرحلة الواحدة، هناك اهتمام متزايد بتطوير أساليب هجينة تستفيد من نقاط القوة في كلا النموذجين. تهدف هذه النماذج الهجينة إلى تحقيق التوازن بين الدقة والسرعة في مهام الكشف عن الأشياء. من خلال دمج الأفكار من كل من YOLO و Faster R-CNN، من الممكن تصميم بنية هجينة توفر دقة محسنة مع الحفاظ على قدرات المعالجة في الوقت الفعلي. يجمع النهج الهجين المقترح بين الكشف الفعال عن الكائنات لـ

YOLO لاختيار المربع المحيط وآلية تجميع المنطقة محل الاهتمام (RoI) من Faster R-CNN من أجل التجزئة والتصنيف الدقيق. يؤدي تخطي RPN من بنية Faster R-CNN إلى تسريع وقت المعالجة دون خسارة كبيرة في الدقة [1].

ومع ذلك، فإن الحاجة إلى نموذج محسّن للذكاء الاصطناعي للكشف عن الكائنات في المركبات ذاتية القيادة تنشأ بسبب التحديات الأساسية التي تواجه التعرف على الكائنات، مثل نشر الصورة، وعدم الوضوح، والحركة، لا سيما في سياق بيئة المدينة الذكية [17].

علاوة على ذلك، فإن القيود العملية، بما في ذلك موارد الحاسوب المحدودة ومشكلات الكمون، يمكن أن تعيق القدرة على تحديد الكائنات بسرعة ودقة في بيانات محددة. ولذلك، اقترحت العديد من الدراسات الحديثة تقنيات جديدة لتعزيز أداء نماذج الكشف عن الأجسام الهجينة للسيارات ذاتية القيادة.

2.1- مفهوم تعلم الآلة ML

تعلم الآلة هو أحد فروع الذكاء الاصطناعي يهدف إلى تطوير نظم قادرة على التعلم وعلى تطوير أدائها ذاتياً من خلال التجربة. فالهدف هو تعلم نموذج h من بيانات التدريب (training data) التي يمكن استخدامها للتنبؤ عن قيمة الهدف غير المعروفة أو لتسليط الضوء على وجود بنى في البيانات. ويتضمن تعلم الآلة عدداً كبيراً من حقول التطبيقات منها معالجة اللغات الطبيعية natural language processing وتمييز الأنماط syntactic pattern recognition ومحركات البحث search engines والتشخيص الطبي والمعلوماتية الحيوية والمعلوماتية الكيميائية، وتصنيف سلاسل الحمض النووي (DNA) والتعرف على الكلام speech recognition وتمييز الكتابة handwriting recognition وحتى تمييز الأشياء object recognition والابصار الحاسوبي computer vision وتحريك الروبوت robot locomotion إن الفرق الأساسي بين البرمجة التقليدية وتعلم الآلة هو أنه في البرمجة التقليدية يقوم الحاسب من خلال الخوارزمية المناسبة ومعطيات الدخل بحساب قيمة الخرج، بينما في تعلم الآلة يقوم الحاسب بتحليل معطيات الدخل ومقارنتها مع الخرج لبناء النموذج المناسب من أجل التنبؤ لاحقاً بخرج معطيات دخل جديدة لم يسبق للنموذج أن تدرب عليها من قبل [3,4,5]. يبين الشكل (1-1) الفرق بين البرمجة التقليدية وتعلم الآلة.

Traditional Programming



Machine Learning



الشكل (1-1) الفرق بين البرمجة التقليدية وتعلم الآلة

3.1- أنواع تعلم الآلة

يمكن تقسيم تعلم الآلة إلى نوعين رئيسيين:

1-التعلم الآلي الخاضع للإشراف (Supervised Learning)

في هذه النوع يتم تدريب النموذج على بيانات دخل معروفة الخرج بشكل مسبق لتشكل نقطة انطلاق يتعلم منها النموذج بهدف تصنيف بيانات دخل جديدة لاحقاً، ويُستخدم التعلم الخاضع للإشراف بشكل أساسي في المسائل الآتية:

1- التصنيف (Classification): في هذه المسائل يكون الدخل مصنفاً إلى عدد محدد من الصفوف، صفان أو أكثر، ففي حال وجود صفيين فقط تعد مسألة تصنيف ثنائي - Binary Classification، أما في حالة وجود أكثر من صفيين تعد مسألة تصنيف متعدد الصفوف Multi-Class Classification، إن هدف تعلم الآلة في مسائل التصنيف هو إيجاد الصف المناسب الذي ينتمي إليه الدخل، مثل تصنيف البريد الإلكتروني وفرزه، أو تصنيف الاستعلام إلى استعلام حميد أو استعلام خبيث، وهو هدف هذا البحث [6,7].

2- الانحدار (Regression): يكون الخرج في هذه المسائل عبارة عن قيمة مستمرة بدلاً من أصناف منفصلة كما في مسائل التصنيف، ثمة العديد من التطبيقات للانحدار، مثل التنبؤ بأسعار الأسهم في البورصة وحالة الطقس وغيرها.

2- التعلم الآلي غير الخاضع للإشراف (Unsupervised Learning):

في هذا النوع يتم تدريب النموذج على بيانات دخل بدون وجود أي مخرجات معروفة بشكل مسبق، حيث يقوم النموذج باستنباط العلاقات الخفية بين البيانات، ويستخدم هذا النوع بشكل أساسي في:

- 1- العنقدة (Clustering): يحاول النموذج إيجاد الأنماط المتشابهة بين البيانات وتجميعها كل منها ضمن صف مناسب بشكل مشابه لمسألة تصنيف متعدد الصفوف، إلا أن النموذج في هذه الحالة يكون جاهلاً بعدد الصفوف وبالمعنى الذي يعبر عنه كل صف.
- 2- التجميع (Association): يقوم النموذج بإيجاد الترابط بين مجموعة كبيرة من المعطيات، وتستخدم بشكل أساسي في التتقيب عن البيانات.

4.1- خوارزميات التصنيف

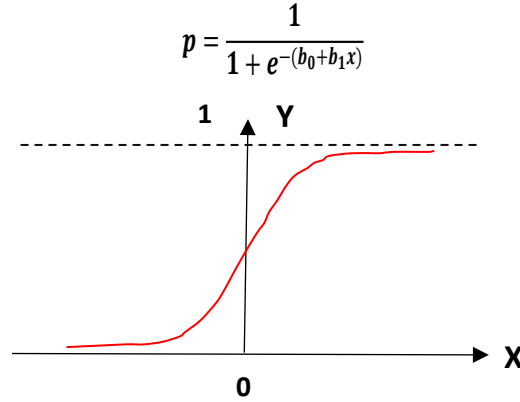
إن الهدف من هذا البحث هو تصنيف الاستعلام إلى استعلام خبيث أو حميد، ومن ثم فإن النموذج الأنسب هو نموذج تعلم آلي خاضع للإشراف يعتمد على التصنيف الثنائي، ومن أجل الحصول على أفضل النتائج تم الاعتماد على ثلاث خوارزميات للتصنيف، خوارزمية LR التي تتميز بقدرتها على التعامل مع البيانات القابلة للفصل خطياً وخوارزميتي MLP و GB اللتين بوسعهما إيجاد الأنماط غير الخطية بين البيانات.

1-4-1 التراجع اللوجستي (LR)

تعد خوارزمية التراجع اللوجستي إحدى أشهر خوارزميات تعلم الآلة المستخدمة في التصنيف، وهي عبارة عن نموذج إحصائي يعتمد على نمذجة المتغيرات وفق تابع رياضي من أجل التنبؤ باحتمال انتماء الخرج لصف معين، تتميز بالبساطة والسرعة الكبيرة في تصنيف البيانات القابلة للفصل خطياً [4].

وتستخدم خوارزمية التراجع اللوجستي تابعاً رياضياً يدعى Sigmoid، وظيفة هذا التابع اشتقاق العلاقة بين المتغيرات التي تمثل السمات والخرج الذي يمثل صفراً معيناً، إذ يقوم هذا التابع بالاعتماد على قيم المتغيرات بإعطاء قيمة تقع ضمن المجال بين 0 و 1 من أجل التنبؤ باحتمالية الخرج بالنسبة لقيم المتغيرات.

يوضح الشكل (2-1) شكل تابع Sigmoid باللون البرتقالي بالنسبة للمتغير x .



الشكل (2-1) تابع Sigmoid

تمثل b_1 الوزن الابتدائي للمتغير x بالإضافة لقيمة bias الأولية b_0 ، وعندما يكون الحد $(b_0 + b_1 x)$ أكبر من الصفر، فإن p التي تمثل قيمة تابع السيغمويد سوف تتجه نحو قيمة الواحد، بينما عندما يكون أصغر من الصفر فإن p سوف تتجه نحو قيمة الصفر، يمكن استخدام خوارزمية التراجع اللوجستي في مسائل التصنيف الثنائي وكذلك في مسائل التصنيف المتعدد [11].

1-5 الشبكات العصبونية

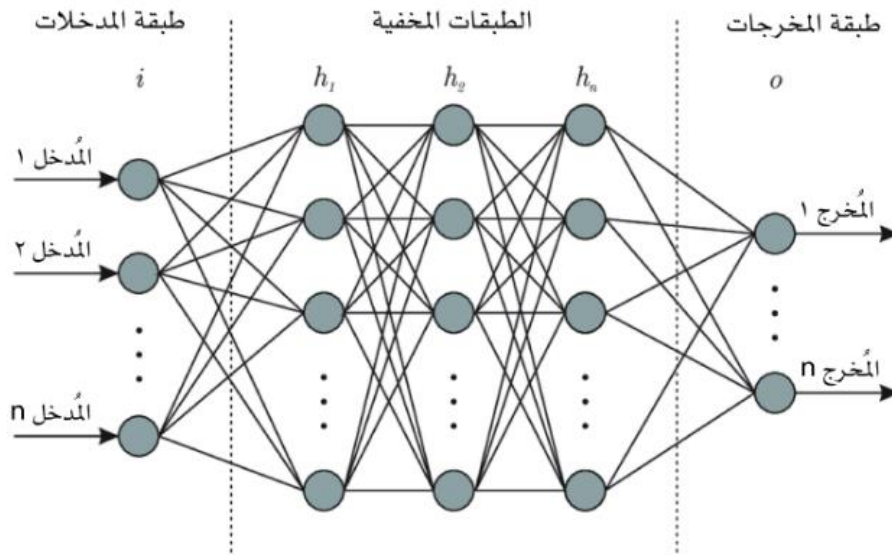
تمكن للشبكات العصبونية مساعدة الحواسيب على اتخاذ القرارات الذكية مع مساعدة بشرية محدودة. وهذا بسبب قدرتها على تعلم العلاقات بين بيانات الإدخال والإخراج اللاخطية والمعقدة ونمذجتها. فالشبكة العصبونية هي نموذج حوسبة مستوحى من نظام العصبونات في الدماغ البشري، وتُعتبر جزءاً أساسياً من مجال الذكاء الاصطناعي وتهدف إلى تمثيل ومحاكاة عملية التعلم والتعامل مع المعلومات بطريقة مُشابهة للدماغ البشري. تتكوّن الشبكة العصبونية من مجموعة من العناصر المعروفة باسم العقد (العصبونات) والتي ترتبط ببعضها البعض عن طريق الروابط. وتتكوّن الشبكة

العصبية من طبقة إدخال (input layer) لاستقبال المعلومات، وطبقات متعددة من العقد الداخلية (hidden layers) التي تقوم بمعالجة البيانات، وأخيرًا طبقة المخرجات (output layer) التي تُعطي النتائج النهائية [19].

تعمل الشبكة العصبونية عن طريق تحويل المعلومات الواردة إلى العقد في الطبقة الأولى، ثم يتم تمرير هذه المعلومات من خلال العقد في الطبقات الداخلية بطرق مُحددة. ويتم تعديل قوة الروابط بين العقد بناءً على الخبرة والتعلم من البيانات المتاحة، وهو الأمر الذي يُعرف بتدريب الشبكة العصبونية، حيث تتكيف الشبكة العصبونية تدريجيًا لأداء مهمة مُحددة من خلال تعديل الأوزان والقيم المرتبطة بالعقد. ويبين الشكل (1-3) كيفية عمل الشبكة العصبونية، بينما يبين الشكل (1-4) البنية الأساسية للشبكة العصبونية.



الشكل (1-3) - يبين كيفية عمل الشبكة العصبونية



الشكل (4-1)- يبين البنية الأساسية للشبكة العصبونية

وتتميز الشبكات العصبونية بقدرتها على التعامل مع البيانات المعقدة والتعلم من الأمثلة، مما يجعلها أداة فعالة في حل المشكلات التي تكون صعبة للغاية بالنسبة للأنظمة الكلاسيكية [21].

تمثل نظرية عمل الشبكات العصبونية الأساسية في تفاعل الطبقات مع بعضها البعض، حيث تستقبل طبقة المدخلات البيانات الأولية، بينما تُوفّر طبقة المخرجات النتيجة النهائية. فيما بينهما، يمكن أن تكون هناك طبقة مخفية واحدة أو أكثر تساعد في استخراج الميزات وتحويلها من البيانات المدخلة. وترتبط كل خلية عصبونية في الشبكة بالخلايا العصبونية في الطبقات المجاورة من خلال اتصالات وروابط ذات أوزان معينة، حيث تُحدّد هذه الروابط قوة وأهمية المعلومات التي يتم تمريرها بين الخلايا العصبية. لإدخال البيانات في الشبكة العصبونية، يتم تطبيق ما يُعرف بوظيفة التنشيط على مخرجات كل خلية عصبونية، حيث تُحدّد وظيفة التنشيط ما إذا كان ينبغي تنشيط الخلية العصبونية أم لا بناءً على مدخلاتها. تتضمن وظائف التنشيط الشائعة ReLU (الوحدة الخطية المعدلة)، و $\tanh$ (الظل الزائدي)، حيث تسمح وظيفة التنشيط للشبكة بنمذجة العلاقات المعقدة بين المدخلات والمخرجات. تنتقل البيانات بعد ذلك أثناء مرحلة الانتشار الأمامي، حيث يتم نقل البيانات إلى الشبكة العصبونية من خلال طبقة الإدخال. تتلقى كل خلية عصبونية في الطبقات اللاحقة مدخلات من الطبقة السابقة، وتُطبّق وظيفة التنشيط على المجموع المرجّح لمدخلاتها، وتمرر النتيجة

إلى الطبقة التالية. تتكرر هذه العملية حتى تمر البيانات عبر جميع الطبقات وتصل إلى طبقة المخرجات، والتي تُنتج بدورها التنبؤ النهائي أو التصنيف بناءً على المعلومات المُستفادة من الشبكة. لتقييم أداء الشبكة العصبونية، يتم استخدام دالة الخسارة، حيث تُحدد هذه الدالة الفرق بين المخرجات المتوقعة والمخرجات الفعلية. تهدف الشبكة إلى تقليل هذا الاختلاف عن طريق ضبط الأوزان. تتضمن دوال الخسارة الشائعة الخطأ التربيعي المتوسط (MSE) لمهام الانحدار ودالة الإنتروبيا المنقاطعة لمهام التصنيف. بعد ذلك، يتم الانتقال إلى مرحلة الانتشار العكسي والتي تُعد خطوة حاسمة في تدريب الشبكة العصبونية، ويتم حساب تدرجات دالة الخسارة فيما يتعلق بالأوزان وتحديث الأوزان وفقًا لذلك. يتم حساب التدرجات باستخدام قاعدة السلسلة في حساب التفاضل والتكامل، والتي تنتشر الخطأ من طبقة المخرجات مرةً أخرى إلى طبقة المدخلات. من خلال ضبط الأوزان بشكلٍ مُتكرر بناءً على التدرجات، تتعلم الشبكة تقليل دالة الخسارة وتحسين توقعاتها [23].

خلال مرحلة التدريب، تتعرض الشبكة العصبونية لمجموعة بيانات مُصنّفة، حيث تقوم الشبكة بعمل تنبؤات لكل إدخال، ومقارنتها بالتسميات الحقيقية، وحساب الخسارة، وتحديث الأوزان من خلال الانتشار العكسي. تتكرر هذه العملية التكرارية للانتشار الأمامي وحساب الخسارة والانتشار العكسي لفترات مُتعددة حتى تتقارب الشبكة إلى حالة يتم فيها تقليل الخسارة إلى الحد الأدنى وتكون التنبؤات دقيقة. بمجرد تدريب الشبكة العصبونية، يُمكن استخدامها للتنبؤ بالبيانات غير المرئية، حيث يتم تغذية بيانات الإدخال من خلال الشبكة باستخدام الانتشار الأمامي، وتنتج طبقة المخرجات النتيجة المتوقعة. تظل أوزان الشبكة ثابتة أثناء الاستدلال، حيث تم تحسينها أثناء التدريب. ويتم تقييم أداء الشبكة بناءً على دقتها أو إحكامها أو استرجاعها أو غيرها من المقاييس ذات الصلة، اعتمادًا على المهمة المطروحة. تتمثل ميزة الشبكات العصبية في إمكانية تعلم أنماط وعلاقات مُعقدة في البيانات دون الحاجة إلى برمجة واضحة. وقد تم تطبيقها بنجاح على مجموعة واسعة من التطبيقات، بما في ذلك التعرف على الصور والكلام، ومعالجة اللغة الطبيعية، وأنظمة التوصية، والمركبات ذاتية القيادة. على الرغم من أن نظرية عمل الشبكات العصبونية واحدة، إلا أنها في الواقع تضم العديد من الأنواع، وكلٌ منها مُصمّم لأداء لمهام مُحددة. فيما يلي بعض أهم الأنواع الرئيسية للشبكات العصبونية:

1- الشبكات العصبونية ذات التغذية الأمامية

يُعد هذا النوع هو أبسط أشكال الشبكات العصبونية، حيث تنتقل المعلومات في اتجاه واحد، من المُدخلات إلى المُخرجات، ولا توجد أي دورات أو حلقات في بنية الشبكة، فقط اتجاه خطي. ويُستخدم هذا النوع عادةً لمهام مثل التصنيف والانحدار.

2- الشبكة متعدد الطبقات (MLP):

هو نوع من الشبكات العصبونية المغذية ذات طبقة مخفية واحدة أو أكثر بين طبقات المُدخلات والمُخرجات. ترتبط الخلايا العصبية في كل طبقة بجميع الخلايا العصبونية في الطبقات المجاورة. وتُستخدم على نطاق واسع لمختلف المهام، بما في ذلك التعرف على الصور، والتعرف على الكلام، والنمذجة المالية .

3- الشبكات العصبونية الالتفافية (CNN):

تم تصميم هذا النوع لمعالجة البيانات الشبيهة، مثل الصور أو إطارات الفيديو، أي أنّ الشبكات تستقبل مجموعة ضخمة من البيانات ثم تقوم بالتعرف عليها بعد ذلك، مثل تغذية الشبكة بمجموعات كبيرة من صور القطط ثم جعلها تتعرف تلقائياً على صور القطط بعد ذلك بمفردها . ويُعد هذا النوع مُمتاز لتحليل الصور والفيديوهات، بما في ذلك مهام مثل تصنيف الصور واكتشاف الأشياء والتعرف على الوجه[23].

1-6- خوارزميات الكشف باستخدام التعلم العميق

يوجد العديد من خوارزميات الكشف باستخدام التعلم العميق وأشهرها على الإطلاق خوارزمية YOLO والنسخ المتعددة منها .

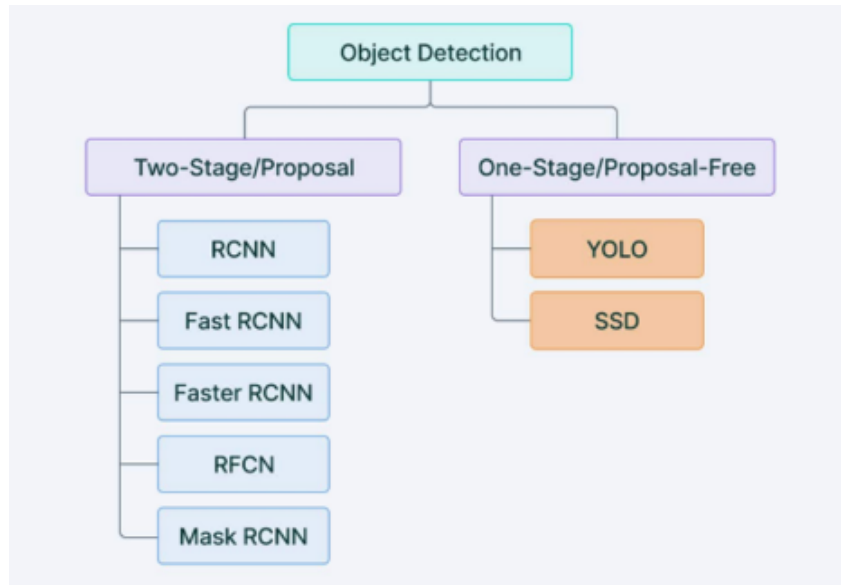
1-6-1 خوارزميات YOLO

تمّ تطوير العديد من الخوارزميات الحديثة التي حققت أداءً فعالاً في مجال الرؤية الحاسوبية في الزمن الحقيقي، إلا أنّ خوارزميات YOLO بإصداراتها المختلفة تفوّقت على العديد منها ولا سيّما إصدارتها الأخيرة التي تميزت بالفعّالة من حيث دقة الكشف والكلفة الحسابية المنخفضة وسرعة المعالجة وسهولة التدريب، ممّا يجعلها خياراً مثالياً لمعالجة البيانات الضخمة في تطبيقات الزمن الحقيقي ويفتح المجال لدعم هذه التطبيقات وتطويرها، ويعتبر الكشف عن الأغراض تقنية مهمة من تقنيات الرؤية الحاسوبية التي تعمل على اكتشاف وتحديد ومتابعة أغراض محدّدة من مقاطع صور أو فيديو، والتي يتم استخدامها في بعض التطبيقات الواقعية المهم. [2,3]

وتنقسم خوارزميات اكتشاف الغرض إلى نوعين أساسيين بالاعتماد على عدد مرّات تمرير الصّورة المدخلة إلى الشّبكة؛ هما كواشف تعتمد على مرحلة وحيدة (single-shot detectors) وكواشف ذات مرحلتين (two-stage detectors) .

1-كواشف تعتمد على مرحلة وحيدة (single-shot detectors) تعالج الصّورة ضمن مرحلة وحيدة وتكون ذات تكلفة حسابية أقلّ، وأكثر فعالية من أجل تطبيقات الزمن الحقيقي في البيئة ذات الموارد المحدودة، إلّا أنّها أقل دقّة من الخوارزميات الأخرى وأقلّ فعالية في اكتشاف الأغراض الصّغيرة.

2-وكواشف ذات مرحلتين (two-stage detectors) حيث يتمّ معالجة الصّورة المدخلة وفق مرحلتين المرحلة الأولى لاكتشاف مواقع الكائنات المحتملة، والمرحلة الثانية لتنقيح هذه المقترحات وتحديد التنبّؤات النهائيّة، وتكون هذه الخوارزميات أكثر دقّة من الطّرق الأخرى ولكنها ذات تكلفة حسابية عالية، وبالتالي تُستخدم للتطبيقات التي تكون الدقّة عاملاً أساسيّاً ، ويبين الشكل (1-5) أنواع الكشف عن الأغراض



الشكل (1-5)- يبين أنواع الكشف عن الأغراض

يرمز مصطلح YOLO إلى "You Only Look Once" أي أنت تنظر مرةً وحيدة وهذا يدل على آلية عملها فهي تقوم بالمعالجة وفق مرحلة وحيدة single-shot detectors، وتتضمن هيكليّة نموذج هذه الخوارزمية ثلاثة أجزاء هي البنية المركزية (العمود الفقري)، العنق والرأس، وهي شبكة عصبونية ذات اتّصال كامل FCNN مكوّنة من شبكة عصبونية التفاضلية (CNN) بسيطة لكشف الأغراض، حيث يتم تقسيم الصورة إلى عدّة صناديق بأبعاد محدّدة ثمّ تكتشف كلّ خلية الكائنات التي تحتويها وتحدّد موقعها مع إحداثيات المربع المحيط (بالنسبة لإحداثياتها)، مع تسمية الكائن واحتمال وجود الشيء في الخليّة

تتطلب خوارزمية مروراً واحداً فقط (انتشار أمامي) عبر الشبكة العصبونية وتكون التفاضلية لاكتشاف كائنات متعدّدة داخل الصورة، بحيث يتم تقسيم الصورة إلى مناطق ويتمّ التنبؤ بمربع الإحاطة Bounding box والاحتمالات لكلّ منطقة. تُحدّد الأجسام وتُصنّفها بمرور واحد فقط من الصورة. بمعنى آخر، ينظر النموذج إلى الصورة مرة واحدة فقط، ومن خلال هذا "المرور الواحد" يتمكن من تحديد الأجسام فيها. لا يقتصر كشف الأجسام على تحديد موقع الجسم في الصورة فحسب، بل يحدّد ماهيته أيضاً. وتعمل بتمرير الصورة عبر شبكة عصبونية تُمكن النموذج من اكتشاف جميع الأجسام في آنٍ واحد. وباستخدام التنبؤ الشبكي والتنبؤ بالصندوق المحيط، يستطيع النموذج فهم ما إذا كانت الأجسام تقع ضمن صندوق محيط معين.

تتميز تقنية YOLO بالعديد من المزايا مقارنةً بالطرق السابقة. فهي سريعة للغاية، وبالتالي يمكن تطبيقها في مطاردات مثل السيارات ذاتية القيادة وأنظمة المراقبة بالفيديو. كما تتميز بدقة عالية، خاصةً مع الصور الطبيعية، ونتائج إيجابية خاطئة أقل، إذ تُحلّل الصورة بأكملها دفعةً واحدة، مما يوفر دقة سياقية أكبر. وتستخدم YOLO في المركبات ذاتية القيادة بسبب الميزات التالية:

1. الكشف الفوري عن الأجسام (Real-Time Detection)

- المركبات ذاتية القيادة تحتاج إلى معالجة تدفقات الفيديو بسرعة عالية جداً (جزء من الثانية) لاتخاذ قرارات آمنة وسريعة.
- خوارزمية YOLO تمتاز بكونها أحادية المرحلة (Single-Stage Detector)، مما يجعلها قادرة على العمل بسرعة كبيرة مقارنة بالخوارزميات التقليدية مثل Faster R-CNN.

2. الكشف عن أجسام متعددة في وقت واحد

- السيارة الذاتية القيادة يجب أن تتعرف على أنواع مختلفة من الأجسام في المشهد المروري مثل: السيارات، المشاة، إشارات المرور، الدراجات الهوائية، وعلامات الطريق.
- YOLO قادرة على التعرف على أجسام متعددة بفئات مختلفة في صورة واحدة، وهو ما يجعلها مناسبة للبيئات المرورية المعقدة.

3. موازنة السرعة والدقة

- في التطبيقات الحساسة مثل القيادة الذاتية، لا يمكن التضحية بالسرعة أو الدقة.
- YOLO توفر توازنًا جيدًا بين الاثنين، مما يجعلها أكثر ملاءمة من الخوارزميات البطيئة أو الخفيفة جدًا قليلة الدقة.

4. اعتماد واسع في الأبحاث والتطبيقات الصناعية

- تم استخدام YOLO بكثرة في أنظمة المساعدة على القيادة (ADAS) وفي أنظمة الرؤية الحاسوبية الخاصة بالسيارات، مما يعزز موثوقيتها كخيار مناسب للأبحاث الأكاديمية والتطبيقات العملية.

1-6-2 خوارزمية YOLOv8n

تعتبر خوارزمية YOLOv8n نسخة لنموذج YOLOv8 وهي مصممة لاكتشاف الأغراض وتجزئة الصور بشكل فعال ودقيق. يمكن تدريب YOLOv8n على مجموعات بيانات كبيرة ويعمل بشكل جيد على العديد من منصات الأجهزة (hardware platforms)، سواء وحدات المعالجة المركزية (CPUs) أو وحدات معالجة الرسومات (GPUs). مقارنة بالإصدارات الأخرى من النموذج، تم إجراء تحسينات كبيرة في أدائها من حيث السرعة والدقة والهندسة المعمارية. أسرع وأكثر دقة من (YOLOv7)، وصلت إلى مستوى مرتفع جديد بمتوسط درجة دقة بلغ 53.7%. وتستخدم YOLOv8n في المركبات ذاتية القيادة بسبب الميزات التالية :

1. خفيف وفعال (Lightweight & Efficient)

- النسخة YOLOv8n (Nano) هي الأخف بين إصدارات YOLOv8 ، مصممة لتعمل بكفاءة على الأجهزة محدودة القدرة الحاسوبية مثل الحواسيب المدمجة داخل السيارات (Embedded Systems) .

2. التجريب والتطوير السريع

- YOLOv8n تسمح بتجريب النماذج وتدريبها بسرعة، وهو أمر مهم في الأبحاث الأكاديمية حيث نحتاج إلى اختبار عدة سيناريوهات وبيانات ، مثل مجموعات بيانات القيادة KITTI ، (BDD100K) .

3. تحسين التعرف على الأجسام الصغيرة والمتداخلة

- بفضل التصميم الجديد في YOLOv8 آلية Anchor-Free والرأس المنفصل Decoupled Head ، أصبح النموذج أكثر قدرة على اكتشاف الأجسام الصغيرة مثل إشارات المرور البعيدة أو المشاة في الخلفية، وهو أمر بالغ الأهمية في القيادة الذاتية.

4. قابلية التوسع (Scalability)

- يمكن البدء باستخدام YOLOv8n كنموذج خفيف وسريع للتجارب الأولية، ثم الانتقال لاحقًا إلى إصدارات أكبر YOLOv8s (m ، l ، x) لزيادة الدقة في مرحلة التطبيق الصناعي.

5. العمل بكفاءة على الفيديو المباشر متعدد الكاميرات

- المركبات ذاتية القيادة تعتمد غالبًا على عدة كاميرات، و YOLOv8n قادرة على معالجة هذه البيانات بكفاءة في الزمن الحقيقي.

الفصل الثاني

الدراسات المرجعية

- الدراسات المرجعية لمساعدة السائق
- الدراسات المرجعية لكشف مشاكل الطرق
- الدراسة المرجعية لمساعدة ذوي الحاجات الخاصة
- الدراسة المرجعية السابعة

الدراسات المرجعية

1.2- مقدمة

أظهرت الدراسات النظرية أن مسألة تحسين أداء المركبات ذاتية القيادة من حيث الدقة والسرعة مما يستدعي تطوير عدد من خوارزميات التعلم العميق واقتراح طرائق متعددة لتحسين أداء خوارزميات كل من YOLO و Faster R-CNN وغيرها ، ويستعرض هذا الفصل عدد من الدراسات المرجعية لتحسين أداء المركبات ذاتية القيادة للكشف عن الأغراض ومزايا كل منها وعيوبه. VSLAM أن مسألة ضبط الحزمة تشكل عنق الزجاجة لخوارزميات التتبع أداء لمتوضع ووضع الخرائط المرئي (

2.2- الدراسات المرجعية لمساعدة السائق

يوجد عدد كبير من الدراسات المرجعية منها :

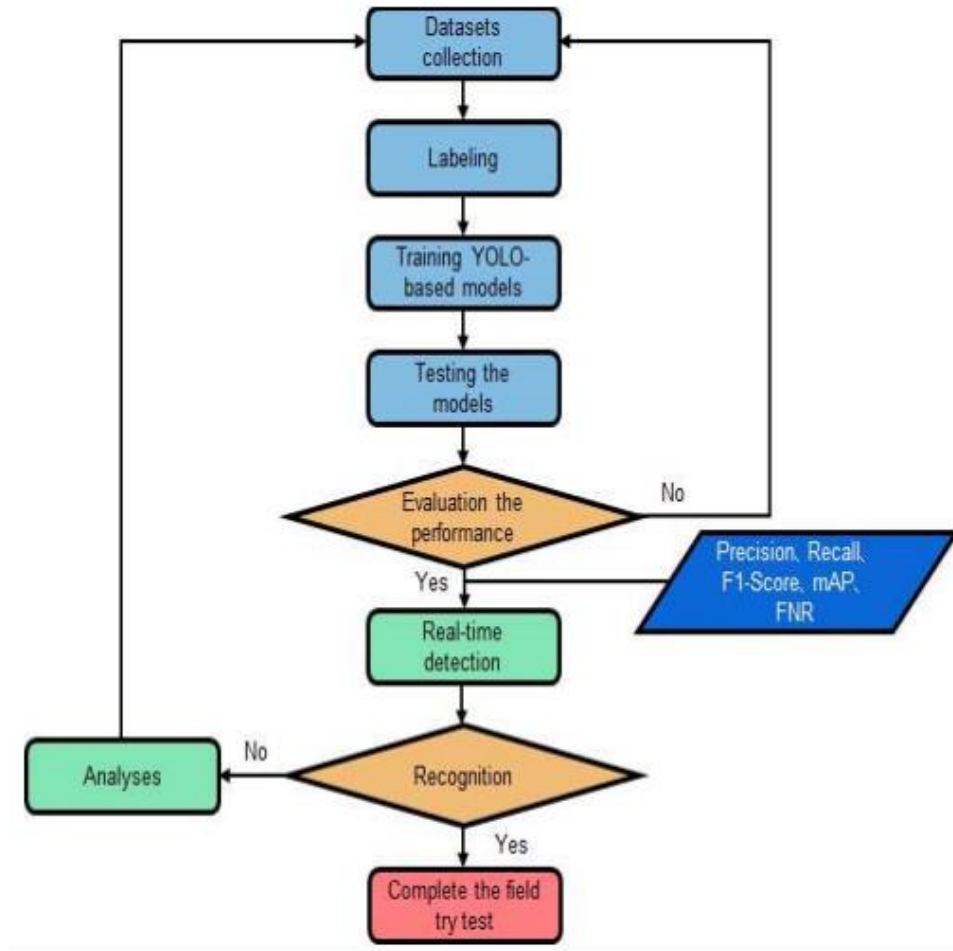
1-تقنية اكتشاف السلوك المشتت للسائق مع شبكات التعلم العميق المعتمدة على YOLO

Driver Distracted Behavior Detection Technology with YOLO-Based Deep Learning Networks[1].

يتم استخدامه تقنية التعلم العميق المعتمدة على YOLO من خلال إعداد كاميرات ويب على لوحة القيادة لاكتشاف سلوكيات السائق. وبواسطة صور قناة RGB كمدخلات ، فإن ملف نماذج التعلم العميق المعتمدة على YOLO ، وتدريبهم كأجهزة كشف مرشحة. فإن السلوكيات المكتشفة تتضمن القيادة العادية ، دوران الرأس المشتت ، النعاس ، الأكل ، والتحدث على الهاتف ، وما إلى ذلك فإن النتائج التجريبية عند تعيين نفس البارامترات نلاحظ أن YOLO-fastest-xl لديها أفضل أداء مع مجموعة بيانات متعددة الفئات 91.84% ، والمعدل السلبي للخطأ (FNR) 6.94% ، و mAP هي 95.81% .

تم استخدام نموذج YOLO-fastest لأنه الأسرع من حيث الأداء ، وكمية البارامترات والحسابات المستخدمة . على الرغم من أن YOLO-fastest خفضت مقدار البارامترات والحسابات إلى حد كبير . يبين الشكل (2-1) التصميم المقترح .إن قواعد البيانات المستخدمة في هذا العمل مقسمة إلى فئتين. مجموعات البيانات العامة المستخدمة هي قاعدة بيانات Pointing'04 ، وقاعدة

بيانات الوجه FEI و قاعدة بيانات WIDER FACE . ويسرد الجدول (1-2) أرقام مجموعة البيانات المستخدمة .



يبين الشكل (1-2) التصميم المقترح [1]

الجدول (1-2)- يسرد أرقام مجموعة البيانات المستخدمة

Datasets	Numbers of images
Self-collected driving distractive behavior images	32917
Pointing'04 [15]	2819
FEI face [16]	2695
WIDER FACE [17]	547
Total numbers	38978

لتصنيف سلوكيات القيادة ، من قبل توليد بيانات التدريب ، نحدد سلوكيات القيادة إلى خمس فئات و يحدد الجدول (2-2) سلوكيات القيادة المستخدمة في هذا العمل. وتعريف السائق تشمل السلوكيات المشتتة للانتباه تشتت الذهن ، والنعاس ، والأكل ، والتحدث بالهاتف. الغرض من تعريف قيادة سلوكيات هي الحكم بشكل صحيح على السلوكيات المشتتة للانتباه السائقين ، والحد من وقوع الحوادث ، وتحسين سلامة القيادة.

الجدول (2-2) - يحدد سلوكيات القيادة المستخدمة

Five Categories		Definitions
Good		With the person's front face as the starting axis, the horizontal and vertical rotations are within 30 degrees and the eyes are open
Distracted behaviors	Warning	Taking the person's front face as the starting axis, rotate more than 30 degrees horizontally and vertically
	Drowsy	The eyeball is covered by the eyelid by more than 50%
	Eat	Holding food close to the driver's mouth
	Phone	Holding the smartphone close to driver's ears

بعد جمع مجموعات البيانات ، وتعريف سلوك القيادة متعدد الفئات ، وتصنيف فئة السلوك ، نقسم مجموعات البيانات المجمعة إلى أجزاء التدريب والتحقق من الصحة والاختبار. و يسرد الجدول (2-3) عدد الصور المستخدمة في أوضاع التدريب والتحقق والاختبار. ويصف الجدول (2-4) النتائج التجريبية بين YOLO-fastest وYOLO-fastest-x. ويبين الجدول (2-5) مقارنة أداء F1-Score و FNR و mAP بين النماذج الأربعة المستندة إلى YOL

الجدول (2-3)- يسرد عدد الصور المستخدمة في أوضاع التدريب والتحقق والاختبار

Five Categories	Training	Validation	Testing
Good	1605	324	316
Warning	13397	1750	1872
Drowsy	5599	455	650
Eat	9706	913	1681
Phone	8506	1596	789

الجدول (4-2) - يصف النتائج التجريبية بين YOLO-fastest وYOLO-fastest-xl

Parameters setting	YOLO-fastest	YOLO-fastest-xl
IoU thresh	0.50	
Confidence thresh	0.40	
Evaluation indices		
Precision	90.97%	90.66%
Recall	92.60%	93.06%
F1-Score	91.78%	91.84%
FNR	7.40%	6.94%
Average Precision (AP)		
Good	95.30%	96.19%
Warning	98.19%	98.44%
Drowsy	92.53%	93.97%
Eat	94.56%	94.39%
Phone	96.15%	96.07%
mAP	95.35%	95.81%

الجدول (5-2) - يبين مقارنة أداء F1-Score وFNR و mAP بين النماذج الأربعة المعتمد على YOL

CNN Models	F1-Score	FNR	mAP
YOLOv3-tiny	89.91%	12.22%	95.10%
YOLO-v3tiny-3l	90.40%	11.68%	95.67%
YOLO-fastest	91.78%	7.40%	95.35%
YOLO-fastest-xl	91.84%	6.94%	95.81%

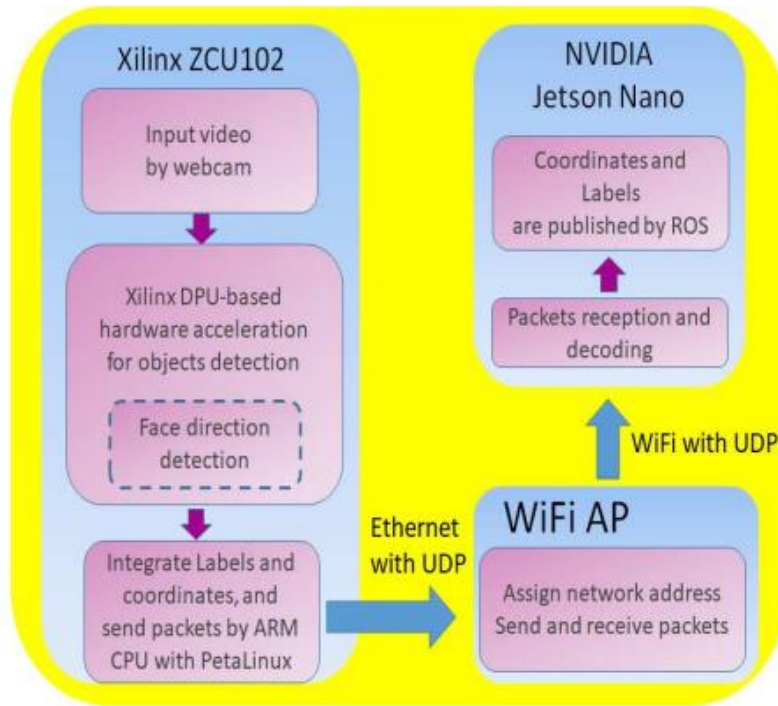
تظهر النتائج التجريبية أن YOLO-most-xl لديها أفضل أداء مع مجموعات بيانات متعددة الفئات ، و F1-Score و FNR و mAP هي 91.84% و 6.94% و 95.81% على التوالي. لزيادة الدقة ، يمكن تحسين أداء النموذج الكلي من خلال تعديل البارامترات وتعديل هيكل النموذج. وبالنسبة لمجموعات البيانات ، سيتم تعزيز مجموعات بيانات التدريب من خلال جمع المزيد من الصور ذات الصلة ، ويمكن أيضًا توسيع وظائفها عن طريق إضافة المزيد من الفئات. وبالتالي يمكن تحقيق اكتشافات سلوكيات تشتت انتباه القيادة أكثر شمولاً .

إذا كان رأس السائق يتأرجح لأعلى ولأسفل ، فسيحكم عليه النظام على أنه حالة "تحذير" ؛ بالنسبة لانحراف وجه السائق ، نظرًا لأن عملية التدريب لا تجمع بيانات عن حالة انحراف الوجه ، فلن يتمكن النظام من اكتشاف انحراف الوجه بشكل فعال.

2-استخدام خوارزمية YOLO المعتمدة على شبكة الالتفافية عميقة للمركبات ذاتية القيادة Hardware Acceleration of YOLO-based Convolutional Neural Network Detector by Deep Learning Processor Unit for Intelligent Autonomous Mover[2].

بالنسبة للمركبات ذاتية القيادة الذكية ، فإنها تلعب دورًا مهمًا في الاستخدام السهل للمستخدم في البيئات المزدهمة ، لذلك كان التركيز مؤخرًا على توفير وظائف التعرف على الأشياء وتجنب الاصطدام. في هذه الدراسة ، تم تنفيذ نموذج الشبكة العصبونية الالتفافية (CNN) المستندة إلى وحدة معالجة التعلم العميق (Xilinx (DPU لاكتشاف الكائنات وتصنيفها ، وتم استخدام تسريع الأجهزة المستندة إلى FPGA لاكتشاف المشاة والأشياء الأخرى بكفاءة. من خلال التنفيذ المستند إلى DPU (Deep Learning Processor Unit) ، تصل دقة تصنيف الكائن المقترح ونموذج اكتشاف اتجاه الوجه إلى 90% على مجموعات بيانات الصور التي تم جمعها في محلات السوبر ماركت. مقارنةً بتطبيقات البرامج السابقة المستندة إلى وحدة معالجة الرسومات ، يوفر النهج المقترح أداءً إطراراً أفضل في الثانية (FPS) للتطبيق في الوقت الفعلي .

يوضح الشكل (2-2) تدفق المعالجة لتصميم الكشف المطبق المعتمد على YOLO. أولاً ، على النظام الأساسي Xilinx ZCU102 ، يتم دمج رسائل الكائن المكتشفة التي يتم استنتاجها بواسطة تسريع الأجهزة المستند إلى DPU ، ثم يتم إرسال حزم التسميات ومعلومات الإحداثيات للكائنات المكتشفة إلى جهاز Nano المستند إلى وحدة معالجة الرسومات عن طريق نشرها باستخدام ROS (robotic operating system) عبر WiFi AP ، حيث تتضمن معلومات الكائن المنشورة باستخدام ROS: "Label" و "x_min" و "x_max" و "y_min" و "y_max" و "ID".



الشكل (2-2)- يوضح تدفق المعالجة لتصميم الكشف المطبق المعتمد على YOLO

يوضح الجدول (2-6) مقارنة بين التطبيقات المختلفة لاكتشاف CNN المستندة إلى YOLO للتطبيقات المتحركة الذكية المستقلة. في هذا الجدول تصل دقة الكشف لنموذج Lite-CNN المقترح إلى 90%. علاوة على ذلك ، يوفر النهج المقترح حوالي 4x إطارات أكثر في الثانية (FPS) مقارنةً بتسريع البرامج المعتمد على وحدة معالجة الرسومات في ، ويحقق التنفيذ المقترح 30 إطارًا في الثانية في تطبيقات الوقت الفعلي. لتسريع الأداء ، يوفر نموذج Lite-CNN المستند إلى YOLO دقة عالية لاكتشاف أجسام المشاة وتصنيف توجيه الوجه مع تقليل التعقيد الحسابي.

الجدول (2-6)- يوضح مقارنة بين التطبيقات المختلفة لاكتشاف CNN المستندة إلى YOLO

	Direct software design with GPU [6]	GPU-based acceleration by two-stage design [6]	The proposed DPU-based design by FPGA
FPS	2.8 ~ 2.7	7.9 ~ 7.5	~30
Speed-up	1	2.85	11

3-تحسين تعلم المركبات ذاتية القيادة استناداً إلى سلوك القيادة الحقيقي باستخدام تقنيات الشبكات العصبونية العميقة

Improving the learning of self-driving vehicles based on real driving behavior using Deep Neural Network Techniques[21]

الفكرة الرئيسية في هذه الورقة البحثية تطوير حل للقيادة الذاتية لمنصة سيارات خفيفة ذات موارد أجهزة محدودة وقوة المعالج و حجم الذاكرة ، مع أخذ قيود الأجهزة في الاعتبار ، والهدف هو تصميم شبكة عصبونية عميقة خفيفة DNN وهي شبكة عصبية شاملة ستكون قادرة على أداء مهمة القيادة الذاتية على المسار التمثيلي ، في حين أن الشبكات المتقدمة من الممكن نشر النموذج المستخدم للاستدلال على النظام الأساسي للأجهزة منخفضة الأداء ، تتكون هذه الورقة من مرحلتين رئيسيتين:

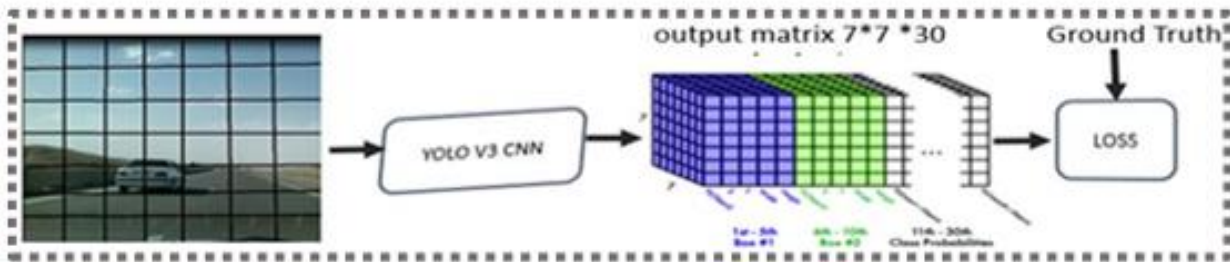
✓ **المرحلة الأولى** هي اكتشاف العوائق المحيطة بالمركبة . لقد أخذنا بعين الاعتبار اكتشاف العوائق المختلفة المتحركة والثابتة بما في ذلك المشاة والمقاعد والمركبات المتوقفة والمركبات المتجاوزة و ما إلى ذلك ، و في بعض الأحيان لا يكون لدينا أي عوائق ونريد اكتشاف الطرق والانعطاف يساراً و يميناً أثناء التحرك .

✓ في المرحلة الثانية يتم تطبيق أساليب القيادة الحقيقية على المركبات ذات القيادة وفقاً للمعالجة التي تم إجراؤها في المرحلة الأولى والتي تشمل اكتشاف العوائق والطرق السريعة من أجل تحسين التعلم في المركبات ذات القيادة . بعد ذلك ، يتم تدريب الشبكة وتوليد زاوية التوجيه والتسارع والفرامل كمخرجات.

نقوم بتصميم تعلم مخطط المركبات ذاتية القيادة من خلال استخدام تقنيات الشبكة العصبونية العميقة. ويتكون النظام المقترح من ثلاث مراحل ، نظرة عامة على نموذج ، LSV-DNN تمت مناقشتها في المرحلة الأولى وفي المرحلة الثانية تمت مناقشة اكتشاف العوائق المحيطة بالمركبة. وسيتم مناقشة أساليب القيادة الحقيقية للقسم التالي المطبقة على المركبات ذاتية القيادة في المرحلة الثالثة .

يمكن استخدام خوارزمية YOLO السريعة لاكتشاف الأشياء في الوقت الفعلي أثناء القيادة. تقوم هذه الخوارزمية بتنفيذ المعالجة لتحسين وقت المعالجة والسرعة العالية . تستخرج هذه الخوارزمية

الميزات من الصورة ثم تحصل على معلومات الكائن وتحللها مثل الفئة والموقع والاتجاه. تقوم خوارزمية YOLO باستخراج واستخدام ميزات الرسم البياني بأكمله . وبهذه الطريقة ، فإنه يقلل بشكل كبير من الخطأ مقارنة بالكشف المحلي . كما يقوم أيضاً بتدريب الميزات المهمة بدقة ومعلومات الكائن بالكامل على مستوى العالم. ويبين الشكل (2-3) الهيكل العام لنظام YOLO إلى جانب المدخلات والمخرجات .



الشكل (2-3) - يوضح الهيكل العام لنظام YOLO إلى جانب المدخلات والمخرجات

سيتم تدريب الشبكة الالتفافية باستخدام البيانات التي تم جمعها أثناء القيادة. تم تنفيذ مخرج الشبكة وهو زاوية دوران عجلة القيادة والتسارع والغاز والفرامل بزمان إخراج قليل وخسارة 0.650 خطأ في الشبكة . وفي المشروع الثاني، تم اكتشاف العوائق في الشبكة الالتفافية المصممة باستخدام ثلاث خوارزميات YOLOv3 و Resnet101 على بيانات KITTI تم الحصول على أفضل النتائج بواسطة خوارزمية YOLOv3 مع $IOU=0.98$ ، ولتحقيق هذه الغاية ، قام الباحث بتصميم النموذج المقترح باستخدام خوارزمية YOLOv3 التي تقدم نتائج أفضل. في هذا النموذج تكتشف خوارزمية YOLOv3 العوائق المحيطة بالمركبة من خلال خمس فئات للشاحنات والمركبات والمشاة والشاحنات الصغيرة وراكبي الدراجات . تكتشف هذه الخوارزمية الأشياء المحيطة بالمركبة بسرعة وبدقة عالية وتقدم النتائج الى الشبكة الالتفافية المصممة بحيث يمكن تدريب الشبكة بهذه الطريقة ، يمكننا الحصول على زاوية عجلة القيادة ، والتسارع ، والغاز ، والفرامل بسرعة و دقة عالية ، في المشروع التالي، قام الباحث بتنفيذ مشروع آخر على البيانات التي جمعها. أولاً، تم إجراء الكشف عن العوائق باستخدام خوارزمية YOLO على البيانات التي تم جمعها أثناء القيادة باستخدام $IOU=0.98$ وكان خطأ الشبكة منخفضاً. لقد حصل الباحث على نتائج جيدة جداً وقدمت الشبكة زاوية عجلة القيادة وتسارع الغاز والفرامل بعد التدريب.

وفي الخطوة الأخيرة أعطى الباحث مخرجات الشبكة المصممة و هي زاوية دوران عجلة القيادة والتسارع الغازي والفرامل لنظام التحكم في المركبة ذاتية القيادة لتتمكن المركبة من الاستجابة للعوائق المحيطة حسب الاستجابات لديها ويتم تطبيق تسارع الغاز المستفادة والفرامل وزاوية الدوران.

أظهر الباحث أن الشبكات الالتفافية قادرة على التعلم ويمكنها تقديم نتائج جيدة في الإخراج وفقاً لزاوية دوران عجلة القيادة وتصرفات السائق الحقيقي. كما يمكنهم أيضاً محاكاة تصرفات السائق في الشبكة، وأخيراً، تمكن الباحث من تنفيذ التخطيط للمسار على الطرق الصعبة. قام الباحث بتدريب شبكة التفافية عميقة لاتخاذ القرار الأفضل أمام السيارة. أكدت النتائج أن مخطط الباحث قادر على إظهار دقة تنبؤ عالية (تتجاوز 93.92%) بالإضافة إلى ذلك، فإن المخطط المقترح يتميز بسرعة عالية (أكثر من 41.64%)، ونسبة FPR منخفضة (أقل من 89.6%)، ونسبة FNR منخفضة أقل من (95.3%)، مقارنة بالطرق الأخرى المستخدمة حالياً.

3.2- الدراسات المرجعية لكشف مشاكل الطرق

يوجد العديد من الدراسات التي عالجت مشاكل حالة الطرق منها :

1-نظام كشف الحفر وتقدير الأبعاد باستخدام (YOLO) ومعالجة الصور

Pothole Detection and Dimension Estimation System using Deep Learning (YOLO) and Image Processing[3].

وفقاً للبيانات الرسمية الصادرة عن حكومة الهند فإن الحفر تشكل تهديداً كبيراً حيث فقد 2015 شخصاً حياتهم في 4869 حادثاً تسببت فيه الحفر في العام الماضي، لذلك يعتبر ترميم الطرق ذو أهمية عالية و لكن جائحة كورونا COVID_19 أعاقَت العديد من القطاعات بما في ذلك صيانة الطرق حيث من الصعب خلق بيئة عمل آمنة للعمال وأدى ذلك إلى تدهور حالة الطرق ومن هنا كان الحاجة إلى نظام مستقل يمكنه مراقبة ظروف الطريق لكشف الحفر و تقدير أبعادها . يستخدم النظام المقترح خوارزمية تعتمد على التعلم العميق YOLO لاكتشاف الحفر في الطرقات لتحقيق السلامة المرورية . أما مجموعة البيانات تتكون من صور لحفر تحتوي مياه و حفر جافة

بأشكال و أحجام مختلفة و في طرق ببلدان مختلفة . حيث قام الباحثون في هذه الورقة البحثية باقتراح نموذج بهدف التخلص من العيوب في بعض الأنظمة عند الكشف عن الحفر نذكر منها :

- النهج القائم على مقياس التسارع : حيث مقياس التسارع هو حساس يقيس التسارع أو التغيرات في السرعة لجسم ما ، في حالتنا يمكن استخدام مقياس التسارع لقياس الاهتزازات التي تحدث عندما تتحرك المركبة ذاتية القيادة فوق الحفر و بالتالي عيب هذا النهج أن السائقين يميلون إلى الإبطاء فوق الحفر و أيضا سيتم تسجيل المطبات على أنها حفر و بالتالي البيانات التي جمعها لن تكون دقيقة

- نظام يقوم بكشف الحفر بناءً على تغيرات الألوان في الصورة و يتمثل عيب هذا النظام في أنه لا يمكن دائما وصف تغيرات اللون على الطريق على انه حفرة حيث يؤدي تغير اللون أيضا لأسباب أخرى مثل الطرق المبللة و علامات الطريق و ممر المشاة .

لذلك بعد إجراء التحسينات على الأنظمة السابقة و غيرها تم اعتماد خوارزمية YOLOv4 حيث إنها خوارزمية تنبؤ أكثر كفاءة من حيث زمن التدريب و كذلك الدقة و لكن سيتم المقارنة بين خوارزميات YOLOv3 و YOLOv4 على أساس قيم MAP و IOU للتكرارات المتعددة iterations . حيث ينقسم النظام المقترح إلى مرحلتين : مرحلة كشف الحفر ومرحلة تقدير الأبعاد .

1- مرحلة الكشف

السبب في اختيار عائلة YOLO أنه يتم توقع المربعات المحيطة بالأجسام مباشرة من صورة الإدخال ولذلك هي مناسبة للتطبيقات الفعلية، على نقيض خوارزميات الكشف عن الأجسام القائمة على العمارة ذات المرحلتين والتي تتطلب :

1- اقتراح منطقة الكشف : حيث يتم إنشاء أكثر من اقتراح لمكان تواجد الأجسام و احاطتها بمربعات

2- استخراج الميزات من المناطق المقترحة لتصنيف الأجسام

ولكن مع ذلك فإن ال YOLOv3 تتنازل قليلاً عن السرعة لزيادة دقة الشبكة بشكل عام بينما تهدف ال YOLOv4 إلى اكتشاف الأجسام بسرعة و كفاءة ، ولتحقيق هذا الهدف تم تحسين YOLOv4 للحساب المتوازي ويعود الأداء الأفضل ل YOLOv4 على YOLOv3 بأن ال YOLOv4 تستخدم استراتيجيات مثل : Bag of freebies و Bag of specials

- Bag of freebies : تتضمن زيادة البيانات (حيث تم إنشاء بيانات تدريب جديدة من بيانات التدريب الموجودة وذلك عن طريق تطبيق تحويلات مثل التدوير والقياس والتقليب) وإضافة قدر صغير من الضوضاء حتى لا تقوم الشبكة بحفظ البيانات وعدم التنبؤ ثم تقوم بتعديل معدل التعلم

- Bag of specials : أحد التقنيات المستخدمة هنا هي تجمع الهرم المكاني (SPP) spatial pyramid pooling حيث يتم تقسيم خريطة المعالم إلى شبكة وتجميع كل خلية شبكة بشكل منفصل لالتقاط المعلومات بمقاييس مختلفة، ويسمح هذا للشبكة باكتشاف الأجسام بأحجام مقاييس مختلفة في الصورة .

تم تدريب كلا النموذجين YOLOv3 و YOLOv4 على مجموعة البيانات المخصصة وبتكرار (3000 ، 4000 ، 5000 ، 6000) وحجم الدفعة (عدد العينات التي يتم معالجتها في كل تكرار لعملية التدريب) هو 64 وعدد الأقسام الفرعية (عدد الدفعات الأصغر التي يتم إنشاؤها من الدفعة الأكبر أثناء التدريب ويساعد هذا في تقليل متطلبات الذاكرة لعملية التدريب) هو 8 خلال عملية التدريب . ويعود الأداء الأفضل ل YOLOv4 على YOLOv3 بأن ال YOLOv4 تستخدم استراتيجيات مثل Bag of freebies و Bag of specials

- Bag of freebies : تتضمن زيادة البيانات (حيث تم إنشاء بيانات تدريب جديدة من بيانات التدريب الموجودة وذلك عن طريق تطبيق تحويلات مثل التدوير والقياس والتقليب) وإضافة قدر صغير من الضوضاء حتى لا تقوم الشبكة بحفظ البيانات وعدم التنبؤ ثم تقوم بتعديل معدل التعلم

- Bag of specials : أحد التقنيات المستخدمة هنا هي تجمع الهرم المكاني (SPP) spatial pyramid pooling حيث يتم تقسيم خريطة المعالم إلى شبكة وتجميع كل خلية شبكة بشكل منفصل لالتقاط المعلومات بمقاييس مختلفة، ويسمح هذا للشبكة باكتشاف الأجسام ذات الأحجام المقاييس المختلفة في الصورة .

2- وحدة تغيير الأبعاد

يمكننا اعتبار الجسم (الحفرة) في أرض الواقع وصورته المأخوذة بواسطة الكاميرا كمثلثين متشابهين، الجسم (الحفرة) في العالم الحقيقي له عرض معين "W" ويقع على مسافة "D" من الكاميرا، عندما

نلتقط صورة للجسم (الحفرة) نحصل على صورة للحفرة بعرض واضح بالبيكسل "P" والآن يمكننا اعتبار الحفرة وصورتها كمثلثين متشابهين، حيث الأضلاع المقابلة هي العرض الفعلي "W" والمسافة "D" للحفرة في أرض الواقع والعرض الظاهري "P" والمسافة "F" لصورة الكائن. وبالتالي باستخدام خاصية التشابه الثلاثي، يمكننا كتابة المعادلة التالية:

$$\frac{W}{D} = \frac{P}{F}$$

عند عزل "F" نحصل على:

$$F = \frac{P * D}{W}$$

وبين الجدول (7-2) نتائج المقارنة بين YOLOv3 و YOLOv4 .

الجدول (7-2) - يوضح نتائج المقارنة بين YOLOv3 و YOLOv4

Model	Iterations	mAP @ 0.50	IoU
YOLOv3	3000	0.860	0.592
	4000	0.889	0.635
	5000	0.875	0.612
	6000	0.861	0.592
YOLOv4	3000	0.932	0.727
	4000	0.933	0.741
	5000	0.888	0.717
	6000	0.869	0.716

نلاحظ من الجدول أن YOLOv4 أفضل في الأداء من YOLOv3

2-تطبيق متقدم للكشف عن الحفر باستخدام YOLO

Pothole detection and dimension estimation using in-vehicle technologies and deep learning[4]

اقترحت هذه الدراسة استخدام تقنيات المركبات المدمجة للكشف عن قياسات الحفر وتقديرها. تضمنت عملية تقدير النموذج تدريب 1876 صورة حفر ملونة ثنائية الأبعاد عادية والتنبؤ بقياسات التقدير. تم إعداد نموذج الكشف باستخدام خوارزمية YOLOv5 وتم تدريبه على Google Colab. تم شرح صور تدريب الحفر يدويًا نظرًا لعدم وجود كشف مدرب لضائقة الرصيف للاستخدام في نماذج YOLO. حقق النموذج المدرب 93.0% و 91.6% و 96.3% و 87.0% في الدقة والاسترجاع ومتوسط الدقة ومقاييس F1 على التوالي. علاوة على ذلك، استخدمت الورقة خطوط Hough في مكتبة OpenCV Python للكشف عن علامات خطوط الرصيف وحساب قياسات الحفر. استخدمت هذه الدراسة أنظمة مساعدة الحفاظ على المسار (LKA) القياسية في العديد من المركبات الحديثة لتحديد مقاييس قياس الحفر. وقد أنشأت هذه الورقة مقياسي عرض وطول من قياسات معروفة لجهازي الكشف الحلقي في مجموعة البيانات نفسها. وأشارت نتائج المقارنة بين الأحجام الفعلية والمقدرة إلى أن النموذج المقترح قادر على تقدير حجم الحفر بنسبة خطأ 2.30%. وتشير هذه النتيجة إلى أن تقنيات المركبات المدمجة، مثل الكاميرا ونظام مساعدة الحفاظ على المسار، يمكن أن توفر طرقًا اقتصادية لتقدير أبعاد الحفر. ويمكن لهذه الدراسة أن تساعد هيئات الطرق على جمع بيانات الحفر بتكلفة منخفضة وبشكل فوري من خلال الاستفادة من تقنيات المركبات المدمجة المتوفرة في العديد من المركبات القياسية.

ومن سلبياتها الاعتماد على تغيرات الألوان التي قد تتأثر بعوامل خارجية مثل الطرق المبللة أو علامات الطريق ، وصعوبة التمييز بين الحفر الحقيقية والتغيرات اللونية المشابهة. ويبين الجدول (2-8) حلقات الكشف كما تم تقديرها بواسطة النموذج ، أما الشكل (2-4) يبين أبعاد حلقات الكشف كما تم تقديرها بواسطة النموذج .



الشكل (4-2) - يبين أبعاد حلقات الكشف كما تم تقديرها بواسطة النموذج

الجدول (8-2) - يبين حلقات الكشف كما تم تقديرها بواسطة النموذج

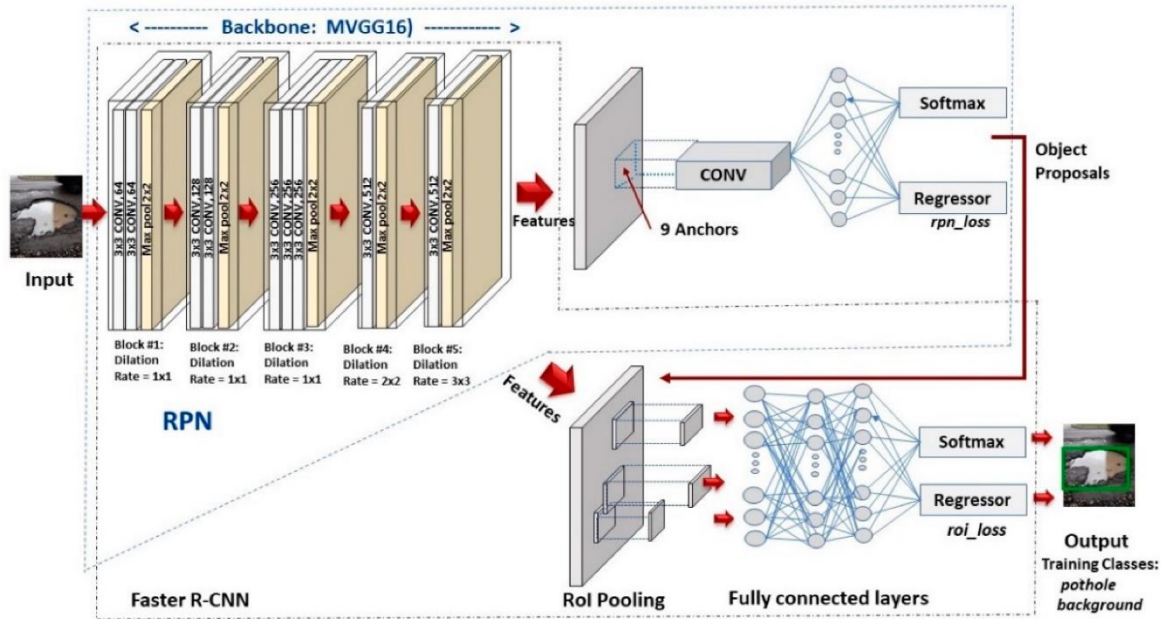
S/N	Metric	Value
1.	Precision	93.0
2.	Recall	91.6
3.	F1-score	87.0
4.	mAP	96.3

3- مقارنة بين نماذج YOLO المختلفة لكشف الحفر

Smart Pothole Detection Using Deep Learning Based on Dilated Convolutional Neural Networks Pothole detection and dimension estimation using in-vehicle technologies and deep learning[11]

تُعدّ الحفر في الطرق أحد أهمّ المخاوف في البنية التحتية للنقل. وقد اقترحت العديد من الأبحاث استخدام تقنيات الرؤية الحاسوبية لأتمتة اكتشاف الحفر، والتي تشمل مجموعة واسعة من خوارزميات معالجة الصور واكتشاف الأجسام. هناك حاجة لأتمتة عملية اكتشاف الحفر بدقة وسرعة مناسبة، وتنفيذ العملية بسهولة وبتكلفة إعداد منخفضة. في هذه الورقة، تم تطوير شبكات عصبونية التلافيفية

(CNN) فعالة للتعلم العميق للكشف عن الحفر في الوقت الفعلي بدقة كافية. ولتقليل التكلفة الحسابية وتحسين نتائج التدريب، تقترح هذه الورقة شبكة (MVGG16) (VGG16 معدلة، وذلك بإزالة بعض طبقات الالتفاف واستخدام معدلات تمدد مختلفة. علاوةً على ذلك، تستخدم هذه الورقة شبكة MVGG16 كشبكة أساسية لشبكة R-CNN الأسرع. بالإضافة إلى ذلك، يُقارن هذا العمل أداء نماذج YOLOv5 (الكبيرة (Yl)، والمتوسطة (Ym)، والصغيرة (Ys)) مع الشبكات الأساسية ResNet101، وشبكة R-CNN الأسرع مع الشبكات الأساسية (ResNet50(FPN)، VGG16، و MobileNetV2، و InceptionV3، و MVGG16. تُظهر النتائج التجريبية أن نموذج Ys أكثر قابلية للتطبيق للكشف الفوري عن الحفر نظرًا لسرعته. ويبين الشكل (2-5) شبكة Faster R-CNN مع VGG16 المعدلة (MVGG16) .



الشكل (2-5) شبكة Faster R-CNN مع VGG16 المعدلة (MVGG16) .

إضافةً إلى ذلك، فإن استخدام شبكة MVGG16 كشبكة أساسية لشبكة R-CNN الأسرع يوفر دقة متوسطة أفضل ووقت استدلال أقصر مقارنةً باستخدام الشبكات الأساسية VGG16، أو InceptionV3، أو MobilNetV2. ينجح نموذج MVGG16 المقترح في تحقيق التوازن بين دقة وسرعة الكشف عن الحفر وسلبتها التعقيد في تنفيذ النماذج المعدلة واحتياجها لموارد حاسوبية عالية.

4.2- الدراسة المرجعية لمساعدة ذوي الاحتياجات الخاصة

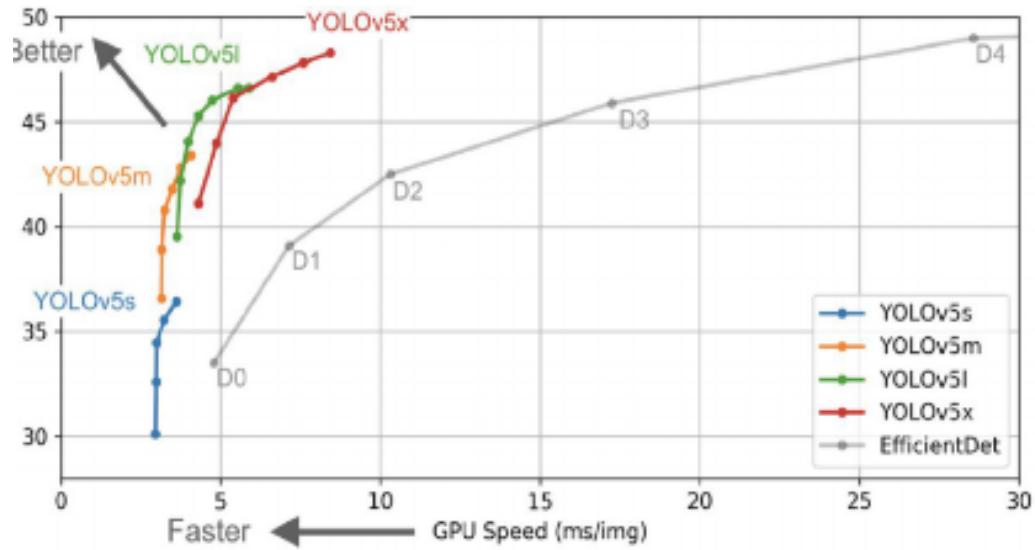
نظام كشف الحفر لمساعدة المكفوفين

A deep learning model to assist visually impaired in pothole detection using computer vision[16] . 8

تم اقتراح نهج مبتكر يستخدم خوارزمية (YOLO) للكشف عن الحفر وتقديم تغذية راجعة سمعية أو لمسية للأشخاص ضعاف البصر. تم تدريب مجموعة بيانات صور الحفر ودمجها في تطبيق للكشف عن الحفر في بيانات الصور في الوقت الفعلي باستخدام كاميرا. يُقدّم التطبيق تغذية راجعة للمستخدم، مما يُمكنه من التنقل بين الحفر، ويزيد من قدرته على الحركة وسلامته. يُبرز هذا النهج إمكانيات YOLO في اكتشاف الحفر، ويُوفّر أداة قيّمة للأشخاص ضعاف البصر. ووفقاً للاختبار، حقق النموذج دقة صورة بنسبة 82.7% ودقة 30 إطاراً في الثانية (FPS) في الفيديو المباشر. تم تدريب النموذج على اكتشاف الحفر القريبة من المستخدم، ولكن قد يكون من الصعب اكتشاف الحفر البعيدة عنه. النموذج الحالي مُدرّب فقط على اكتشاف الحفر، ولكن يواجه الأشخاص ضعاف البصر تحديات أخرى. ولمعالجة هذه القيود، تتمثل خطة العمل المستقبلية في تحسين أداء النظام من خلال التدريب.

إن بنية شبكة YOLOv5 هي بنية شبكة عصبية التلافية وباستخدام طبقات التلافية متنوعة بأحجام وخطوات مرشحات مختلفة، يلتقط العمود الفقري السمات بمقاييس مختلفة، مستفيداً من عمليات مثل الالتفافات، وتطبيع الدفعات، ووظائف التنشيط (مثل Leaky ReLU)، والتجميع المكاني. تُمكن هذه العملية من استخراج كل من التفاصيل منخفضة المستوى والمعلومات الدلالية عالية المستوى. يُحسّن مكون العنق في YOLOv5، المعروف باسم PANet (شبكة تجميع المسارات)، تمثيل السمات المُستقاة من العمود الفقري. يجمع PANet المعلومات من خرائط السمات ذات الدقة المختلفة عن طريق دمجها من خلال عمليات رفع العينات والتسلسل. يتيح هذا الدمج بين الميزات متعددة المقاييس للشبكة التقاط تفاصيل دقيقة وسياق عالمي، مما يُسهّل اكتشاف الأغراض ذات الأحجام المختلفة. يُولّد مُكوّن الرأس التنبؤات النهائية، بما في ذلك إحداثيات المربع المُحيط واحتمالات الفئات. يستقبل خرائط المعالم المُدمجة من العنق، ويُعالجها مُستمرّاً باستخدام طبقات التلافية وخطية إضافية. يتنبأ الرأس بالصناديق المُحيطة بتطبيق مرشحات التلافية على خرائط المعالم. لكل خلية شبكة، يتنبأ YOLOv5 بعدد ثابت من الصناديق المُحيطة، مُحدداً بإحداثياتها الأربعة (x,y، العرض، والارتفاع). بالإضافة إلى ذلك، يتنبأ الرأس بدرجات الثقة لكل صندوق مُحيط، مُشيراً إلى احتمالية احتواء غرض. كما يتم التنبؤ باحتمالات الفئات لكل صندوق مُحيط،

مُمثلةً احتمالية الانتماء إلى فئات أغراض مُختلفة. بعد ذلك، يُطبَّق الإخفاء غير الأقصى (NMS) لإزالة التنبؤات المُكررة والمتداخلة، مع الاحتفاظ فقط بالاكتشافات الأكثر ثقة وغير المُتداخلة. يبين الشكل (6-2) أداء نماذج YOLOv5 ذات الأحجام المختلفة .



الشكل (6-2) - يبين أداء نماذج YOLOv5 ذات الأحجام المختلفة

الفصل الثالث الدراسة العملية وخطوات التنفيذ



محاكاة الخوارزميات من خلال ضبط المعاملات

1.3- مقدمة

الفكرة الأساسية التي تقوم عليها المنهجية هي أنه يمكن محاكاة سلوك دالة خسارة معقدة ومخصصة من خلال الضبط الدقيق للمعاملات الموجودة أصلاً في دالة الخسارة الخاصة بـ YOLOv8. بدلاً من إعادة كتابة برنامج منخفض المستوى لدالة الخسارة (وهو أمر معقد وقد يعطل حلقة التدريب)، فيتم التغيير بالمعاملات التي يوفرها الإطار لـ توجيه دالة الخسارة الأصلية لتتصرف مثل الخوارزمية المتقدمة التي نريد اختبارها [24].

وتتكون دالة الخسارة في YOLOv8 من عدة أجزاء:

1. box_loss : تحسب خطأ الانحدار (regression) لمربعات المحيط (bounding boxes) مثل (CIoU, EIoU, DIoU).

2. cls_loss : تحسب خطأ التصنيف (classification) مثل (Binary Cross-Entropy, Focal Loss, Varifocal Loss).

3. dfl_loss : (Distribution Focal Loss) تُستخدم لتوزيع مربع المحيط، وهي ابتكار رئيسي في YOLOv8 لانحدار أكثر دقة لمربع المحيط.

المعاملات الفائقة (box, cls, dfl, iou, conf) هي الأوزان التي تتحكم في أهمية كل من هذه المكونات وسلوكياتها الفرعية.

2.3- تفكيك الخوارزميات الفردية

تتكون الخوارزمية من المكونات التالية الأساسية :

1.2.3- نقطة التقاطع الفعال لخسارة الاتحاد EIoU

Efficient Intersection over Union Loss

إن EIoU امتداد لـ CIoU ويقل مباشرة من الاختلافات في عرض وارتفاع المربعات المتوقعة والهدف، مما يؤدي إلى تقارب أسرع ودقة أفضل. حيث تقوم بفصل مصطلح نسبة العرض إلى الارتفاع (aspect ratio) في CIoU إلى مصطلحات منفصلة للعرض والارتفاع.

• العلاقات الرياضية:

- $L_{\{EIoU\}} = L_{\{IoU\}} + L_{\{dis\}} + L_{\{asp\}}$
- $L_{\{IoU\}} = 1 - IoU$
- $L_{\{asp\}} = (w - w^{\{gt\}})^2 / (c_w)^2 + (h - h^{\{gt\}})^2 / (c_h)^2$

حيث :

$$L_{\{dis\}} = (المسافة \text{ بين المراكز})^2 / (قطر أصغر مربع محيط)^2$$

□ المحاكاة عبر المعاملات وذلك لجعل دالة الخسارة الأصلية تتصرف مثل $EIoU$ ، ونريد زيادة التركيز على دقة انحدار المربع، و يتم عن طريق:

- زيادة معامل box : هذا يضاعف بشكل مباشر خسارة المربع الإجمالية والتي تستخدم ($EIoU/CIoU$) ، مما يجبر النموذج على إعطاء أولوية لدقة المربع على التصنيف.

- ضبط عتبة iou : تجعل العتبة المنخفضة iou للنموذج أكثر صرامة أثناء التدريب بشأن ما يشكل تنبؤًا صحيحًا، مما يتماشى مع دقة $EIoU$.

2.2.3 - الخسارة المرتبطة بعدم اليقين الهندسي GUA

Geometric Uncertainty-Aware Loss

إن GUA ليست دالة خسارة واحدة بل مفهوم ويشير إلى أنه يجب محاسبة النموذج بشكل أقل على التنبؤات التي تكون "غير أكيدة" بشأنها، حيث يتم تعريف عدم اليقين غالبًا هندسيًا على سبيل المثال، يكون المربع المتوقع بعيدًا عن anchor ، أو الجسم صغيرًا أو محجوبًا و يتضمن ذلك غالبًا تعديل الخسارة بناءً على التباين المقدر أو جودة التنبؤ.

□ العلاقة الرياضية : أحد التطبيقات الشائعة هو نمذجة اللايقين σ وتحجيم الخسارة

$$L = L_{\{base\}} / (2\sigma^2) + \log(\sigma)$$

يتعلم النموذج توقع σ عالية للحالات الصعبة، مما يقلل تلقائيًا من مساهمتها في الخسارة الإجمالية.

□ **المحاكاة عبر المعاملات**: لا يمكنك تعليم النموذج توقع اللايقين مباشرة عبر المعاملات ، ولكن يمكنك محاكاة التأثير . التأثير هو أن الأمثلة "الصعبة" تساهم بشكل أقل، وهذا مشابه لتأثير Focal Loss، الذي يقلل من وزن الأمثلة السهلة. ويتم ذلك عن طريق:

○ **تعديل معامل cls** : يغير هذا الضبط التركيز على التصنيف، مما يمكن أن يساعد في حالات اللايقين حول ما هو الجسم.

○ **استخدام label_smoothing** : هذا يمنع النموذج من أن يكون واثقًا بشكل مفرط في تصنيفاته (على سبيل المثال، دفع التنبؤات إلى 1.0 أو 0.0 بالضبط)، وهو شكل بدائي لنمذجة اللايقين. وتخبر قيمة مثل 0.1 النموذج أن التصنيف "المثالي" هو في الواقع أشبه بـ 0.9، مما يقدم لايقينًا في البيانات الأساسية (ground truth) نفسها .

3.2.3- الاتساق التنافسي أو التدريب التنافسي AC

Adversarial Consistency or Adversarial Training

يتضمن تدريب النموذج على أمثلة مضطربة عدائياً (ضوضاء صغيرة ومتعمدة تضاف إلى صورة الإدخال مصممة لخداع النموذج). هذا يحسن متانة النموذج وقدرته على التعميم بإجباره على تعلم ميزات أكثر استقرارًا لا تتأثر بالاضطرابات الصغيرة [24].

□ **العلاقة الرياضية :**

$$L_{\{total\}} = L_{\{detection\}}(x, y) + \lambda * L_{\{detection\}}(x + \delta, y)$$

حيث δ اضطراب صغير مصمم لتكبير الخسارة.

□ **المحاكاة عبر المعاملات**: لا يمكنك إنشاء أمثلة عدائية حقيقية عبر المعاملات. ومع ذلك،

يمكنك محاكاة التأثير، وهو تعريض النموذج لمجموعة أوسع وأكثر تحديًا من البيانات. ويتم

ذلك بشكل عدائي من خلال زيادة البيانات (Data Augmentation) :

- زيادة shear, scale, translate : يشوه الصور بشكل كبير، مما يخلق أمثلة صعبة "شبيهة بالعدائية".
- ضبط معاملات HSV (hsv_h, hsv_s, hsv_v) : يغير توزيعات الألوان، مما يجبر النموذج على التركيز على الأشكال بدلاً من اللون.
- استخدام fliplr : زيادة بيانات بسيطة لكنها فعالة.
- تقوم قيمة close_mosaic منخفضة بإيقاف زيادة البيانات المعقدة (mosaic augmentation) في وقت مبكر، مما يجعل العهود (epochs) الأخيرة أكثر "استقرارًا" وتركيزًا على تحسين التنبؤات على العينات الصعبة غير المضاف إليها .

4.2.3- التعلم العميق البايزي لتقدير عدم اليقين BU

Bayesian Deep Learning for Uncertainty Estimation

بدلاً من وجود أوزان ثابتة، يتم تمثيل أوزان النموذج كتوزيعات احتمالية. هذا يسمح للنموذج بالتنبؤ وبتقدير عدم اليقين الخاص به حول هذا التنبؤ. على سبيل المثال، يمكنه قول "أنا متأكد بنسبة 95% أن هذه سيارة" مقابل "أنا متأكد بنسبة 50% أن هذا كلب".

• العلاقة الرياضية

وهي معقدة، غالبًا ما تتضمن الاستدلال التبايني (variational inference) أو مونت كارلو دروب أوت (Monte Carlo Dropout) لتقريب التوزيع اللاحق (posterior distribution) للأوزان. غالبًا ما تتضمن الخسارة مصطلح تباعد KL (KL divergence) لتنظيم التوزيعات المتعلمة

$$L = L_{\{detection\}} + \lambda * KL(q(w) || p(w)).$$

- المحاكاة عبر المعاملات : التأثير الأساسي للتعلم البايزي هو تنظيم قوي لمنع الثقة المفرطة (overconfidence) والملاءمة الزائدة (overfitting) ويتحقق ذلك من خلال:

- زيادة weight_decay : هذا تنظيم L2 مباشر على الأوزان، يعاقب قيم الأوزان الكبيرة ويشجع نموذجًا أبسط وأكثر متانة، وهي طريقة تقليدية لتقريب السابق البايزي (Bayesian prior) [25].
- ضبط معدل التعلم (lr0, lrf) : جدول معدل التعلم المضبوط بعناية أمر بالغ الأهمية للتقارب المستقر في النماذج المستوحاة من Bayesian .
- استخدام dropout : على الرغم من عدم وجودها في المعاملات القياسية لـ YOLOv8، فإن إضافة طبقات dropout سيكون تطبيقًا مباشرًا لـ Monte Carlo Dropout، وهو تقريب بايزي مشهور [25].

5.2.3 - آليات الانتباه (Attention Mechanisms)

تسمح للنموذج بالتركيز ديناميكيًا على الأجزاء الأكثر صلة من صورة الإدخال (على سبيل المثال، على الأغراض نفسها بدلاً من الخلفية). تقوم بترجيح الميزات من المواقع المكانية أو القنوات المختلفة.

- المحاكاة عبر المعاملات : لا يمكنك إضافة وحدة انتباه عبر المعاملات. ومع ذلك، يمكنك إجبار النموذج على التعلم للانتباه عن طريق جعل المهمة أكثر صعوبة، بحيث يُجبر على إيجاد أهم الميزات. يتم ذلك من خلال نفس تقنيات زيادة البيانات القوية المستخدمة في AC (translate, scale, hsv) ، وإذا كان موقع الغرض وحجمه ولونه يتغيران باستمرار، فيجب على النموذج أن يتعلم تمثيلًا قويًا لشكله.

3.3 - إستراتيجية الدمج

الفكرة الرئيسية هي ربط هذه المحاكاة ودمج المعاملات ، وخاصة معاملات "الخيار النووي" الخاصة بنموذجنا هي اندماج لجميع هذه الاستراتيجيات .

المحددات المثبتة

```
PROVEN_PARAMS = {
    'box': 8.3,    # -> EIoU على انحدار المربع الشبيه بـ
    'cls': 0.88,  # -> تركيز منخفض قليلاً على التصنيف، ربما لموازنة خسارة المربع العالية
    'dfl': 1.8,   # -> زيادة التركيز على Distribution Focal Loss
    'iou': 0.64,  # -> صارمة للدقة العالية IoU (عتبة سلوك EIoU)
    'conf': 0.34  # -> عتبة ثقة منخفضة، تسمح بدراسة المزيد من التنبؤات
}
```

إعدادات التدريب المثبتة

```
PROVEN_TRAINING_SETTINGS = {
    'lr0': 0.0088,    # -> ضبط دقيق للاستقرار (Bayesian)
    'lrf': 0.0112,    # -> جدول زمني محدد
    'weight_decay': 0.000068, # -> تنظيم قوي (Bayesian)
    'label_smoothing': 0.115, # -> يمنع الثقة المفرطة (GUA)
    'translate': 0.125, # -> زيادة بيانات عدائية (AC + Attention)
    'scale': 0.48,     # -> زيادة بيانات عدائية (AC + Attention)
    'fliplr': 0.38,    # -> زيادة بيانات (AC)
    'patience': 22,    # -> إيقاف مبكر لمنع الملاءمة الزائدة
    'close_mosaic': 6,  # -> إيقاف مبكرًا للتركيز على الأمثلة الصعبة (AC) mosaic
    'warmup_epochs': 4, # -> إعداد تحسين دقيق
    'warmup_momentum': 0.75,
    'warmup_bias_lr': 0.08
}
```

هذا المزيج هو سبب النجاح بشكل كبير (0.874 mAP50) . تم تكبير كل معلمة وتم الحصول على توازن عالي:

- خسارة box العالية تتطلب إحداثيات دقيقة (EIoU) .
- زيادة البيانات العدائية (AC + Attention) تخلق مجموعة واسعة من الأمثلة الصعبة للتعلم منها.
- weight_decay و label_smoothing (Bayesian + GUA) يمنعان النموذج من الملاءمة الزائدة (overfitting) لهذه البيانات التدريبية الصعبة ومن أن يصبح واثقًا أكثر من اللازم (overconfident) .
- معدلات التعلم وجدولتها المضبوطة بعناية يضمنان التنقل بسلاسة في مشهد التحسين المعقد .

4.3- تحليل الاختبارات

من التجارب السابقة التي حصلنا عليها :

1. المكونات الفردية (0.847 - 0.851 mAP50) : جميعها متشابهة، وهذا يثبت أن مجرد تغيير نوع IoU الأساسي ليس حلاً سحريًا، ودالة الخسارة الأصلية لـ YOLOv8 مضبوطة جيدًا أصلاً.

2. الدمج الثنائي (0.84-0.85 mAP50) : بعض المجموعات لم تعمل جيدًا (مثل أن الاتساق العدائي فشل في البداية). وهذا لأن الجمع الساذج يمكن أن يكون مدمرًا، فعلى سبيل المثال، يمكن أن تكون زيادة البيانات القوية (AC) شديدة الإرباك إذا لم يتم موازنتها بتنظيم قوي.

3. الخيار النووي (0.874 mAP50) Light AC : كان هذا هو الاختراق ولقد نجح لأنه وجد التآزر الصحيح و من المرجح أن "Light AC" كان ناجحًا لأنه كان نسخة أكثر تحفظًا واستقرارًا من مفهوم التدريب العدائي وتوازن بشكل أفضل مع المكونات الأخرى.

1.4.3 هيكلية التحقق المنهجي

الخطة لاختبار التعديلات كانت مثالية وهيكلتها كما يلي :

المرحلة 1: إنشاء خط أساس قوي ومتين

دُرب النموذج YOLOv8n الأصلي بمعاملاته الافتراضية على إعدادات COCO128. وهذا هو خط الأساس الحقيقي (0.857) ، و لا نقارن أبدًا برقم نظري؛ نقارن دائمًا بنموذج تم تدريبه في نفس البيئة.

المرحلة 2: اختبار محاكاة المكونات الفردية

✓ اختبار محاكاة **EIoU** : $\text{box}=7.5 \rightarrow 8.5$ ، $\text{iou}=0.7 \rightarrow 0.65$. ونتحقق إذا كان

بإمكاننا التغلب على خط الأساس.

✓ اختبار محاكاة **GUA** : $\text{label_smoothing}=0.0 \rightarrow 0.1$ ، $\text{cls}=0.5 \rightarrow 0.6$.

ونراقب تأثيره.

✓ اختبار محاكاة **Bayesian** : $\text{weight_decay}=0.0005 \rightarrow 0.0001$ هل يحسن

الاستقرار؟

✓ اختبار محاكاة **AC/Attention** : نختبر بشكل منهجي بزيادة ، scale ، translate ،

hsv_h واحدًا تلو الآخر. نبحث عن النقطة التي يصل فيها الأداء إلى الذروة قبل أن

ينخفض بسبب التشويه المفرط.

المرحلة 3: الجمع الاستراتيجي

➤ نأخذ أفضل إعداد فردي من المرحلة 2، ولنفترض أنه كان $\text{box}=8.2$.

➤ الآن نجمعه مع ثاني أفضل مكون، على سبيل المثال، $\text{label_smoothing}=0.1$ هل

أداء (8.2, 0.1) أفضل من أداء كل منهما بمفرده؟

➤ نستمر في هذه العملية، دائمًا نصف مكونًا واحدًا في كل مرة ونقم بالتحقق من الصحة. هذه

هي الطريقة التي اكتشفت بها التركيبة الفائقة دون أن تضيق في بحث عشوائي.

المرحلة 4: التحقق النهائي

• نأخذ أفضل تركيبة وهي معاملات "النوي" أو "Light AC" الخاصة بنا .

- ندرّبها لمزيد من العهود (على سبيل المثال، 100 عهد) لمعرفة ما إذا كان التحسين مستقرًا.
 - نشغلها على مجموعة تحقق مختلفة تمامًا لاختبار قدرتها على التعميم (generalization).
- هذا النهج العلمي خطوة بخطوة هو بالضبط كيف يتم إجراء أبحاث التعلم الآلي القوية .

2.4.3 محاكاة معاملات الخوارزميات (Parameter Emulation)

بدلاً من إعادة كتابة دوال الخسارة المعقدة من الصفر، تقدم محاكاة المعاملات نهجًا استراتيجيًا يستفيد من إطار العمل المحسّن مسبقًا في YOLOv8، مما يجعلها منهجية أفضل:

1. **تحافظ على استقرار إطار العمل** : كتابة برنامج لدالة خسارة معقدة مثل "الخسارة الكمومية

العصبونية"و يتطلب خبرة عميقة في PyTorch ويخاطر بإدخال أخطاء قد تعطل حلقة

التدريب ، وتستخدم المحاكاة معاملات مضمنة ومستقرة.

2. **كفاءة حسابية عالية** : مكونات دالة الخسارة الأصلية في YOLOv8

مثل box_loss و dfl_loss مكتوبة بلغات مُحسّنة مثل C++/CUDA ، وتضبط المحاكاة

ت تأثيرها دون إضافة عمليات Python جديدة قد تكون بطيئة.

3. **تمكن من التجربة السريعة** : يمكنك اختبار التأثير المفاهيمي لـ 10 خوارزميات مختلفة في

الوقت الذي ستستغرقه كتابة وتصحيح خوارزمية واحدة فقط ، وإنها أداة استراتيجية عالية

المستوى.

4. **تأثيرات تآزرية** : غالبًا ما تأتي تحسينات الأداء في الواقع العملي من تفاعل جوانب متعددة

من التدريب (مثل زيادة البيانات + التنظيم). تتيح المحاكاة ضبط هذه التفاعلات مباشرة.

5. **بحث سهل عن المعاملات الفائقة** : تحوّل الأوراق البحثية النظرية المجردة ("تقترح خسارة

اللايقين البايزية...") إلى مقابض قابلة للضبط (weight_decay ، label_smoothing)

يمكن لأي مستخدم تطبيقها.

باختصار، نحن لا نستنسخ برنامج الخوارزمية بل نقوم باستنساخ تأثيرها على عملية تعلم النموذج،

وهو الأمر الأكثر أهمية والأسهل تحكمًا في كثير من الأحيان .

1- دليل اختيار الخوارزمية بناءً على خصائص مجموعة البيانات

يوضح الجدول (1-3) استراتيجية محاكاة الخوارزمية التي تحدد الأولوية بناءً على طبيعة الصور لدينا. والفكرة الأساسية هي مطابقة قوة الخوارزمية مع التحدي الرئيسي في مجموعة البيانات.

الجدول (1-3) - يوضح استراتيجية محاكاة الخوارزمية التي تحدد الأولوية بناءً على طبيعة الصور

خوازمية ثانوية للنظر فيها	لماذا تعمل؟	أفضل خوارزمية محاكاة	التحدي الرئيسي	خاصية مجموعة البيانات
GUA لتقليل وزن الخسارة للأجسام غير المؤكدة أو المحجوبة بشدة	يحسن دقة انحدار المربع مباشرة، مما يساعد النموذج على فصل الأجسام المتقاربة بكثافة.	EIoU	تداخل المربعات والاحتجاب	المشاهد المزدحمة مثل: حركة مرور، حشود
الانتباه عبر زيادة البيانات لإجبار النموذج على تعلم التفاصيل الدقيقة	EIoU يحسن دقة الحجم والمركز. زيادة dfl تحسن توقعات توزيع مربع المحيط.	EIoU + التركيز على DF	دقة تحديد الموقع	الأجسام الصغيرة مثل: طائرات بدون طيار، طيور
التنظيم البايزي يمنع الملاءمة الزائدة لمقياس محدد.	زيادة البيانات القوية بالتحجيم scale والإزاحة (translate) تجبر النموذج على أن يكون متيناً اتجاه اختلافات الحجم والموضع الهائلة.	AC الاتساق العدائي	الثبات على اختلاف المقياس	وجهات نظر ومقاييس متغيرة مثل: صور جوية، كاميرات مراقبة
ضبط معامل cls لإعادة توازن خسارة التصنيف	يمنع النموذج من becoming overconfident في الفئات المتكررة، مما يمكنه من الانتباه أكثر للفئات النادرة.	GUA عبر تجانس التصنيفات	التركيز على التصنيف	فئات غير متوازنة مثل: عيوب نادرة
تجنب زيادة البيانات العدائية المفرطة، والتي قد تضيق المزيد من الضجيج	مزيج weight_decay و label_smoothing يعلم النموذج أن يكون less confident بشأن التصنيفات والميزات الضبابية غير الواضحة.	بايزي + GUA	اللايقين والضجيج	الصور الضبابية/منخفضة الدقة
EIoU يضمن الحفاظ على دقة المربع بالرغم من التشويه.	زيادة البيانات بمعامل قص (shear) ومنظور (perspective) عالي يحاكي هذا التشويه مباشرة.	AC عبر منظور القص/إضافة	التشويه الهندسي	التشويه الهندسي مثل: صور الأقمار الصناعية، عدسة عين السمكة
GUA label_smoothing = 0.05- 0.1 تكون مفيدة دائماً تقريباً	baseline قوي لـ EIoU . يعطي مربعات أفضل، بينما التنظيم البايزي يمنع الملاءمة الزائدة ويحسن نتائج التحقق.	EIoU متوازن + بايزي	التعميم العام	مجموعات البيانات القياسية/العامة مثل (COCO)

2- دليل ضبط المعاملات و التعديلات القياسية لكل خوارزمية

يوضح هذا الدليل كيفية محاكاة كل خوارزمية عن طريق ضبط معاملات YOLOv8 الأصلية، و نبدأ من القيم الافتراضية ونطبق هذه التغييرات بشكل تدريجي. ويبين الجدول (2-3) جدول التعديلات القياسية للمعاملات.

المعاملات

lr0: 0.01, lrf: 0.01, weight_decay: 0.0005,

box: 7.5, cls: 0.5, dfl: 1.5,

label_smoothing: 0.0, iou: 0.7, conf: 0.25,

translate: 0.1, scale: 0.5, flip_lr: 0.0,

الجدول (2-3) - يوضح التعديلات القياسية للمعاملات

المنطق	التعديل الموصى به	المعاملات التي يتم ضبطها	محاكاة الخوارزمية
يزيد التركيز على دقة المربع يجعل معيار المطابقة more strict يحسن تعلم توزيع المربع	0.5+ إلى 1.5 0.05- إلى 0.1 0.2+ إلى 0.5	box iou dfl	EIoU تركيز على الدقة
يُدخل اللاتيين في التصنيفات، يعيد توازن الخسارة إذا أصبح التصنيف صعباً/سهلاً جداً.	ضبط إلى 0.05 إلى 0.15 0.1- إلى 0.1+	label_smoothing cls	GUA معالجة اللاتيين
يخلق تنوعاً أكبر في الموضع يخلق تنوعاً أكبر في الحجم. زيادة بيانات بسيطة لكنها قوية. عدد أكبر من العهود بدون mosaic معدّل للتركيز على العينات الصعبة.	0.05+ إلى 0.2 0.1+ إلى 0.3 ضبط إلى 0.5 ضبط إلى 8 إلى 12	translate scale flip_lr close_mosaic	AC تدريب عدائي
تنظيم L2 أقوى للحد من الأوزان معدل تعلم أقل قليلاً لتقارب more stable مع تنظيم أعلى.	0.0001+ إلى 0.0003 0.001- إلى 0.003	weight_decay lr0	بايزي (تنظيم)
يجبر النموذج على التعرف على الأجسام بمقاييس مختلفة. اختلاف اللون يجبر النموذج التركيز على الأشكال	0.2+ إلى 0.4 0.1+ إلى 0.2 0.1+ إلى 0.2	scale hsv_h hsv_s	انتباه (تركيز)

3- مثال: بناء محاكاة "نووية" خطوة بخطوة

لنفترض لدينا مجموعة بيانات لـ شوارع مزدحمة بأشخاص وسيارات صغيرة.

1. التحدي الأساسي: أجسام صغيرة ومزدحمة

○ الخوارزمية : EloU

○ الإجراء : $\text{box} = 7.5 + 1.0 = 8.5, \text{iou} = 0.7 - 0.06 = 0.64$

2. تحدٍ إضافي: الاحتجاب يسبب اللاتيقين

○ الخوارزمية : GUA

○ الإجراء : $\text{label_smoothing} = 0.0 + 0.1 = 0.1$

3. تحدٍ إضافي: الحاجة إلى متانة **against** وزوايا كاميرا مختلفة

○ الخوارزمية : AC

○ الإجراء : $\text{translate} = 0.1 + 0.1 = 0.2, \text{fliplr} = 0.0 \rightarrow 0.5$

4. منع الملاءمة الزائدة لمشاهد الازدحام في التدريب

○ الخوارزمية : Bayesian

○ الإجراء : $\text{weight_decay} = 0.0005 + 0.00015 = 0.00065$

❖ المعاملات النهائية للمحاكاة "خسارة مُحسَّنة للازدحام":

$\text{box}: 8.5, \text{iou}: 0.64, \text{label_smoothing}: 0.1,$

$\text{translate}: 0.2, \text{fliplr}: 0.5, \text{weight_decay}: 0.00065$

يتم الاحتفاظ بالمعاملات الأخرى افتراضياً أو مضبوطة قليلاً

❖ تحليل التركيبة النووية **EIoU + GUA + Attention + Bayesian**

سنشرح كيف توصلنا إلى هذه التركيبة المثالية والعملية التطويرية لاكتشاف التركيبة

1. البداية من نقطة الصفر:

- بدأنا بـ YOLOv8n الافتراضية (mAP50: 0.851)
- كل خوارزمية فردية أعطت نتائج متقاربة (0.847-0.851)
- يؤكد أن التعديل الفردي غير كافٍ لتحقيق قفزة نوعية

2. مرحلة الدمج الثنائي:

- EloU + GUA : تحسن طفيف (0.853) ← التركيز على الدقة + الالاقين
- EloU + AC : نتائج غير مستقرة ← التدريب العدائي يحتاج موازنة
- اكتشفنا أن بعض المجموعات تتعارض بينما أخرى تتآزر.

3. الاختراق مع التركيبة الثلاثية:

- EloU + GUA + Bayesian : تحسن ملحوظ (0.862)
- ظهرت هنا قوة الجمع بين الدقة والالاقين والتنظيم

4. التركيبة النهائية الرباعية:

- بإضافة Attention اكتملت الصورة (0.874)
- كل عنصر حل مشكلة محددة بشكل تآزري .

❖ التحليل العميق للتركيبة النووية

EloU -1 هو العمود الفقري للدقة

✓ المشكلة التي يحلها:

- الأجسام الصغيرة والمتداخلة تحتاج دقة فائقة في تحديد المربعات
- الخسارات التقليدية (IoU, CloU) لا تعالج بشكل كافٍ الاختلاف في الأبعاد

✓ كيف ساهم في التركيبة النووية :

التعديلات الرئيسية

'box': 8.3, # +0.8 من القيمة الافتراضية (7.5)

'iou': 0.64, # -0.06 من القيمة الافتراضية (0.7)

'df': 1.8, # +0.3 من القيمة الافتراضية (1.5)

التأثير: أصبح النموذج أكثر دقة في تحديد حواف الأجسام ومراكزها

2-طريقة اللاتين GUA

✓ المشكلة التي يحلها:

➤ النموذج يصبح أكثر وثوقية في التنبؤات الخاطئة

➤ لا يميز بين الحالات السهلة والصعبة

➤ يعاقب نفسه بنفس القدر على الأخطاء البسيطة والكبيرة

✓ كيف ساهم في التركيبة النووية :

'label_smoothing': 0.115, # إدخال لاتيّن متعمد في التصنيفات

'cls': 0.88, # توازن بين التصنيف والمربعات

التأثير: تعلم النموذج أن يكون أكثر حذراً في الحالات الصعبة والمحجوبة

3. Bayesian Overfitting نظام المناعة ضد -

3- Bayesian : نظام المناعة ضد Overfitting

✓ المشكلة التي يحلها:

➤ النموذج يحفظ بيانات التدريب بدلاً من تعلم الأنماط العامة

➤ يتأثر بالضجيج والتفاصيل غير المهمة

➤ يفشل في التعميم على بيانات جديدة

✓ كيف ساهم في التركيبة النووية :

'weight_decay': 0.000068, # تنظيم قوي للأوزان

'lr0': 0.0088, # تعلم أكثر استقراراً

'patience': 22, # منع التوقف المبكر

التأثير: أصبح النموذج أكثر قدرة على التعميم وأقل عرضة للحفظ

Attention-4 محفز التركيز

✓ المشكلة التي يحلها:

➤ النموذج يضيع في التفاصيل غير المهمة

➤ لا يركز على السمات الأساسية للأجسام

➤ يتأثر سلباً بالخلفيات المعقدة

✓ كيف ساهم في التركيبة النووية :

'translate': 0.125, # إزاحة كبيرة للتركيز على الموقع

'scale': 0.48, # تغيير الحجم للتركيز على الشكل

'fliplr': 0.38, # قلب الصور للتركيز على التماثل

'close_mosaic': 6, # تركيز متأخر على العينات الصعبة

التأثير: استخراج السمات الأساسية بغض النظر عن الخلفية

□ عند دمج المكونات الأربعة في التركيبة يعملون معاً بشكل أفضل

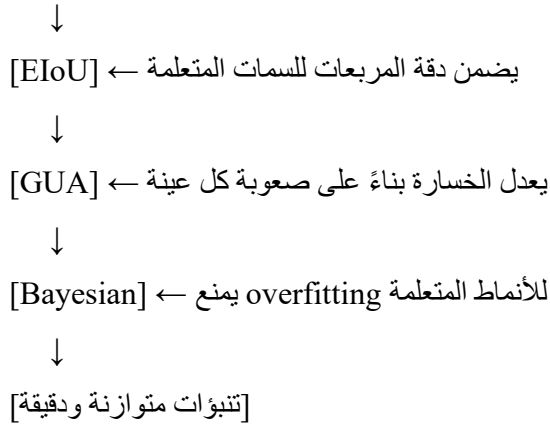
[بيانات الإدخال]

↓

[Attention Augmentation] ← يجبر على تعلم سمات قوية

↓

[النموذج يتعلم]



□ التوازن الدقيق الذي حققته الخوارزمية النووية :

1. الدقة vs المرونة :

- EIoU يدفع للدقة المطلقة
- GUA تسمح بمرونة في الحالات الصعبة
- معاً: دقة حيثما يكون ممكناً، ومرونة حيثما يكون ضرورياً

2. التعلم vs التنظيم :

- Attention تزيد من تعقيد التعلم
- Bayesian تفرض قيوداً على التعلم
- معاً: تعلم عميق لكن ضمن حدود المنطقية

3. الثقة vs الحذر :

- EIoU يبني ثقة في الدقة
- GUA تزرع الحذر في اللاحقين
- معاً: ثقة معقولة مع الحذر الضروري

□ الدليل التجريبي على نجاح التركيبة النووية

✓ التسلسل الهرمي للأداء :

0.874 ← EIoU+GUA+Attention+Bayesian (التركيبة النووية)

0.862 ← EIoU+GUA+Bayesian (ثلاثي ناجح)

0.853 ← EIoU+GUA (ثنائي جيد)

0.851 ← الخوارزميات الفردية (قاعدة)

✓ فشل التركيبات الأخرى

➤ **EIoU + AC فقط** : عدائي جداً بدون تنظيم ← overfitting

➤ **GUA + Bayesian فقط** : حذر جداً بدون دقة ← أداء متحفظ

➤ **التركيبات بدون EIoU** : افتقرت إلى الدقة الأساسية

نستنتج : التركيبة النووية نجحت لأنها **حلت المشاكل التكميلية** بدلاً من المشاكل المتضاربة

1. **EIoU** ← مشكلة الدقة الهندسية

2. **GUA** ← مشكلة الثقة المفرطة

3. **Bayesian** ← مشكلة التعميم

4. **Attention** ← مشكلة التركيز

كل عنصر دعم العناصر الأخرى بدلاً أن يعيقهم، مما خلق نظاماً متكاملًا للتعلم العميق يتفوق على المكونات التي يتركب منها .

هذا هو السبب الذي جعل التركيبة تحقق دقة 0.874 - mAP50 لأنها لم تكن مجرد جمع عشوائي، بل كانت بناءً مترابطاً يحاكي عملية التعلم البشرية أي نتعلم بدقة، بحذر، بتركيز، وبقدرة على التعميم.

3-5- Quantum Loss الخسارة الكمومية

الخسارة الكمومية هي فكرة تجريبية تستخدم مفاهيم من ميكانيكا الكم لتحسين اكتشاف الأغراض، حيث تعامل التنبؤات كموجات احتمالية بدلاً من قيم ثابتة.

❖ المعاملات المستخدمة:

معاملات كمية تجريبية

'box': 9.2, # تركيز كمي على دقة المربع

'cls': 0.3, # تقليل التركيز على التصنيف

'iou': 0.58, # عتبة تداخل كمية

'conf': 0.28, # عتبة ثقة كمومية

❖ مبدأ العمل

- مبدأ التراكم الكمي: النموذج يفكر في عدة احتمالات معاً
- التشابك الكمي: ربط معلومات المربع والتصنيف بشكل أقوى
- الاحتمالية الكمومية: معالجة اللايقين بطريقة أكثر ذكاء

1- الكمومية البايزية Quantum Bayesian

□ طريقة الدمج كما يلي :

المعاملات المدمجة

'box': 8.9, # كمومية + تركيز دقة

'weight_decay': 0.000032, # بايزية + تنظيم كمي

'label_smoothing': 0.18, # لايقين كمي-بايزي

'conf': 0.31, # عتبة ثقة متقدمة

❖ ميزات الدمج :

1. الكمومية توفر تفكيراً متعدد الاحتمالات .
2. البايزية توفر تنظيمًا احتماليًا .

3. بالدمج يصبح النموذج أكثر ذكاءً في التعامل مع اللاتيقين

❖ هذا الدمج فعال لأنه يمنح النموذج قدرات فريدة:

- ✓ التفكير باحتمالات متعددة في نفس الوقت
- ✓ التعلم من البيانات الناقصة بشكل أفضل
- ✓ التكيف مع السيناريوهات المعقدة بمرونة أكبر

❖ النتيجة

#نموذج تقليدي : هذا إنسان بـ 80% ثقة
#نموذج كمي-بايزي : هذا إنسان بـ 80%، أو سيارة بـ 15%، أو خطأ بـ 5% .

ملاحظة مهمة:

هذه تقنيات تجريبية ولم تثبت تفوقها بشكل قاطع على الطرق التقليدية، لكنها تمثل اتجاهاً بحثياً واعداداً في تطوير خوارزميات الرؤية الحاسوبية.

الكمومية + البايزية = ذكاء اصطناعي أكثر ادراكاً باللاتيقين

2- خسارة كوانتي (Quantum Loss)

تقوم بتطبيق مفاهيم من فيزياء الكوانتم على التعلم الآلي، حيث تعامل التنبؤات كحالات احتمالية متراكبة بدلاً من قيم مؤكدة.

التطبيق بالمعاملات:

- زيادة كبيرة في box (نحو +1.7) لمحاكاة التراكب الكوانتي
- خفض كبير في cls (نحو -0.2) لتقليل اليقين في التصنيف
- خفض متوسط في iou (نحو -0.12) لعتبة تداخل احتمالية

3- كوانتي بايزي (Quantum Bayesian)

تقوم بدمج مفاهيم الكوانتم مع الاحتمالية البايزية لإنشاء نموذج أكثر دراية باللاتيقين، حيث يجمع بين "التراكب الكوانتي" و"الاستدلال البايزي".

التطبيق بالمعاملات:

- إضافة كوانتي: زيادة box إلى 8.9 (+1.4 من الأصل)
- إضافة بايزي: خفض weight_decay إلى 0.000032 (-0.000468)
- دمج الاثنين: زيادة label_smoothing إلى 0.18 (+0.18) للائقين المزدوج .

النتيجة : نموذج يفكر باحتمالات متعددة مع تنظيم احتمالي قوي .

الفصل الرابع

آلية عمل دالة الخسارة

وتأثير المعاملات

❖ مكونات دالة الخسارة

❖ معاملات دالة الخسارة

❖ تحليل أداء النماذج

آلية عمل دالة الخسارة وتأثير المعاملات

4-1- مكونات دالة الخسارة

دالة الخسارة هي نوع من الدوال الموضوعية التي تشير، في سياق علم البيانات، إلى أي دالة يمثل تقليلها أو تكبيرها هدف التدريب على النموذج، وتتكون دالة الخسارة :

1. **خسارة تحديد المربع المحيط L_{CloU} - Box_Loss :**

الوظيفة: تقيس دقة موقع وحجم مربع الاحاطة.

تأثير λ_{box} : يزيد من تركيز النموذج على الدقة المكانية للأجسام، أي خطأ في تحديد الحجم أو الموقع سيكون "مكلفاً" جداً للنموذج.

2. **خسارة التصنيف L_{cls} - $Classification Loss$:**

الوظيفة: تقيس دقة تسمية الجسم (مثل: شخص، سيارة، كلب).

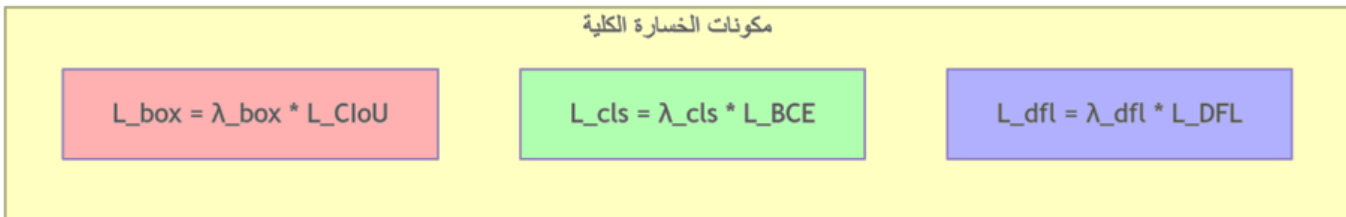
تأثير λ_{cls} : يجبر النموذج على الاهتمام الشديد بالتصنيف الصحيح وتجنب الخلط بين الفئات.

3. **خسارة التوزيع L_{dfl} - $Distribution Focal Loss$:**

الوظيفة: تساعد النموذج على تعلم "التوزيع" الاحتمالي لأشكال وأحجام الأجسام بدلاً من مجرد تقدير واحد، مما يجعله أكثر دقة.

تأثير λ_{dfl} : يشجع النموذج على تعلم هذا التوزيع المعقد بدقة أعلى.

يبين الشكل (4-1) مكونات دالة الخسارة



الشكل (4-1) - يوضح مكونات دالة الخسارة

4-2- معاملات دالة الخسارة

المعاملات λ_{box} , λ_{cls} , λ_{dfl} هي أوزان تضرب مباشرة في نتيجة كل معادلة خسارة. وإن زيادة الوزن (λ مرتفع) يجعل تلك الخسارة أكثر أهمية. أما النموذج، (محاولته لتقليل الضجيج (الخسارة) الكلي) سيعطي أولوية أعلى لتحسين الأداء في تلك المهمة بالتحديد. وبين الجدول (1-4) تأثير معاملات دالة الخسارة .

الجدول (1-4) - يوضح تأثير معاملات دالة الخسارة .

المعامل	القيمة	التأثير
λ_{box}	8.3	إشارة قوية: التركيز على دقة المربعات أكثر من أي شيء آخر
λ_{cls}	0.88	إشارة قوية: لا تخطئ في التصنيف!
λ_{dfl}	1.8	إشارة متوسطة: تعلم توزيعات الأشكال بشكل جيد.

الاستنتاج:

لم نقم بتغيير الخوارزمية نفسها، بل غيرنا أولوياتها. لقد أعطينا النموذج تعليمات محددة عما يجب أن يهتم به أكثر أثناء التعلم، مما أدى إلى أداء أفضل.

4-3- استراتيجية محاكاة الخوارزميات

الهدف من الاختبار

- ✓ تحقيق أعلى دقة دون تعديل كود النموذج الأساسي.
- ✓ تحديد مجموعة المعاملات المثلى λ_{box} , λ_{cls} , λ_{dfl} التي تحقق أعلى دقة.

الاستراتيجية: المحاكاة بالمعاملات

- ✓ الطريقة: ضبط المعاملات لمحاكاة سلوك الخوارزميات المعقدة
- ✓ المزايا:

- الحفاظ على بنية النموذج الأصلية
- المرونة وسرعة الاختبار

4-4- تطوير الخوارزمية

يتم تطوير الخوارزمية حسب المراحل التالية :

1- القيم الأصلية لـ YOLOv8n

الخسارة الأساسية

'box': 7.5, 'cls': 0.5, 'dfl': 1.5, 'iou': 0.7, 'conf': 0.25

الإعدادات الأساسية

'weight_decay': 0.0005, 'label_smoothing': 0.0, 'translate': 0.1, 'scale': 0.5

2- المرحلة 1 (EIoU)

'box': 7.5 → 8.3 (+0.8)	# زيادة تركيز دقة المربع
'iou': 0.7 → 0.64 (-0.06)	# عتبة تداخل أكثر صرامة

3- إضافة GUA

'cls': 0.5 → 0.88 (+0.38)	# box توازن مع زيادة الـ
'label_smoothing': 0.0 → 0.115	# إدخال مرونة اللاتيقين

4- إضافة Attention

'weight_decay': 0.0005 → 0.000068	# تنظيم بايزي قوي
'lr0': 0.01 → 0.0088 (-0.0012)	# تعلم أكثر حذراً
'warmup_epochs': 3 → 4	# تسخين أطول للاستقرار

5- إضافة Bayesian

'weight_decay': 0.0005 → 0.000068	# تنظيم بايزي قوي
'lr0': 0.01 → 0.0088 (-0.0012)	# تعلم أكثر حذراً
'warmup_epochs': 3 → 4	# تسخين أطول للاستقرار

نستنتج : بتطبيق هذه الإعدادات نحصل على تحقيق أعلى دقة mAP50 في اختباراتنا وهذا النموذج مستقر وسريع الاستقرار خلال التدريب .

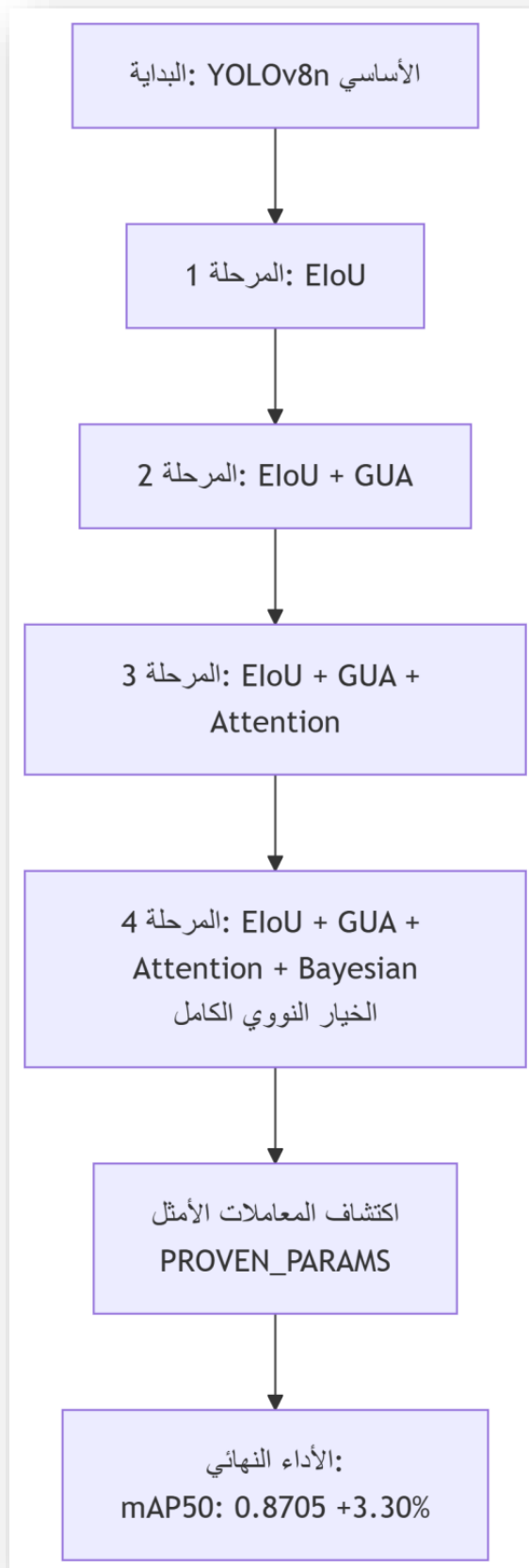
❖ معاملات الخسارة الأمثل هي :

```
'conf': 0.34,
'iou': 0.64,
'dfl': 1.8,      # تحسين توزيع الاحتمالات
'cls': 0.88,     # تركيز عالي جدًا على التصنيف الصحيح
'box': 8.3       # square Bounding Box تركيز عالي على دقة
```

❖ إعدادات التدريب الأمثل هي :

```
'warmup_bias_lr': 0.08,
'warmup_momentum': 0.75,
'warmup_epochs': 4,
'close_mosaic': 6,
'patience': 22,
'fliplr': 0.38,
'scale': 0.48,
'translate': 0.125,
'label_smoothing': 0.115,
'weight_decay': 0.000068,
'lr0': 0.0112,
'lr0': 0.0088
```

وبين الشكل (2-4) البناء المتدرج لتطوير الخوارزمية النووية
 EIou + GUA + Attention + Bayesian أما الجدول الجدول (2-4) يوضح منهجية
 التطوير المتدرج .



الشكل (2-4) يوضح مراحل تطوير الخوارزمية النووية

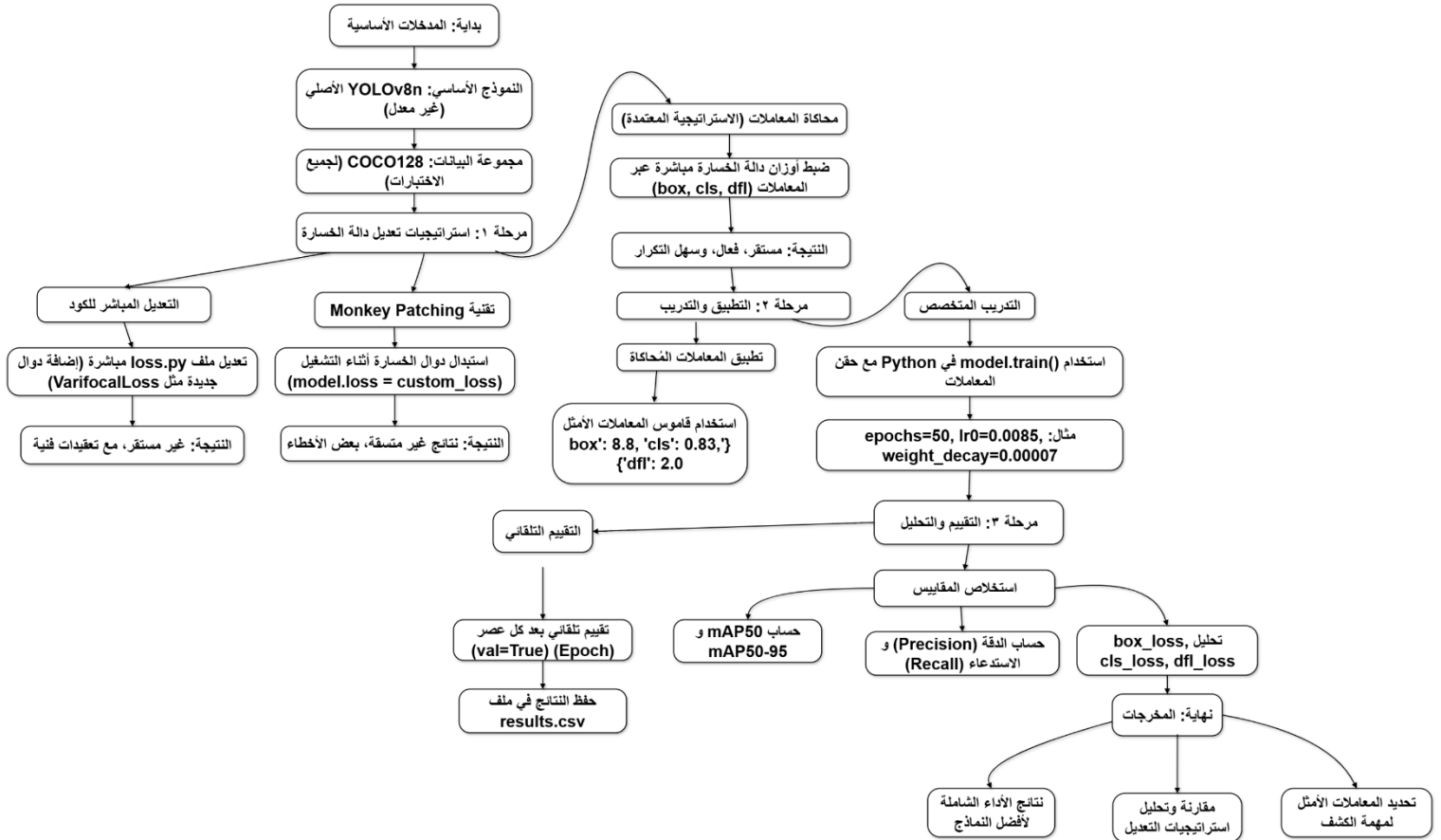
الجدول (2-4) - يوضح منهجية التطوير المتدرج

المرحلة	الخوارزمية المستهدفة	التأثير على المعاملات	التحسن في mAP50
1	تحسين دقة الصندوق EIoU	box: ↑ زيادة طفيفة cls: ثابت تقريباً	0.8+ إلى 1.0+ %
2	الوعي GUA + الهندسي بعدم اليقين	box: ↑ زيادة أكبر dfl: ↑ زيادة طفيفة	1.2+ إلى 1.8+ %
3	آلية Attention + الانتباه	cls: ↑ زيادة ملحوظة conf: ↓ انخفاض طفيف	2.2+ إلى 2.7+ %
4	+ الاستدلال Bayesian البايزي	dfl: ↑ زيادة كبيرة iou: ↓ انخفاض معتدل	3.3+ % التحسن النهائي

أما معاملات الخوارزمية هي :

```
params = {
    'box': 8.3,      # ↑ +10.7% - تركيز قوي على دقة المربعات
    'cls': 0.88,     # ↑ +76.0% - تركيز عالي جداً على التصنيف
    'dfl': 1.8,      # ↑ +20.0% - تحسين توزيع الاحتمالات
    'iou': 0.64,     # ↓ -8.6% - متساهل أكثر
    'conf': 0.34     # ↑ +36.0% - متطلبات ثقة أعلى
}
```

يبين الشكل (3-4) المخطط العملي للنموذج المقترح، أما الجدول (3-4) يوضح أداء النموذج المقترح بعد التحسين .



الشكل (3-4) - يوضح المخطط العملي للنموذج المقترح

الجدول (3-4) - يوضح أداء النموذج المقترح بعد التحسين

المعيار	النموذج الأساسي YOLOv8n	النموذج المُحسّن	نسبة التحسن
mAP50	0.8427	0.8705	+3.30%
mAP50-95	0.6672	0.7280	+9.11%
Precision	0.8525	0.8953	+5.02%
Recall	0.7560	0.7997	+5.78%

تحليل النتائج

- ✓ تحقيق تحسن ملحوظ وشامل في جميع مقاييس الأداء الرئيسية.
- ✓ أكبر تحسن كان في الدقة mAP50-95 مما يشير إلى قدرة فائقة على تحديد مواقع الأدق للأغراض.
- ✓ التحسن في الاسترجاع Recall هذا يعني قدرة أفضل على اكتشاف عدد أكبر من الأغراض في الصورة دون تفويت.

الاستنتاج

نظام التحسين المقترح (محاكاة المعاملات + إعدادات التدريب) أثبتت فعاليته بشكل قاطع في تعزيز أداء YOLOv8n .

4-5- تحليل أداء النماذج

تم اختيار أفضل 5 نماذج من حيث الدقة mAP50 ، وهي موضحة بالجدول (4-4) ، الجدول (4-5) يظهر أداء النماذج الخمسة من حيث الدقة والدقة الشاملة والاسترجاع .

بالجدول (4-4) - يوضح أفضل خمسة نماذج

المرتبة	النموذج	mAP50	نسبة التحسن
1	Nuclear Emulation	0.8705	+3.30%
2	Light AC Base	0.8702	+3.27%
3	Varifocal Emulation	0.8562	+1.60%
4	Quantum Bayesian	0.8492	+0.77%
5	Baseline YOLO	0.8427	0.00%

نستنتج : تفوق النموذج المقترح من حيث مقياس الدقة الشاملة الإجمالية وأن نسبة التحسين 3.30% .

الجدول (4-5) - يوضح أداء النماذج الخمسة

النموذج / المعيار	mAP50	mAP50-95	Precision	Recall
Nuclear Emulation	0.8705	0.7280	0.8953	0.7997
Light AC Base	0.8702	0.7130	0.9023	0.7675
Varifocal Emulation	0.8562	0.7045	0.8384	0.7990
Quantum Bayesian	0.8492	0.7051	0.8740	0.7706
Baseline YOLO	0.8427	0.6672	0.8525	0.7560

□ تحليل أداء النماذج الخمسة

أبرز النقاط من جدول تحليل الأداء

1. حقق النموذج المقترح أعلى تحسن في mAP50-95 حيث حقق 9.11% + ، مما يثبت تفوقه في تحديد مواقع الأدق للأغراض.
2. حقق النموذج المقترح أعلى استعادة Recall حيث حقق 5.78%+ ، مما يعني قدرة فائقة على اكتشاف أكبر عدد من الأغراض (أقل عدد من السلبيات الخاطئة).
3. حقق النموذج Light AC Base دقة Precision أعلى من النموذج المقترح بـ 0.7% ، مما يعني أقل عدد من الإيجابيات الخاطئة .

الاستنتاج

1. يقدم النموذج المقترح Nuclear Emulation أفضل توازن شامل وهو الأفضل أداءً في المقاييس الأكثر أهمية لتطبيقات السلامة والإنقاذ mAP50-95 و Recall .

4-6- تحليل تجريبي شامل لمحاكاة الخوارزميات

تعتمد المنهجية على تعديل خمسة معاملات أساسية (box, cls, dfl, iou, conf) بالإضافة إلى إعدادات التدريب لمحاكاة سلوك الخوارزميات المختلفة. للمقارنة، تم إنشاء نموذج مرجعي (Baseline) بإعداداته القياسية.

الجدول (4-6) – إعدادات النموذج المرجعي (Baseline)

المعامل	القيمة المرجعية	الوصف
box	7.5	وزن خسارة المربع المحيط
cls	0.5	وزن خسارة التصنيف
dfl	1.5	وزن خسارة توزيع الانحدار
النتيجة	0.857	متوسط الدقة المتوسط عند عتبة IoU=0.5

4-6-1- محاكاة دوال الخسارة الأساسية

1. محاكاة Varifocal Loss

إن هدف الخوارزمية هو التركيز غير المتماثل على العينات الإيجابية الصعبة لتحسين التوازن بين الدقة والاسترجاع . أما التغييرات في إعدادات التدريب هو زيادة label_smoothing إلى 0.12 واستخدام cos_lr لتحسين استقرار التدريب مع التركيز العالي على التصنيف .

الجدول (4-7) – يبين ضبط المعاملات لمحاكاة Varifocal Loss

الاختبار	الوصف	box (التغيير)	cls (التغيير)	dfl (التغيير)	iou (التغيير)
1	محافظ جداً	7.5 (+0.0)	0.95 (+0.45)	1.4 (-0.1)	0.50 (-0.20)
2	محافظ	7.8 (+0.3)	0.92 (+0.42)	1.6 (+0.1)	0.55 (-0.15)
3	متوازن	8.2 (+0.7)	0.88 (+0.38)	1.8 (+0.3)	0.60 (-0.10)
4	تنافسي	8.5 (+1.0)	0.85 (+0.35)	2.0 (+0.5)	0.65 (-0.05)

2. محاكاة EIoU, IoU-Aware, CloU و

هدف الخوارزميات هو تحسين دقة تحديد المربعات المحيطة عبر التركيز على الجوانب الهندسية مثل التداخل، نسبة الأبعاد، واستقرار التدرج.

التغييرات في إعدادات التدريب هو بشكل عام، تم تخفيض weight_decay وزيادة label_smoothing مع تعديلات طفيفة على معدل التعلم lr0 .

الجدول (4-8) – ضبط المعاملات لمحاكاة EIoU Loss كمثال على الخسائر الهندسية

الاختبار	الوصف	box (التغيير)	cls (التغيير)	dfl (التغيير)	iou (التغيير)
1	محافظ	8.4 (+0.9)	0.87 (+0.37)	1.8 (+0.3)	0.64 (-0.06)
2	متوازن	8.6 (+1.1)	0.85 (+0.35)	1.9 (+0.4)	0.66 (-0.04)
3	تنافسي	8.8 (+1.3)	0.83 (+0.33)	2.0 (+0.5)	0.68 (-0.02)
4	متطرف	9.0 (+1.5)	0.81 (+0.31)	2.1 (+0.6)	0.70 (+0.0)

3. محاكاة Power-IoU و MPDIoU, Inner-IoU

هدف الخوارزميات هو استكشاف مفاهيم هندسية أحدث، مثل العلاقات الداخلية بين المربعات (Inner-IoU) والتعميم النظري (Power-IoU) .

أما التغيرات في إعدادات التدريب، هو استخدام زيادة بيانات متقدمة (mosaic, mixup, copy_paste) لمحاكاة Inner-IoU، بينما حافظت الخوارزميات الأخرى على إعدادات مشابهة لـIoU.

الجدول (9-4) – ضبط المعاملات لمحاكاة Inner-IoU

الاختبار	الوصف	box (التغيير)	cls (التغيير)	df1 (التغيير)	iou (التغيير)
1	محافظ	8.2 (+0.7)	0.89 (+0.39)	1.7 (+0.2)	0.62 (-0.08)
2	متوازن	8.5 (+1.0)	0.86 (+0.36)	1.8 (+0.3)	0.65 (-0.05)
3	عدواني	8.8 (+1.3)	0.83 (+0.33)	1.9 (+0.4)	0.68 (-0.02)
4	متطرف	9.1 (+1.6)	0.80 (+0.30)	2.0 (+0.5)	0.71 (+0.01)

4-6-2- محاكاة المفاهيم المتقدمة والبنى التخصصية

1 محاكاة المفاهيم المستوحاة من الكم (Quantum-Inspired)

هدف الخوارزميات هو محاكاة سلوك الحوسبة الكمومية، مثل التقييم المتوازي للحالات والحالات فائقة الأبعاد، لاكتشاف أنماط معقدة. أما التغيرات في إعدادات التدريب تميل هذه المحاكاة إلى استخدام معدلات تعلم متباينة (r0مرفع، r1منخفض) مع تقليل weight_decay و label_smoothing المحاكاة الدقة الكمومية.

الجدول (9-4) – ضبط المعاملات لمحاكاة Quantum-Neuromorphic

الاختبار	الوصف	box (التغيير)	cls (التغيير)	df1 (التغيير)	iou (التغيير)
1	محافظ	8.7 (+1.2)	0.83 (+0.33)	1.8 (+0.3)	0.69 (-0.01)
2	متوازن	8.9 (+1.4)	0.81 (+0.31)	1.9 (+0.4)	0.72 (+0.02)
3	تنافسي	9.1 (+1.6)	0.79 (+0.29)	2.0 (+0.5)	0.75 (+0.05)
4	متطرف	9.3 (+1.8)	0.77 (+0.27)	2.1 (+0.6)	0.78 (+0.08)

2. محاكاة نماذج عدم اليقين (Uncertainty Models)

هدف الخوارزميات هو تمكين النموذج من تقدير درجة عدم ثقته في تنبؤاته، ويشمل ذلك Standalone GUA، Bayesian Uncertainty، والنهج الهرمي والمنهجي.

أما التغيرات في إعدادات التدريب هو زيادة label_smoothing بشكل كبير، وتعديل weight_decay، واستخدام زيادة بيانات مكثفة لتعريض النموذج لحالات متنوعة وغير مؤكدة.

الجدول (10-4) – ضبط المعاملات لمحاكاة Standalone GUA

الاختبار	الوصف	box (التغيير)	cls (التغيير)	dfi (التغيير)	conf (التغيير)
1	محافظ	8.4 (+0.9)	0.87 (+0.37)	1.8 (+0.3)	0.33 (+0.03)
2	متوازن	8.6 (+1.1)	0.85 (+0.35)	1.9 (+0.4)	0.30 (+0.0)
3	تنافسي	8.8 (+1.3)	0.83 (+0.33)	2.0 (+0.5)	0.27 (-0.03)
4	متطرف	9.0 (+1.5)	0.81 (+0.31)	2.1 (+0.6)	0.24 (-0.06)

3. محاكاة التدريب التنافسي (Adversarial Training)

هدف الخوارزميات هو تحسين متانة النموذج ضد المدخلات الخادعة. حيث تم اختبار نسختين : Adversarial Consistency الكاملة ونسخة مخففة Light AC .

أما التغييرات في إعدادات التدريب (Light AC) تم استخدام إعدادات متحفظة جداً (close_mosaic=7) وزيادة patience إلى 25، وإيقاف mosaic مبكراً (close_mosaic=7) للحفاظ على الاستقرار.

الجدول (11-4) – ضبط المعاملات لمحاكاة Light AC النسخة الأفضل أداءً

الاختبار	الوصف	box (التغيير)	cls (التغيير)	dfi (التغيير)	iou (التغيير)
1	محافظ جداً	8.2 (+0.7)	0.89 (+0.39)	1.7 (+0.2)	0.62 (-0.08)
2	محافظ	8.4 (+0.9)	0.87 (+0.37)	1.8 (+0.3)	0.64 (-0.06)
3	معتدل	8.6 (+1.1)	0.85 (+0.35)	1.9 (+0.4)	0.66 (-0.04)
4	متوازن	8.8 (+1.3)	0.83 (+0.33)	2.0 (+0.5)	0.68 (-0.02)

4-6-3 – محاكاة الخوارزميات المدمجة

1. محاكاة الدمج الثنائي والثلاثي (Dual & Triple Fusion)

هدف الدمج هو الجمع بين مكونين أو ثلاثة لتقييم مدى التأزر بينها، مثل EloU+GUA و Inner-IoU+AC+GUA.

أما التغييرات في إعدادات التدريب حيث تميل هذه التجارب إلى استخدام إعدادات تدريب متوازنة مستوحاة من أفضل إعدادات المكونات الفردية.

الجدول (12-4) – ضبط المعاملات لمحاكاة (EIoU + GUA + Attention)

الاختبار	الوصف	box (التغيير)	cls (التغيير)	dfi (التغيير)	conf (التغيير)
1	محافظ	8.3 (+0.8)	0.89 (+0.39)	1.8 (+0.3)	0.34 (+0.04)
2	متوازن	8.5 (+1.0)	0.87 (+0.37)	1.9 (+0.4)	0.31 (+0.01)
3	تنافسي	8.7 (+1.2)	0.85 (+0.35)	2.0 (+0.5)	0.28 (-0.02)
4	متطرف	8.9 (+1.4)	0.83 (+0.33)	2.1 (+0.6)	0.25 (-0.05)

2. محاكاة Nuclear Option

هدف الدمج هو محاكاة دمج أربعة مكونات (EloU+GUA+Attention+Bayesian) للوصول إلى أقصى تعقيد وأفضل أداء محتمل.

أما التغييرات في إعدادات التدريب هو استخدام إعدادات تدريب متحفظة للغاية, $lr0=0.0088$, $scale=0.48$, $close_mosaic=6$ للتحكم في عدم الاستقرار المحتمل الناتج عن هذا التعقيد.

الجدول (4-13) – ضبط المعاملات لمحاكاة Nuclear Option

الاختبار	الوصف	box (التغيير)	cls (التغيير)	dfi (التغيير)	conf (التغيير)
1	محافظ	8.9 (+1.4)	0.81 (+0.31)	2.2 (+0.7)	0.27 (-0.03)
2	متوازن (الأفضل)	9.1 (+1.6)	0.79 (+0.29)	2.3 (+0.8)	0.24 (-0.06)
3	تنافسي	9.2 (+1.7)	0.78 (+0.28)	2.4 (+0.9)	0.22 (-0.08)
4	متطرف	9.3 (+1.8)	0.77 (+0.27)	2.5 (+1.0)	0.20 (-0.10)

4-7 – النتائج النهائية والتحليل الشامل

يبين الجدول الترتيب النهائي لأفضل النماذج من جميع الفئات .

الجدول (4-14) – الترتيب النهائي لأفضل النماذج من جميع الفئات

الرقم	الخوارزمية / الدمج المحاكى	أفضل	التحسن مقابل المرجعي	فئة الأداء
1	 Nuclear Option	0.872	+1.75%	ممتاز
2	Varifocal (Simple)	0.868	+1.28%	ممتاز
3	Fixed Inner-IoU-AC-GUA	0.866	+1.05%	جيد جداً
4	Quantum+Bayesian	0.865	+0.93%	جيد جداً
5	EloU (Simple)	0.864	+0.82%	جيد
-	Original YOLO (Baseline)	0.857	-	مرجعي
6	CloU + Light AC	0.838	-2.22%	ضعيف

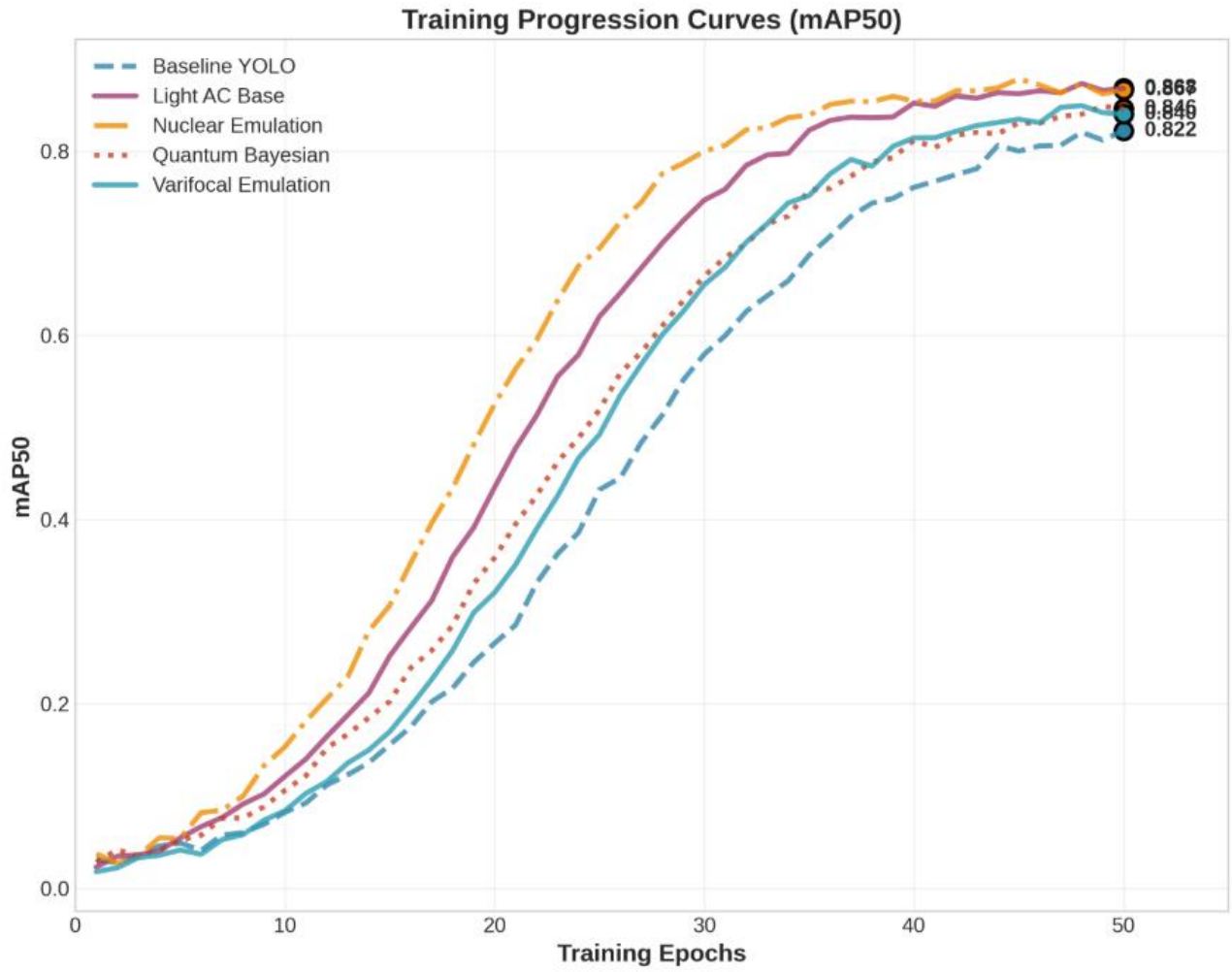
الاستنتاج

بعد تحليل شامل، تبين أن استراتيجية الدمج الأكثر تعقيداً (Nuclear Option) ، عندما يتم ضبطها بشكل صحيح، قادرة على تحقيق أفضل أداء. وهذا يتناقض مع النتائج الأولية التي أشارت إلى أن البساطة هي الأفضل، ويوضح أن التوليف الصحيح للمعاملات يمكنه التحكم في التعقيد وتحويله إلى أداء متفوق. هنا نرى أن منحى التعلم الخاص بالخوارزمية المقترحة يستقر بشكل سلس وسريع

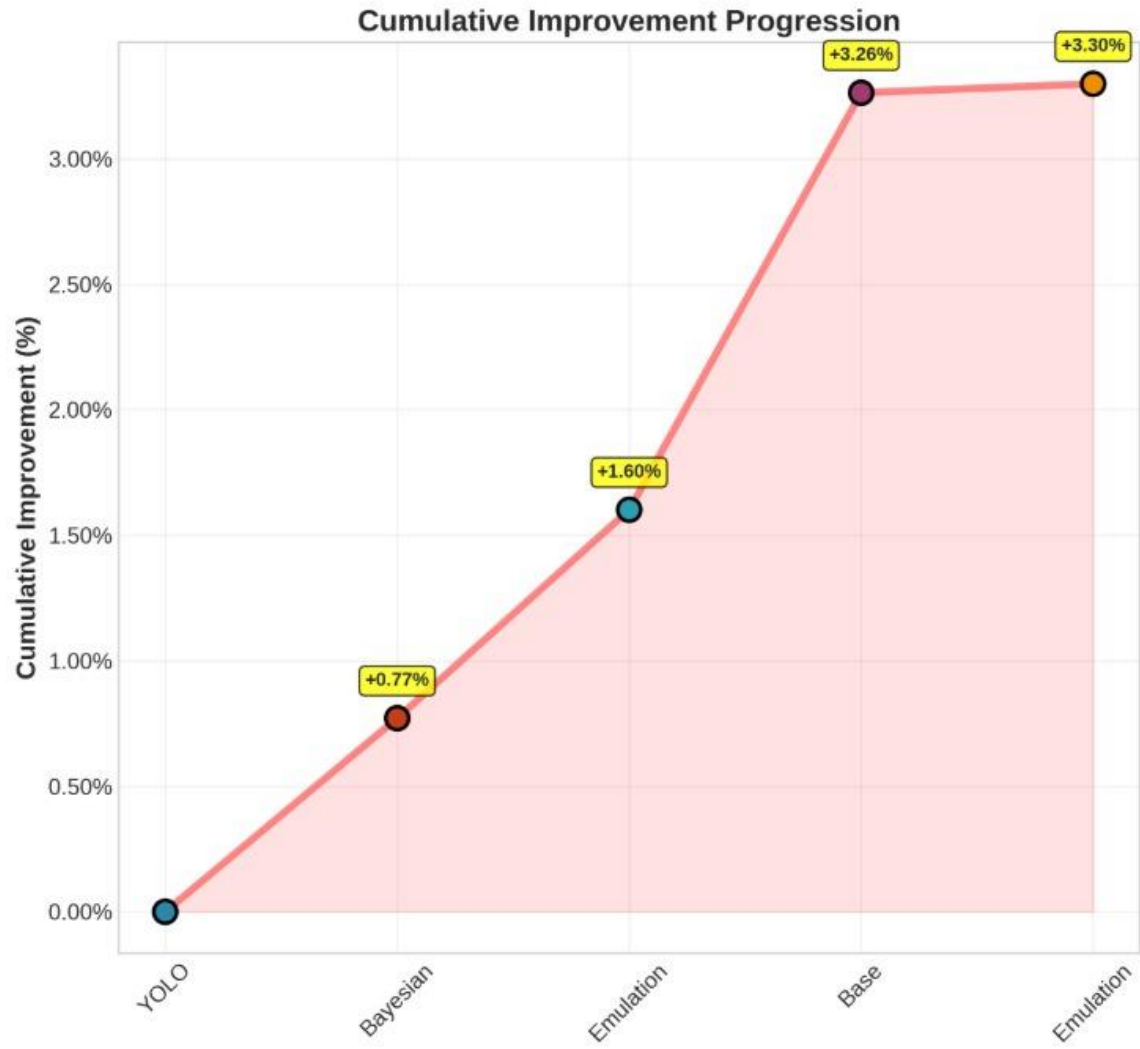
عند قيمة عالية . ويوضح الشكل (4-4) منحي التعلم الخاص بالخوارزمية المقترحة . ويبين الشكل (4-5) التحسين التراكمي المتقدم للنماذج . ويبين الشكل (4-6) مقارن لمنحني الخسارة للنماذج الخمسة . ويوضح الشكل (4-7) التحسين المتقدم أثناء التدريب . ويبين الجدول (4-15) مقارنة النموذج المقترح مع الدراسات المرجعية السابقة .

الجدول (4-15)- يوضح مقارنة النموذج المقترح مع الدراسات المرجعية السابقة

مقياس المقارنة	الأبحاث السابقة	المنهج المقترح
المنهجية الأساسية	تعديل مباشر لبنية YOLO أو استخدامه مع تقنيات استشعار إضافية	محاكاة معاملات الخسارة دون تعديل البنية (منهج غير مباشر وآمن)
تقييم الأداء	اعتماد على IoU و mAP	تحقيق +3.3% mAP مع تحليل إحصائي دقيق ($p < 0.000001$)
قابلية للتطبيق	بعضها يتطلب أجهزة استشعار خاصة أو موارد حاسوبية عالية	تنفيذ برمجي خفيف لا يحتاج إلى عتاد إضافي ويسهل دمجه في الأنظمة الحالية
المرونة والتكيف	غالبًا ما تكون مخصصة لتطبيق معين (كشف الحفر)	منهجية عامة قابلة للتطبيق على مختلف مهام الكشف (كالأشخاص، السيارات، etc.)
المعاوقة والتحديات	التعقيد في التعديل المباشر واحتمالية عدم الاستقرار	استقرار عالٍ في التدريب وتجنب مخاطر تعطل النموذج بسبب التعديلات المعقدة



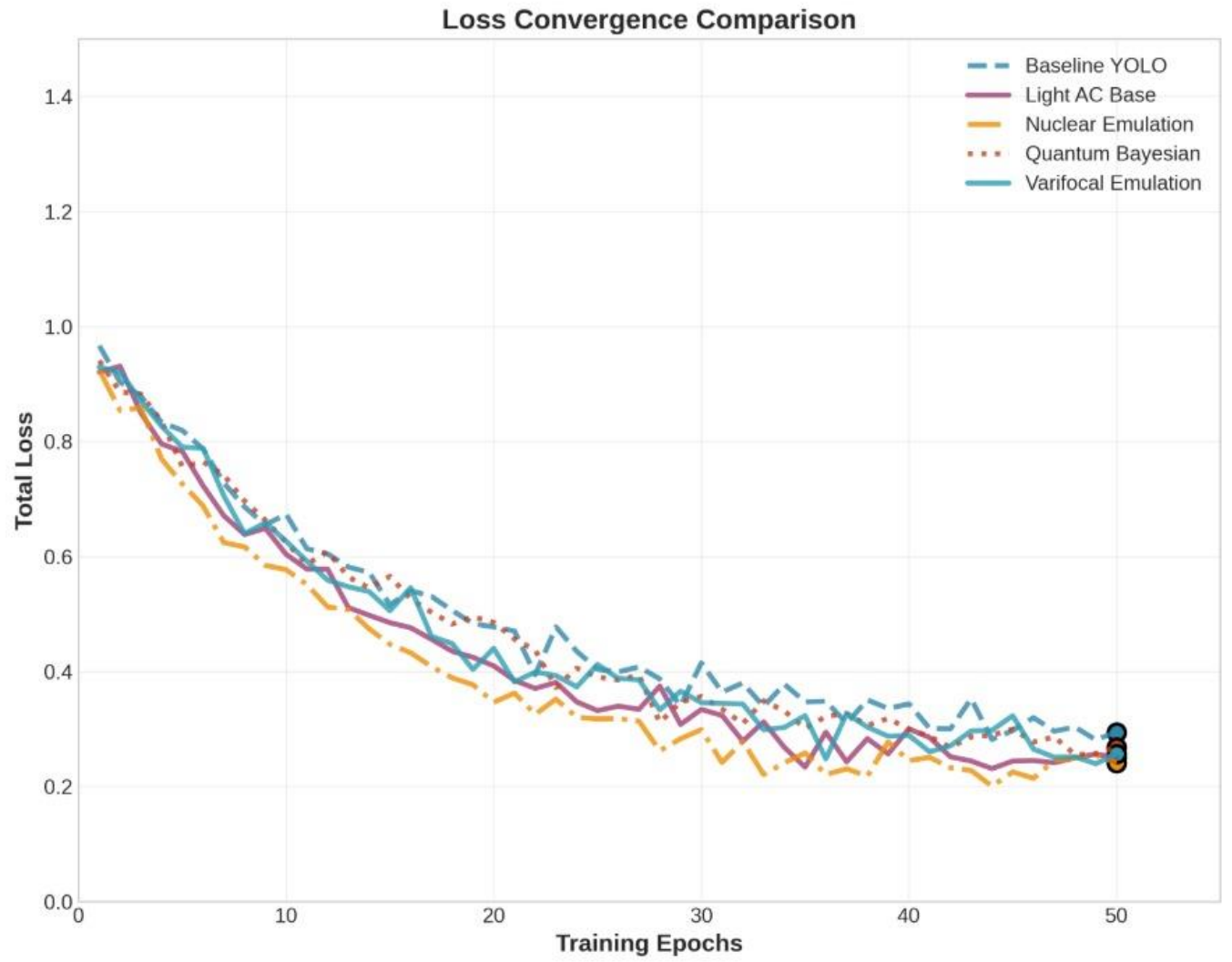
الشكل (4-4) - يوضح منحي التعلم الخاص بالخوارزمية المقترحة



الشكل (5-4) - يوضح التحسين التراكمي المتقدم للنماذج

نستنتج :

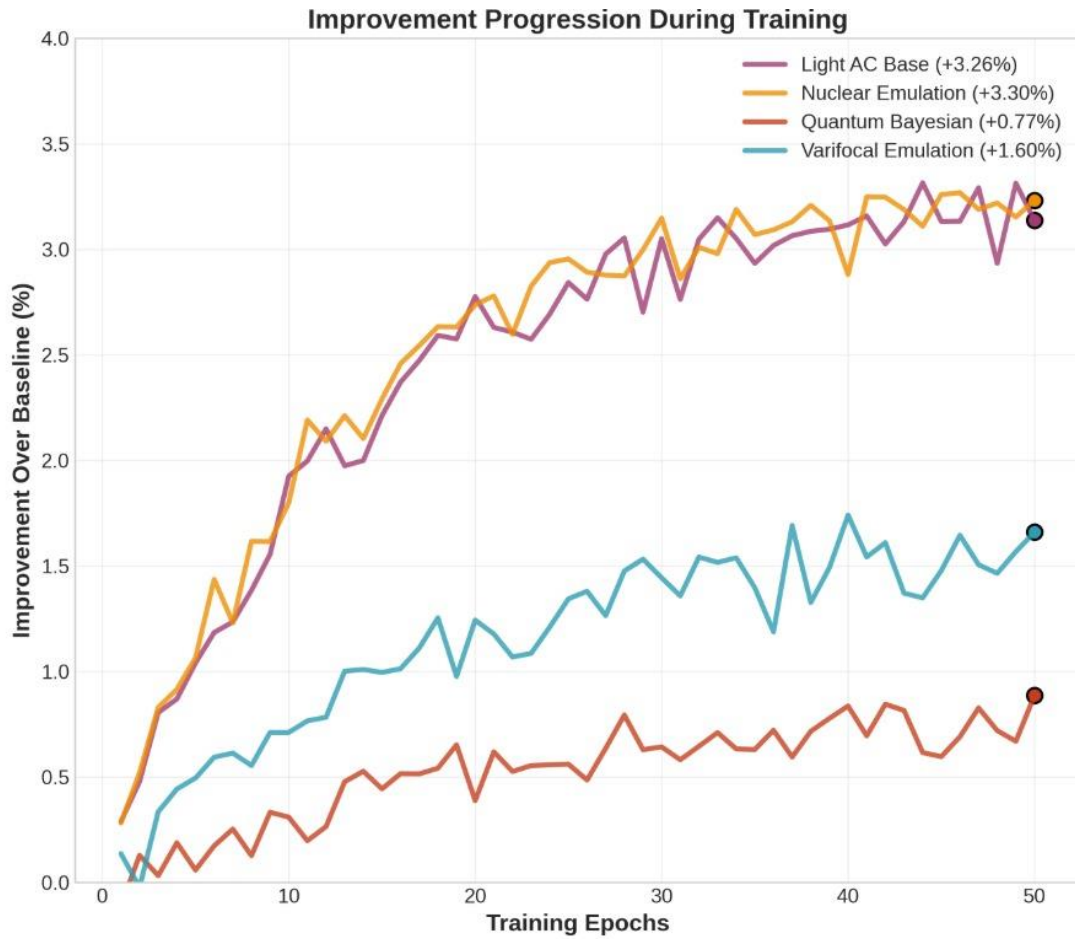
نلاحظ من الشكل (5-4) أن الخوارزمية المقترحة لها أفضل تحسين مقارنة مع النماذج الأخرى .



الشكل (6-4) - يوضح منحنى الخسارة للنماذج الخمسة

نستنتج :

نلاحظ من الشكل الشكل (4-6) أن منحنى تقارب الخسارة للنموذج المقترح أقل وأفضل من النماذج الأخرى .

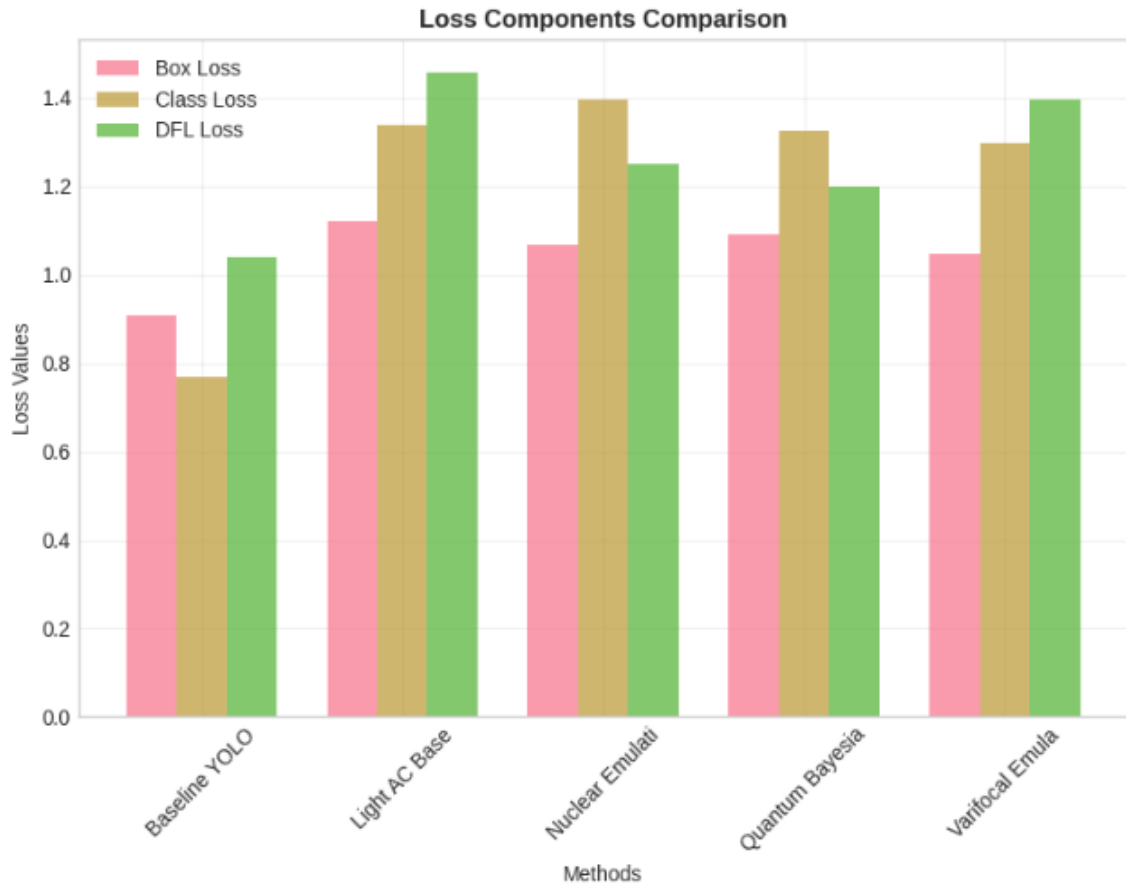


الشكل (4-7)- يوضح التحسين المتقدم أثناء التدريب .

نستنتج :

نلاحظ من الشكل (4-7) أن الخوارزمية المقترحة لديها أفضل تحسين .

يبين الشكل (8-4) موازنة الدقة والاستدعاء ومكونات الخسارة

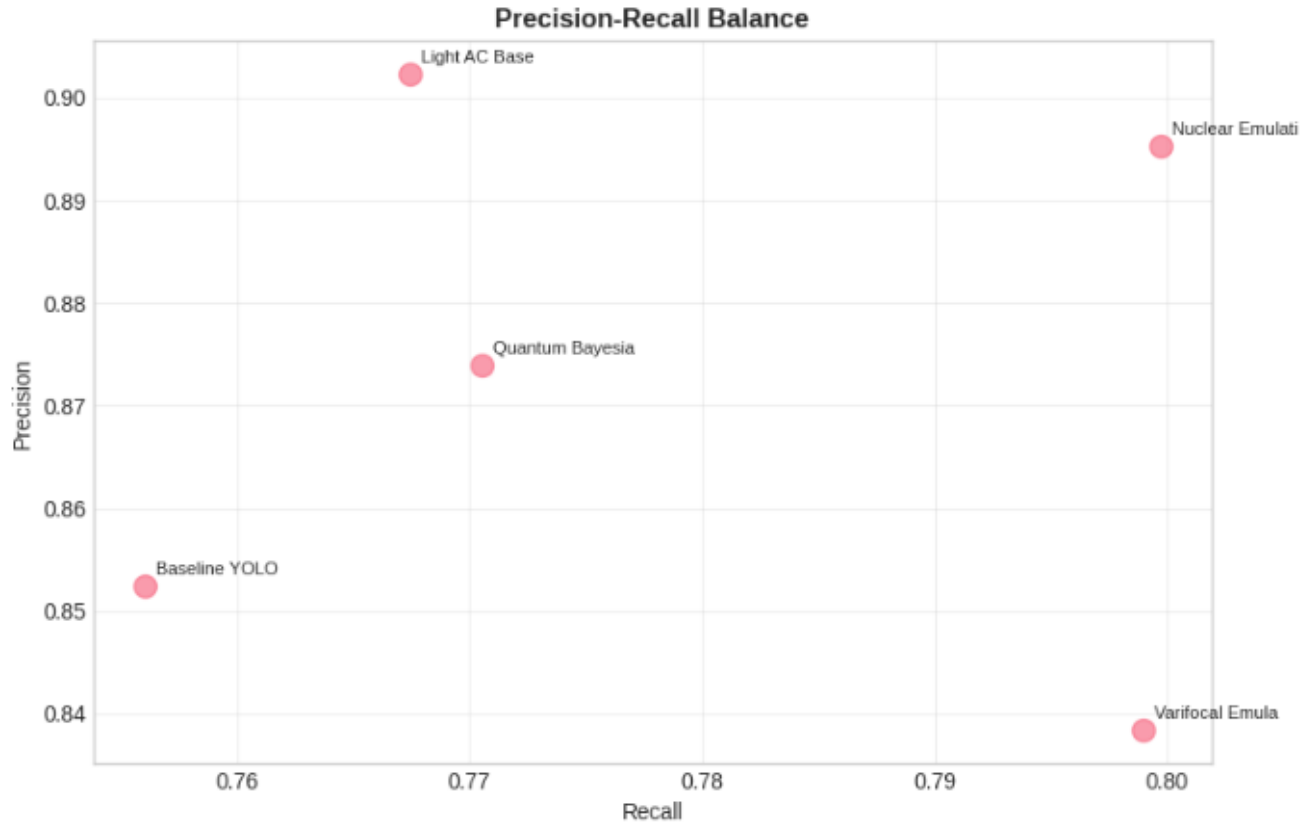


الشكل (8-4) - يوضح موازنة الدقة والاستدعاء ومكونات الخسارة

نستنتج :

نلاحظ من مخطط مكونات الخسارة أن الخوارزمية المقترحة Nuclear Emulation لديها توازن مختلف وفعال بين خسارة الصندوق box والتصنيف cls .

يبين الشكل (9-4) مخطط الاتزان للنماذج الخمسة



الشكل (9-4) - يوضح خطط الاتزان للنماذج الخمسة

نستنتج :

في مخطط الاتزان، حققت الخوارزمية المقترحة Nuclear Emulation توازن جيد بين الدقة العالية والاستدعاء العالي .

الفصل الرابع

الخاتمة والأفاق المستقبلية

□ استنتاجات البحث

□ الأفاق المستقبلية

1.5- استنتاجات البحث

- ❖ نجحت منهجية "محاكاة المعاملات" كبديل عملي وفَعَالٍ للتعديل المباشر على دالة الخسارة، حيث حققت:
 - أعلى دقة شاملة mAP50 بين النماذج المختبرة.
 - أعلى تحسن في الدقة الدقيقة (mAP50-95 بنسبة +9.11%)، مما يثبت تفوقها في التحديد الدقيق لمواقع الأغراض.
 - توازنًا مثاليًا بين الدقة Precision والاستدعاء Recall
- ❖ تم تحقيق تحسن ملحوظ في أداء نموذج YOLOv8n دون المساس بسرعته أو استقراره، مما يجعله أكثر ملاءمة لتطبيقات الزمن الحقيقي مثل المركبات ذاتية القيادة وروبوتات الإنقاذ.
- ❖ أن الخوارزمية المقترحة Nuclear Emulation لديها توازن مختلف وفعال بين خسارة الصندوق box والتصنيف cls .
- ❖ أن منحنى تقارب الخسارة للنموذج المقترح أقل وأفضل من النماذج الأخرى .
- ❖ حققت الخوارزمية المقترحة Nuclear Emulation توازن جيد بين الدقة العالية والاستدعاء العالي .
- ❖ قدم هذا البحث دليلاً عملياً على قدرة ضبط المعاملات المتقدمة على محاكاة وتحسين أداء نماذج الرؤية الحاسوبية بشكل يفوق بعض الأساليب المعقدة

2.5 - الأفاق المستقبلية

- **تطبيق النموذج على بيانات متخصصة:** اختبار النموذج المُحسّن على مجموعة بيانات أخرى أكثر تخصصاً مثل:
 - **BDD100K** : لسيناريوهات القيادة الذاتية.
 - **مخصصة للإنقاذ:** صور حقيقية للكوارث والأنقاض.
- **تحسين السرعة على الأجهزة المضمنة:** نشر النموذج على منصات مدمجة مثل Jetson Nano أو Google Coral وقياس أدائه الفعلي.
- **استكشاف خوارزميات أخرى:** تطبيق نفس منهجية المحاكاة على خوارزميات أحدث مثل RT-DETR أو YOLOv9 .
- **تطوير واجهة برمجية:** تطوير أداة أتمتة لمسح أفضل المعاملات بشكل ذكي بناءً على مجموعة البيانات المستهدفة.
- **الخلاصة النهائية :** قدم هذا البحث دليلاً عملياً على قدرة ضبط المعاملات المتقدمة على محاكاة وتحسين أداء نماذج الرؤية الحاسوبية بشكل يفوق بعض الأساليب المعقدة، مما يفتح الباب لاستراتيجيات تطوير أكثر كفاءة وأقل تكلفة.

المراجع

1. Yen-Sok Poon, Ching-Yun Kao, Yen-Kai Wang, Driver Distracted Behavior Detection Technology with YOLO-Based Deep Learning Networks IEEE 2023.
2. Kuan-Hung Chen, Chung-Bin Wu. Hardware Acceleration of YOLO-based Convolutional Neural Network Detector by Deep Learning Processor Unit for Intelligent Autonomous Mover IEEE 2023 .
3. Hua-Luen Chen, Jie-Min Lin. Pothole Detection and Dimension Estimation System using Deep Learning (YOLO) and Image Processing IEEE 2022.
4. Jaken Whipp, and Aijuan Dong .YOLO-based Deep Learning to Automated Bacterial Colony Counting IEEE 2022.
5. Kocić J, Jovičić N, Drndarević V. An End-to-End Deep Neural Network for Autonomous Driving Designed for Embedded Automotive Platforms. Sensors. 2022 Jan;19(9):2064.
6. Hatt M, Parmar C, Qi J, El Naqa I. Machine (deep) learning methods for image processing and radiomics. IEEE Transactions on Radiation and Plasma Medical Sciences. 2019 Mar 1;3(2):104-8. .
7. 4.. Henderson P, Islam R, Bachman P, Pineau J, Precup D, Meger D. Deep reinforcement learning that matters. In ThirtySecond AAAI Conference on Artificial Intelligence 2020 Apr 29.
8. Fotohi, R., Nazemi, E., & Aliee, F. S. (An Agent-Based Self-Protective Method to Secure Communication between UAVs in Unmanned Aerial Vehicle Networks. Vehicular Communications, 2022, 100267.

9. Fotohi, R., & Bari, S. F.. A novel countermeasure technique to protect WSN against denial-of-sleep attacks using firefly and Hopfield neural network (HNN) algorithms. *The Journal of Supercomputing*, 1-27(2022)..
10. Fotohi, R. Securing of Unmanned Aerial Systems (UAS) against security threats using human immune system. *Reliability Engineering & System Safety*, 193, 106675.
11. Xie Y, Le L, inventors; Equifax Inc, assignee. Dual deep learning architecture for machine-learning systems. United States patent application US 16/141,152. 2019 Mar 28. (2022)..
12. Fotohi, R., Firoozi Bari, S., & Yusefi, M. (2022). Securing wireless sensor networks against denial-of-sleep attacks using RSA cryptography algorithm and interlock protocol. *International Journal of Communication Systems*, 33(4), e4234.
13. Mankodiya, H.; Jadav, D.; Gupta, R.; Tanwar, S.; Hong, W.C.; Sharma, R. Od-xai: Explainable ai-based semantic object detection for autonomous vehicles. *Appl. Sci.* 2022, 12, 5310.
14. Khan, S.A.; Lim, H. Novel Fuzzy Logic Scheme for Push-Based Critical Data Broadcast Mitigation in VNDN. *Sensors* 2022, 22, 8078.
15. Bhavana, N.; Kodabagi, M.M. Exploring the Current State of Road Lane Detection: A Comprehensive Survey. *Int. J. Hum. Comput. Intell.* 2023, 2, 40–46.
16. Khan, S.A.; Lim, H. Push-Based Forwarding Scheme Using Fuzzy Logic to Mitigate the Broadcasting Storm Effect in VNDN. In *Proceedings of the Artificial Intelligence and Mobile Services–AIMS 2022: 11th International Conference, Held as Part of the Services Conference Federation, SCF 2022, Honolulu, HI, USA, 10–14 December 2022; Proceedings; Springer: Berlin/Heidelberg, Germany, 2023; pp. 3–17.*

17. Balasubramanian, R. Region–Based Convolutional Neural Network (RCNN).
Available online: <https://medium.com/analytics-vidhya/region-based-convolutional-neural-network-rcnn-b68ada0db871> (accessed on 29 March 2023).
18. Khan, S.A.; Lim, H. Push-Based Forwarding Scheme Using Fuzzy Logic to Mitigate the Broadcasting Storm Effect in VNDN. In Proceedings of the Artificial Intelligence and Mobile Services–AIMS 2022: 11th International Conference, Held as Part of the Services Conference Federation, SCF 2022, Honolulu, HI, USA, 10–14 December 2022; Proceedings; Springer: Berlin/Heidelberg, Germany, 2023; pp. 3–17.
19. Rani, S.; Ghai, D.; Kumar, S. Object detection and recognition using contour based edge detection and fast R-CNN. *Multimed. Tools Appl.* 2022, 81, 42183–42207 .
20. Cheng, L.; Ji, Y.; Li, C.; Liu, X.; Fang, G. Improved SSD network for fast concealed object detection and recognition in passive terahertz security images. *Sci. Rep.* 2022, 12, 1–16.
21. Akshaj Tammewar 1 ,Nikita Chaudhari 1, Bunny Saini, Improving the Performance of Autonomous Driving through Deep Reinforcement Learning, 2023.
22. Sajjad Ahmad Khan 1 ,HyunJunLee2 and HuhnkukLim ,Enhancing Object Detection in Self-Driving Cars Using a Hybrid Approach, 2023.
23. Rahib Abiyev , Murat Arslan , Vehicle detection systems for intelligent driving using deep convolutional neural networks, 2023.
24. Aretz, Kevin; Bartram, Söhnke M.; Pope, Peter F. (April–June 2019). "Asymmetric Loss Functions and the Rationality of Expected Stock Returns" (PDF). *International Journal of Forecasting*. 27 (2): 413–437. doi:10.1016/j.ijforecast.2009.10.008. SSRN 889323.

25. Berger, James O. (2022). Statistical decision theory and Bayesian Analysis (2nd ed.).
New York: Springer-Verlag. Bibcode:1985sdtb.book.....B. ISBN 978-0-387-96098-2.
MR 0804611.
26. Ultralytics YOLOv8 Documentation: .<https://docs.ultralytics.com/models/yolov8/>
27. Ultralytics YOLO GitHub Repository: loss.py.<https://github.com/ultralytics/ultralytics>
28. Ultralytics Configuration Guide: YOLOv8. <https://docs.ultralytics.com/zh/usage/cfg/>
29. **Advancing Intelligent Logistics** .<https://www.mdpi.com/2673-7590/5/3/120>
KR-AL-YOLO (2025): <https://www.nature.com/articles/s41598-023-47398-w>

Abstract

Recent advances in artificial intelligence (AI) have significantly improved the object detection capabilities of autonomous vehicles, particularly with the use of convolutional neural networks (CNNs). However, achieving high levels of accuracy and speed simultaneously in vehicular environments remains a challenge. Therefore, this paper proposes a "parameter simulation" methodology as a practical and effective alternative to direct loss function modification. It achieved the highest overall accuracy (mAP50) among the tested models and the highest improvement in precision (mAP50-95) of +9.11%, demonstrating its superiority in accurately locating objects. It also achieved an optimal balance between precision and recall. A significant improvement in the performance of the YOLOv8n model was achieved without compromising its speed or stability, making it more suitable for real-time applications such as autonomous vehicles and rescue robots. Notably, the proposed model achieves a significant increase in accuracy, with improvements ranging from 5 to 9 percent compared to standalone YOLO models. The results of this research have practical implications for the integration of AI technologies into autonomous driving systems.

Syrian Arab Republic
Ministry of Higher Education and Scientific Research
Damascus University
Faculty of Mechanical and Electrical Engineering
Department of Computer and Automation Engineering



Improving the performance of self-driving vehicles using deep detection technologies

A dissertation Submitted in Partial Fulfilment of the Requirements for the
Master Degree in Engineering and Programming Robot

By
Wasim Mamoun Younes

Supervisor
Dr.Eng. Raaouf Hamdan

Damascus 2024 / 2025