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

تطوير خوارزمية تمييز الكائنات المتداخلة في الزمن الحقيقي باعتماد السمات الطيفية الحرارية والحركية

أطروحة مقدمة لنيل درجة الدكتوراه في هندسة الإلكترونيات التطبيقية

إعداد

المهندس عبده شحاده هلاله

المشرف العلمي

الدكتور عبد الرزاق خالد البدوية

المشرف المشارك

الدكتور شادي خليل البيطار

العام الدراسي

2025-2024

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

Real Time Algorithm Development of Overlapping Objects Recognition Based On Thermal and Kinetic Features

A thesis Prepared for Obtaining a PhD Degree in Applied Electronics Engineering

Prepared by

Eng. Abdo Shehada Hilala

Scientific Supervisor

Prof. Dr. Abdelrazak Khaled Badawieh

Co- Supervisor

Prof. Dr. Shadi Khalil Al Bitar

Academic Year

2024-2025

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

بطاقة شكر

أتقدم بأسمى آيات الشكر والامتنان لأعضاء لجنة الحكم الموقرة لقبولهم قراءة خلاصة بحثي والحكم على أطروحة الدكتوراه التي أنجزتها.

أعضاء لجنة الحكم

الدكتور عبد الرزاق البدوية	عضواً ومشرفاً
الدكتور عبد الكريم السالم	عضواً
الدكتور طلال حمود	عضواً
الدكتور نضال زيدان	عضواً
الدكتورة جمان أبو جيب	عضواً

المشكلة بموجب قرار مجلس البحث العلمي والدراسات العليا رقم /3258/ المتخذ بالجلسة رقم /23/ تاريخ 28/ 7 /2025

د. عبد الرزاق البدوية د. عبد الكريم السالم د. طلال حمود د. نضال زيدان د. جمان أبو جيب

تصريح خطي

أنا الموقع أدناه أصرح بعلمي الكامل وقبولي بـ:

1- أن كل ما ينتج من البحث والاطروحة هو ملكية فكرية ومالية بالتشارك مع جامعة دمشق، واني ألتزم بأخذ موافقة الجامعة في حال رغبتي بنشر البحث أو الأطروحة أو جزء منها نصاً أو مضموناً خارج إطار الجامعة (من دور نشر أو مكتبات أو مواقع إلكترونية وغيرها من وسائل النشر)

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

الاسم والتوقيع

م. عبده شحاده هلاله

شكر وتقدير

أتقدم بخالص الشكر والتقدير لكل من كان له الفضل في انجاز هذا البحث وأخص بالشكر:

- أستاذي المشرف الدكتور عبد الرزاق البدوية لدعمه العلمي والمعنوي وتوجيهاته التي بذلها لكي يبصر هذا العمل النور.
- أستاذي المشرف المشارك الدكتور شادي البيطار الذي لم يبخل بالوقت والجهد والمعرفة والدعم لإيصال البحث إلى نهايته.
- أساتذتنا في الهيئة التدريسية في قسم هندسة الالكترونيات والاتصالات بكلية الهندسة الميكانيكية والكهربائية بجامعة دمشق على دعمهم وتوجيهاتهم.
- المعهد العالي للعلوم التطبيقية والتكنولوجيا على دعمه وتقديم الوسائل المساعدة لوصول البحث إلى نتائجه.

كما أتقدم بالشكر الجزيل لأعضاء لجنة الحكم الأفاضل لملاحظاتهم وإرشاداتهم التي أغنت الاطروحة.

م. عبده شحاده هلاله

المحتويات

الفصل الأول: الإطار العام للبحث 1
1-1 مقدمة 2
2-1 الدراسات المرجعية 2
3-1 مشكلة البحث 6
4-1 هدف البحث 7
5-1 أهمية البحث 7
6-1 مسوغات البحث 8
7-1 محددات البحث 8
8-1 حدود البحث 8
9-1 منهج البحث وأدواته 9
10-1 بنية الاطروحة 9
الفصل الثاني: الصورة الرقمية وتجزئتها 11
1-2 مصادر الصورة 12
2-2 الصورة الرقمية 14
1-2-2 تمييزية الصورة 15
2-2-2 التوابع الهامة 16
3-2 المرشحات المكانية الخطية 17
1-3-2 الترابط والتلاف المكاني 18
2-3-2 مرشح التمرير المنخفض 18
3-3-2 نوى مرشح غاوص منخفض التردد 19
4-3-2 التنعيم التكيفي 20
5-3-2 مرشح التمرير المرتفع (المرشحات الحدية) 20
4-2 تجزئة الصورة 21
1-4-2 كشف النقاط والخطوط والحواف 22
2-4-2 تابع العتبة 23
5-2 العمليات المورفولوجية 24
1-5-2 الحت 25
2-5-2 التمددDilation 25
3-5-2 الفتح والاغلاق 27
الفصل الثالث: بنى وخوارزميات ملاحقة الكائنات المتعددة 29
1-3 السمات وطرائق تحديدها 30
1-1-3 خوارزميات كشف الحواف 30
2-1-3 بعض خوارزميات كشف الزوايا 33

2-3 الذكاء الاصطناعي 35
1-2-3 تعلم الآلة 35
2-2-3 التعلم العميق 36
3-3 خوارزميات كشف الكائنات 37
1-3-3 الشبكات العصبونية التلافيفية CNN 38
2-3-3 خوارزمية الكشف YOLO 42
4-3 خوارزميات ملاحقة الكائنات المتعددة MOT 46
1-4-3 المبدأ العام لعمل خوارزمية ملاحقة الكائنات MOT 47
2-4-3 أنواع خوارزميات ملاحقة الكائنات المعتمدة على التعلم العميق 50
5-3 نماذج الحركة 53
1-5-3 النموذج الديناميكي 54
2-5-3 مرشح $\boldsymbol{\alpha - \beta}$ 55
3-5-3 مرشحات α−β−γ 57
4-5-3 مرشح كالمن 58
6-3 خوارزميات الملاحقة 63
1-6-3 خوارزمية الملاحقة SORT 63
2-6-3 خوارزمية DeepSORT 64
3-6-3 معايير الاختبار لخوارزمية ملاحقة الكائنات 64
الفصل الرابع: الخوارزمية المقترحة لملاحقة وتمييز الكائنات الحرارية المتداخلة 68
1-4 مشاكل عملية ملاحقة الكائنات 70
1-1-4 سرعة الملاحقة 70
2-1-4 تشويه الخلفية 70
3-1-4 مقاييس مكانية متعددة 70
4-1-4 مشكلة ربط المسارات 70
5-1-4 الاحتجاب 71
2-4 ملاحقة الكائنات الحرارية 72
1-2-4 ميزات العمل في مجال الصورة الحرارية 72
2-2-4 مشاكل الملاحقة ضمن الصور الحرارية 73
3-2-4 تمييز الكائنات ضمن الصور الحرارية 73
3-4 إطار العمل والحلول 74
1-3-4 تعزيز سمات الكائنات الحرارية المكتشفة 74
2-3-4 استخدام سمات الحركة 78
3-3-4 إعادة تعريف الكائن ضمن الصور الحرارية 84
4-3-4 المنهجية الأساسية للخوارزمية المقترحة 90
الفصل الخامس: التنفيذ والنتائج العملية 98
1-5 أدوات البحث 99

5-1-1 العتاد المستخدم 99
5-1-2 البرمجيات 99
5-1-3 المكتبات..... 100
5-2 مجموعات البيانات 101
5-3 مقاييس الأداء..... 102
5-3-1 معيار دقة الحل 102
5-3-2 معيار معدل الخطأ..... 103
5-3-3 معيار زمن التنفيذ 103
5-4 إجراءات التنفيذ العملي 104
5-4-1 موافقة الصورة الحرارية..... 104
5-4-2 الملاحقة بالخوارزمية المطورة الأساسية..... 113
5-4 تحسين الخوارزمية المطورة..... 114
5-4-1 مشكلة مركز ثقل الكائن وحلها 114
5-4-2 مشكلة عدم ثبات شكل الكائن الحراري أثناء الحركة وطريقة الحل..... 115
5-4-3 مشكلات نموذج الحركة والاجراءات 116
5-4-4 إجراءات التحسين على الخوارزمية..... 117
5-4-5 اختبارات الخوارزمية المحسنة..... 118
5-5 الاستنتاجات..... 126
5-6 التوصيات 126
الأبحاث المنشورة..... 127
الآفاق المستقبلية 127
قائمة المراجع..... 129

فهرس الأشكال

الرقم	وصف	الصفحة
1-2	القدرة على تمييز الكائنات في الصورة (Szeliski, 2022)	12
2-2	الطيف المرئي ضمن الطيف الكهرومغناطيسي (Frank, 2023)	13
3-2	الطيف المرئي وطيف الأشعة تحت الحمراء (Lea, 2011)	13
4-2	تحصيل الصورة الرقمية (Stanford, 2024)	14
5-2	عدد البايتات التي تحتاجها الصوة للتخزين (Szeliski,2022)	15
6-2	هيستوغرام لصور عائدة لصورة وحيدة بظروف مختلفة (Gonzalez,2018)	16
7-2	قالب المرشح (Gonzalez, 2018)	17
8-2	قالب مرشح التنعيم (Gonzalez, 2018)	19
9-2	تابع التوزيع الغاوصي (Rudrabhatla, 2020)	19
10-2	قالب المرشح الغاوصي 3 x 3 (Gonzalez, 2018)	20
11-2	المرشح اللابلاسي (Hardan, 2018)	21
12-2	المشتق الأول والثاني (Hardan, 2018)	22
13-2	قالب Sobel لكشف الحواف (Hande, 2021)	23
14-2	ناتج كشف الحواف (Kalipayun, 2018)	32
15-2	تابع العتبة في الهيستوغرام (Gonzalez,2018)	24
16-2	استخدام الهيستوغرام في التجزئة (Gonzalez,2018)	24
17-2	الحت (Gonzalez,2018)	25
18-2	مثال عن التمدد (Gonzalez,2018)	25
19-2	ناتج التمدد (Gonzalez,2018)	26
20-2	اصلاح الأحرف المتآكلة بالتمدد (Gonzalez,2018)	26
21-2	الحت والتمدد لإزالة الأجزاء الصغيرة (Tarabalka,2022)	26
22-2	تقنية الفتح (Gonzalez,2018)	27
23-2	تقنية الاغلاق (Gonzalez,2018)	27

24-2	بعض الاصلاحات التي تقدمها تقنيات الفتح والإغلاق (Gonzalez,2018)	28
1-3	الحافة تمثل انتقال بالمستوى اللوني (Malche, 2024)	31
2-3	كشف الحواف (Tian, 2021)	31
3-3	كشف الحواف بخوارزمية Sobel (Yuheng, 2017)	32
4-3	كشف الزوايا (Juranek, 2018)	33
5-3	مصفوفة لاختبار Moravec (Thapa, 2022)	33
6-3	حالات اختبار المنطقة المنبسطة والحافة الشاقولية والافقية (Thapa, 2022)	34
7-3	مبدأ الكشف السريع للزاوية (Medium, 2019)	34
8-3	التعلم العميق في الذكاء الاصطناعي (الباحث)	35
9-3	البرمجة التقليدية (الباحث)	36
10-3	التعلم الآلي (الباحث)	36
11-3	الذكاء الصناعي وتعلم الآلة (Seseri, 2023)	37
12-3	بنية الشبكة العصبونية (Aabidi,2023)	38
13-3	نموذج أولي لشبكة عصبونية اصطناعية (Liang, 2020)	39
14-3	مثال على تابع تفعيل (Ganguly, 2021)	40
15-3	عملية الالتفاف في الشبكات العصبونية (Manganiello, 2021)	40
16-3	التجميع حسب العدد الأعظمي/ المتوسط (Yani, 2019)	41
17-3	شبكة عصبونية لكشف الأحرف (Saha, 2023)	41
18-3	خوارزميات الكشف بمرحلة واحدة وخوارزميات الكشف بمرحلتين (Kang.2025)	42
19-3	تقسيم الصورة إلى شبكة من الصور الجزئية (Kotai, 2022)	43
20-3	معاملات المتجه Y (Kotai, 2022)	44
21-3	اختيار الشبكة ذات النسبة الأعلى لـ IOU (Kotai, 2022)	45
22-3	الناتج النهائي لخوارزمية YOLO (Kotai, 2022)	45
23-3	مخرجات MOT، رقم يحدد كائنًا معينًا في الفيديو (Bhuiyan, 2023)	46
24-3	مخطط مسلكي مبسط لخوارزمية ملاحقة الكائنات المتعددة (الباحث)	47

25-3	الكشف وتعريف الكائنات (الباحث)	47
26-3	الكشف في الإطار اللاحق (الباحث)	48
27-3	البحث عن ترابط الكائنات (الباحث)	48
28-3	تحديد الكائنات المترابطة (الباحث)	49
29-3	التعريف في الإطار اللاحق (الباحث)	49
30-3	تحديد كائن جديد في الملاحقة (الباحث)	50
31-3	تمثيل لخوارزمية TBD (Guo, 2022)	51
32-3	تمثيل لخوارزمية JDT (Guo, 2022)	51
33-3	بنية خوارزمية المحول (Devi, 2023)	52
34-3	تقدير الحالة الجديدة للكائن أثناء عملية الملاحقة (Sadli, 2023)	53
35-3	تقدير الحالة (Becker, 2023)	55
36-3	خوارزمية التقدير (Becker, 2023)	56
37-3	مخطط القيم الحقيقية والقيم المقدرة من قبل مرشح كالمن (Becker, 2023)	56
38-3	تقدير الموقع بتغيير السرعة (Becker, 2023)	57
39-3	تقدير المسافة للمرشح α−β−γ (Becker, 2023)	57
40-3	التوزيع الغاوصي (Becker, 2023)	58
41-3	مثال لعشرة قياسات (Becker, 2023)	59
42-3	تقدير الحالة (Becker, 2023)	60
43-3	المخطط الصندوقي لمرشح كالمان (Becker, 2023)	62
44-3	الكثافة الاحتمالية لكل من القيم المقاسة والقيم الحالية المخمنة والقيم المقدرة سابقاً (Becker, 2023)	62
45-3	بنية خوارزمية DeepSORT (Borah, 2022)	64
46-3	معيار التقاطع فوق الاتحاد (karen, 2019)	65
47-3	الدقة المحلية (Luiten, 2021)	66
48-3	مطابقة المسارات (Encord2, 2024)	67
1-4	مراحل التداخل في الطريقة المطورة (الباحث)	69

2-4	الاحتجاب (Occlusion) (Mk Bashar, 2022)	71
3-4	صورتان لمشهد واحد من كاميرتين مرئية وحرارية (Flir, 2023)	75
4-4	هيستوغرام الصورة المرئية (الباحث)	76
5-4	هيستوغرام الصورة الحرارية (الباحث)	76
6-4	قلب مخطط الهيستوغرام للصورة الحرارية السابقة (الباحث)	77
7-4	ضغط مخطط الهيستوغرام لجهة اليسار (الباحث)	77
8-4	مفدار ضغط مخطط الهيستوغرام لجهة اليسار (الباحث)	78
9-4	البحث عن الربط (Sadli,2023)	79
10-4	تقدير الحالة الجديدة للكائن أثناء عملية الملاحقة (Sadli, 2023)	80
11-4	مثال لمعطيات إطارين (Sadli, 2023)	81
12-4	مصفوفة التكلفة عند الربط (Sadli, 2023)	81
13-4	تصرف نظام ملاحقة عند مراقبة متحركين متداخلين (الباحث)	82
14-4	التداخل الناتج عن تقاطع متحركين مع نموذج حركة (الباحث)	82
15-4	فشل المرشح بسبب تغيير اتجاه الحركة (Guo,2022)	83
16-4	مراقبة أكثر من نقطة لكائن متغير الشكل (الباحث)	83
17-4	التعرف على الوجه. (Olszewska, 2016)	84
18-4	مراحل التعرف على الوجه. (Olszewska, 2016)	84
19-4	الرقم التعريفي للكائن (Serpush, 2020)	85
20-4	صورة حرارية لمجموعة أشخاص (الباحث)	86
21-4	كشف الكائنات البشرية بـ YOLO-V8 (الباحث)	87
22-4	تابع العتبة Threshold (الباحث)	87
23-4	تجزئة الكائنات ضمن الأقنعة بتابع العتبة (الباحث)	87
24-4	عمليات مورفولوجية لإزالة الأجزاء الصغيرة (الباحث)	88
25-4	تحديد مراكز الأجزاء (الباحث)	88
26-4	مواقع (الاحداثيات النسبية) لأجزاء الهدفين الأيمن والأوسط (الباحث)	89

27-4	تعريف هوية كل هدف (الباحث)	89
28-4	مخطط مسلكي لتعامل الخوارزمية قبل التداخل (الباحث)	90
29-4	ناتج الكاشف في الطريقة المطورة (الباحث)	91
30-4	تابع معامل الثقة في الطريقة المطورة (الباحث)	92
31-4	تابع التفعيل ReLU (Machine learning, 2024)	92
32-4	توضيح لتصرف الخوارزمية في مرحلة ما بعد التداخل (الباحث)	93
33-4	مطابقة صورتين (Elliott, 2022)	94
34-4	اختيار معاملات أعلى ترابط (Elliott, 2022)	94
35-4	تمثيل ثلاثي الابعاد لقيم تابع الترابط في مختلف نقاط الصورة (Wieneke,2017)	95
36-4	احتمالية وجود الهدفين Target1 وTarget2 (الباحث)	96
37-4	عمل الخوارزمية لاعتماد النتائج بعد التداخل (الباحث)	96
1-5	نتائج YOLO-V8 على صورة مرئية (الباحث)	104
2-5	نتائج YOLO-V8 على الصورة المماثلة الحرارية (الباحث)	104
3-5	نتائج YOLO-V8 وفق الطريقة المحسنة (الباحث)	105
4-5	صورة ليلية لشخص (Flir_ADAS_V2)	105
5-5	ناتج YOLO-V8 (الباحث)	106
6-5	تابع العتبة المستخدم (الباحث)	106
7-5	مشهد حراري (Flir_ADAS_V2)	107
8-5	تجزئة المشهد (الباحث)	107
9-5	تحديد مستطيل مناطق الشدة المرتفعة (الباحث)	107
10-5	البحث ضمن مستطيل مناطق الشدة المرتفعة (الباحث)	108
11-5	ناتج YOLO-V8 بعد اجراء التجزئة بتابع العتبة (الباحث)	108
12-5	المخطط المسلكي لخوارزمية تعديل التصحيحات (الباحث)	110
13-5	التحسين بتعديل الإضاءة والتباين (الباحث)	111
14-5	علاقة معامل الثقة بإضاءة الصورة (الباحث)	112

15-5 التأخيرات الزمنية (الباحث) 113
16-5 مركز الهدف بالاعتماد على ناتج خوارزمية الكشف (الباحث) 115
17-5 مركز الهدف بالاعتماد على الطريقة المحسنة (الباحث) 115
18-5 اهمال الأجزاء البعيدة عن المركز أفقياً (الباحث) 115
19-5 فشل المرشح بسبب تغيير اتجاه الحركة (Guo,2022) 116
20-5 الخوارزمية المحسنة بتصحيح المراكز قبل التداخل (الباحث) 117
21-5 الخوارزمية المحسنة بتصحيح المراكز بعد التداخل (الباحث) 117
22-5 اختبار الخوارزمية في المرحلة الأولى على مشهد معقد (الباحث) 119
23-5 اختبار الخوارزمية في المرحلة الثانية على مشهد معقد (الباحث) 119
24-5 اختبار الخوارزمية في المرحلة الثالثة على مشهد معقد (الباحث) 120
25-5 اختبار الخوارزمية في المرحلة الرابعة على مشهد معقد (الباحث) 120
26-5 نتائج اختبار الخوارزمية DeepSORT باستخدام YOLO (الباحث) 121
27-5 نتائج الملاحقة باعتماد الخوارزمية المطورة. (الباحث) 122
28-5 نتائج اختبار مؤشر استعادة الهوية DF1 (الباحث) 124
29-5 نتائج اختبار مؤشر الدقة MOTA (الباحث) 124
30-5 نتائج اختبار مؤشر زمن التنفيذ (الباحث) 125
31-5 نتائج اختبار مؤشر السرعة (الباحث) 125

فهرس الجداول

الرقم	الوصف	الصفحة
1-3	معادلات مرشح كالمن (Becker, 2023)	61
1-5	ثوابت التحسين (الباحث)	111
2-5	نتائج اختبار الطريقة المطورة (الباحث)	112
3-5	اختبارات أولية خوارزمية الملاحقة المطورة (الباحث)	114
4-5	اختبارات أولية خوارزمية الملاحقة المطورة (الباحث)	118
5-5	نتائج اختبار مؤشر استعادة الهوية IDF1 (الباحث)	123
6-5	نتائج اختبار مؤشر الدقة لملاحقة الكائنات المتعددة MOTA (الباحث)	123
7-5	نتائج اختبار دورة المعالجة T (الباحث)	123
8-5	نتائج اختبار معدل الإطارات FPS (الباحث)	123

قائمة المصطلحات

المقابل العربي	الاختصار	المصطلح الانكليزي
الضبط		Accuracy
تابع التنشيط		Activation function
أنظمة مساعدة السائق المتقدمة	ADAS	Advanced Driver Assistance Systems
خوارزمية		Algorithm
المتوسط	Avg	Average
التجميع حسب المتوسط		Average pooling
متوسط دقة الكشف	AP	Average Precision
النقل الخلفي عبر الشبكة		Back propagation
المستطيل المحيط بالكائن		Bounding Box
الإضاءة		Brightness
القوة المسيطرة		Brute-force
الصنف		Class
معامل ثقة الكشف	Conf	Confidence
التباين		Contrast
التلاف	CONV	Convolution
الشبكات العصبونية التلافيفية	CNN	Convolutional Neural network
اكتشاف الزوايا		Corner detection
الترابط		Correlation
نهج التعلم القياسي جيب التمام		Cosine metric learning approach
تابع الترابط المتصالب		Cross correlation
التعلم العميق	DL	Deep Learning
التمدد		Dilation
اكتشاف الحواف		Edge detection
الطيف الكهرومغناطيسي		Electromagnetic spectrum
الحت		Erosion
مجموعة بيانات للوجه		Face dataset
كشف الوجه		Face detection
التعرف على الوجه		Face recognition
سلبي كاذب	FN	False Negative
إيجابي كاذب	FP	False Positive

Far-infrared		الأشعة تحت الحمراء البعيدة
Fast corner detector		خوارزمية كاشف الزوايا السريع
Fast R-CNN		خوارزمية CNN السريعة ذات الذاكرة
Feature classification		الخصائص المشتركة
Frame		اطار الفيديو
Full Connected layer	FC	الطبقة المتصلة بالكامل
Gaussian function		تابع التوزيه الطبيعي غاوص
Hestogram Equalization		تسوية الهيستوغرام
Hidden layers		الطبقات المخفية
Hierarchical Spatial-aware Siamese convolutional neural network	HSSNet	شبكة عصبونبة تلافيفية مضاعفة هرمية واعية بالمكان
High Pass Filter	HPF	مرشح التمرير العالي
Image histogram		التوزيع اللوني للصورة (هيستوغرام)
Hungarian algorithm		الخوارزمية المجرية لتطابق المسارات
Identification	ID	معرف الهوية
Identification Precision	IDP	دقة التعريف
Identification Recall	IDR	تذكر التعريف
Identification Switch	IDSW	مؤشر تبديل الهوية
Identification F1-score	IDF1	درجة الحفاظ على الهوية
Image background		خلفية الصورة
Image segmentation		تجزئة الصورة
Infra red		تحت الحمراء
Infra-Red Tracking	TIR	الملاحقة بالأشعة تحت الحمراء
Input layer		طبقة الادخال
Intelligence Artificial	AI	الذكاء الاصطناعي
Intersection Over Unions	IOU	نسبة التقاطع على الاجتماع
Inverse image		معكوس الصورة
Joint detection and tracking	JDT	نهج الكشف والملاحقة المشترك
Kalman Filter	KF	مرشح كالمان
Kernel		النواة
Kinetic features		سمات حركية
Kirsch		خوارزمية كيرش لكشف الحواف
Learning Machine	ML	تعلم الآلة
Linear spatial filter		المرشح المكاني الخطي
Long Short-Term Memory	LSTM	الذاكرة طويلة المدى

Low Pass Filter	LPF	مرشح التمرير المنخفض
Max disappear		معامل المسافة الأعظمية للاختفاء
Max pooling		التجميع حسب الرقم الأعلى
Measurement uncertainty		عدم اليقين في القياس
Mid-wavelength-infrared		الأشعة تحت الحمراء متوسطة الطول الموجي
Moravec		خوارزمية مورافيك لكشف الزوايا
Morphological processes		العمليات على الشكل
Multiple Object Tracking	MOT	ملاحثة الكائنات المتعددة
Multiple Object Tracking accuracy	MOTA	مقياس أداء الملاحقة
Multiple spatial scales		الأبعاد المتنوعة للكائنات
Natural Language Processing	NLP	معالجة اللغة الطبيعية
Near-infrared		الأشعة تحت الحمراء القريبة
Neural network		الشبكات العصبونية
Object detection		كشف الكائنات
Objects		كائنات
Occlusion		الاحتجاب
Opening & Closing		الفتح والاغلاق
Output layer		طبقة الاخراج
Overlapped		متداخل
Perceptron		أول نموذج لشبكة عصبونية اصطناعية
Person		صف الأشخاص
POOL		طبقة التجميع
Precision		الدقة
Prediction equation		معادلة التنبؤ
Probability Density Function	PDF	تابع الكثافة الاحتمالي
Real time		الزمن الحقيقي
Recall		التذكر
Recognition		تمييز
Red Green Blue System	RGB	نظام الألوان (الأحمر الأخضر الأزرق)
Resolution	Res	التمييزية
Robustness		المتانة
Simple Online Realtime Tracking	SORT	خوارزمية ملاحقة
Sobol		خوارزمية سوبل لكشف الحواف
Standard deviation		الانحراف المعياري
State extrapolation equation		معادلات استقراء الحالة

State update equation		معادلات تحديث الحالة للمرشح
Target		الهدف
Thermal		حراري
Thermal spectrum		الطيف الحراري
Threshold		تابع العتبة
Track		مسار
Tracking		ملاحقة
Tracking By etection	TBD	طريقة الملاحقة بالكشف
Tracking speed		سرعة الملاحقة
Transformers		الملاحقة المعتمدة على المحولات
Transition equation		معادلة الانتقال
True Nigative	TN	إيجابي حقيقي
True Positive	TP	إيجابي حقيقي
Ultra violet		فوق البنفسجي
Variance		التباين
Visible spectrum		الطيف المرئي
You only look once	YOLO	خوارزمية الكشف (أنت تنظر لمرة واحدة)

الملخص

يهدف هذا البحث إلى التركيز على إيجاد حلول لبعض مشاكل الملاحقة بالصورة الحرارية وخصوصاً مشاكل التداخل، حيث يواجه تصنيف الكائنات وملاحقتها في الرؤية الحرارية صعوبات كبيرة، وذلك بسبب فقدانها للألوان والمعالم الداخلية وتشوه أجزائها، لذلك يتم اقتراح منهجية لملاحقة الأشخاص المتواجدين في مشهد ملتقط من كاميرا حرارية والتمييز بينهم في ظروف تداخلهم ضمن المشهد، هذه المنهجية تعتمد على مجموعة من الحلول التي تدمج بين الطرائق الذكية والطرائق التقليدية، وتشمل في تفاصيلها تعزيز السمات بالصورة الحرارية وتقديم طريقة مناسبة لتصحيح مركز الكائن الحراري وخصوصاً متغير الشكل، بالإضافة لاستخدام النماذج الحركية بفعالية أكبر، وذلك لتقدير الموقف لكل كائن أثناء مرحلة التداخل، وإعادة ربط الكائنات وتمييزها بعد مرحلة التداخل.

تم اعتماد الطريقة التجريبية والتحليلية والوصفية لفهم المواقف المختلفة في الملاحقة، ومن ثم بناء الخوارزمية المقترحة في كافة مراحلها، وتم استخدام أحدث اصدار من الكاشف السريع YOLO للكشف عن الكائنات، ومن ثم تم تجزئة هذه الكائنات إلى مجموعات، واستخدام مرشح كالمن كنموذج حركة لكافة النقاط والمجموعات المراقبة، بالإضافة لذلك استُخدمت خوارزمية الملاحقة DeepSORT كخوارزمية لتجريب الطريقة المطروحة، والمقارنة والحصول على النتائج بتمايز الكائنات ضمن المشهد الواحد، حيث تم خفض الخطأ المرتكب عند إعادة تعريف الكائنات بعد انتهاء التداخل بنسبة جيدة تبلغ 6.5% في المشاهد المختبرة، مع بقاء عملية الملاحقة ضمن الزمن الحقيقي. تم الاعتماد في تجاربنا على خوارزمية الملاحقة DeepSORT لكن الطريقة المقترحة مستقلة وغير مرتبطة بأي خوارزمية، لذلك تعتبر حلا منهجياً جيداً، عوضت عن انخفاض السمات في الكائن الحراري، ويمكن اعتمادها في العديد من خوارزميات الملاحقة.

الكلمات المفتاحية: الكائن الحراري، معالجة الصورة الحرارية، كاشف الكائنات، السمات، معرف الكائن، مركز الكائن، خوارزميات الملاحقة، نموذج الحركة، تداخل الكائنات، تصنيف الكائنات.

Abstract

This research aims to focus on finding solutions to some of the problems of thermal imaging tracking, particularly overlap problems. Object classification and tracking in thermal vision face significant difficulties due to the loss of color and internal features, as well as the distortion of their parts. Therefore, a methodology is proposed for tracking and distinguishing between persons present in a scene captured by a thermal camera when they overlap within the scene. This methodology relies on a set of solutions that combine intelligent and traditional methods. Its details include enhancing features in thermal imaging and providing a suitable method for correcting the position of the thermal object, especially one with a variable shape. It also uses more effective kinematic models to estimate the position of each object during the interference phase, and to reconnect and distinguish objects after the interference phase.

We adopted an experimental, analytical, and descriptive approach to understand the different tracking situations, and then developed the proposed algorithm in all its stages. The latest version of the YOLO fast detector was used to detect objects. These objects were then segmented into groups, and a Kalman filter was used as a motion model for all monitored points and groups. In addition, the DeepSORT tracking algorithm was used to test the proposed method, compare, and obtain results by distinguishing objects within a single scene. The error committed when re-identifying objects after overlapping was reduced by a significant 6.5% in the tested scenes, while the tracking process remained in real time. Our experiments relied on the DeepSORT tracking algorithm, but the proposed method is independent and not tied to any algorithm. Therefore, it is considered a good methodological solution, compensating for the reduced features in the thermal object and can be adopted in many tracking algorithms.

Keywords: Thermal object, thermal image processing, object detector, features, object identifier, object center, tracking algorithms, motion model, object overlap, object classification.

الفصل الأول

الإطار العام للبحث

Research General Framework

1-1 مقدمة

أدت طرائق اكتشاف الكائنات القائمة على التعلم العميق إلى تقدم كبير في خوارزميات ملاحقة الكائنات المتعددة Multi Object Tracking (MOT). تركز طرائق MOT الحالية بشكل أساسي على مشاهد المشاة أو المركبات، ولكن ملاحقة الكائنات لازالت تعتبر أحد المشكلات الأساسية والعملية في مجالات أبحاث الرؤية الحاسوبية. تُظهر العديد من المشاريع التطبيقية أن ملاحقة الكائنات لا يزال يواجه العديد من المشكلات، مثل الاختلافات الضعيفة في المظهر بين الهدف والخلفية، والخلفية المعقدة، والتداخلات (Hu, 2024).

على الرغم من نجاح الطرائق نسبياً في مقاطع الفيديو بالنظام المرئي Red Green Blue (RGB)، إلا أنه بالنسبة للمشهد الحراري فالأمر معقد (Wu,2022)، فمهمة ملاحقة الهدف بالأشعة تحت الحمراء الحرارية لا تتأثر بتغيرات الإضاءة ويمكن تتبعها في الليل، وفي الأيام الممطرة، والأيام الضبابية، وغيرها من الظروف الجوية القاسية، لكن لا تزال هذه المهمة تواجه العديد من التحديات، مثل الحجب والتشوه وتداخل التشابه، بسبب أن الهدف يكون عادةً باهتًا وعديم المعالم. من الصعب حتى على البشر تحديد موقع الهدف فقط من خلال النظر إلى إطار ثابت واحد من فيديو الأشعة تحت الحمراء (Yuan, 2023)، لذلك لا تستطيع الطرائق الحالية القائمة على اكتشاف المنطقة المحلية التقاط الهدف بالأشعة تحت الحمراء بشكل فعال.

تعتمد أنظمة الملاحقة على طيف واسع من الطرائق، فمنها من لايزال يتبع الأساليب التقليدية بإرجاع الملاحقة لمسألة ملاحقة نقطة مادية، أو ملاحقة نموذج شكل ما، إلى ما هنالك. تعتمد الكثير من خوارزميات الملاحقة حديثاً على كاشف الكائنات، بينما تعتمد خوارزميات أخرى على الحركة، وهناك من يمزج الحلول، وتتنافس الخوارزميات على تذليل الصعاب المتمثلة بالتعرف على الشكل وتمييزه وملاحقته حتى أثناء التداخل.

2-1 الدراسات المرجعية

أحرزت الأبحاث الحديثة في كشف الكائنات في المجال المرئي تقدماً كبيرة، وفي مجال الأشعة تحت الحمراء لازالت المشاكل تؤخر الوصول إلى حلول منافسة للمرئية، وقد ركز الكثير منها على اتجاه واحد من الطرائق إما على الاكتشاف أو الملاحقة. نبين في الدراستين التاليتين تركيز الأولى على الكشف، بينما ركزت الأخرى على الملاحقة.

في البحث (Liu, 2017) يتم الاعتماد على الشبكات العصبونية التلافيفية CNN المدربة مسبقًا والقائمة على الصور المرئية للملاحقة في مجال الأشعة تحت الحمراء الحرارية، وذلك بسبب القدرة التمثيلية القوية للشبكات العصبونية التلافيفية وتطبيقها الناجح في الملاحقة البصرية، وهو ينطلق من ملاحظة أن السمات في الطبقة الأخيرة المسمات "متصلة بالكامل" ليست مناسبة لتتبع الأشعة تحت الحمراء الحرارية بسبب نقص المعلومات المكانية للهدف، في حين أن السمات في طبقات التلاف مناسبة. إلى جانب ذلك، فإن السمات في طبقة التلاف الواحدة ليست

قوية في مواجهة التحديات المختلفة. بناءً على هذه الملاحظة، أُقترح ملاحقاً قائماً على مرشح الارتباط مع سمات تلافيفية متعددة الطبقات للملاحقة ضمن مجال الأشعة تحت الحمراء الحرارية.

تنقسم الخوارزمية إلى مرحلتين: التدريب والبحث. في مرحلة التدريب، الهدف هو الحصول على قوالب المرشح على الإطار المتوقع (الإطار t). ثم استخدام شبكة مدربة مسبقًا لاستخراج السمات التلافيفية للهدف الحراري بالأشعة تحت الحمراء على طبقات تلاف متعددة. بعد ذلك، يتم الحصول على المرشح المقابل باستخدام هذه السمات التلافيفية وعلامة الشكل الغاوسي. في مرحلة البحث، الهدف هو الحصول على نتيجة الملاحقة المثلى ضمن إطار البحث (الإطار (t + 1)). كما هو الحال في مرحلة التدريب، يتم الحصول على خريطة السمات للهدف على كل طبقة تلاف. بعد ذلك، يتم جمع خرائط السمات وقوالب المرشح لحساب الارتباط المتبادل.

وأخيراً، يقترح طريقة تجميعية تعتمد على تباعد كولباك-ليبلر "الانتروبية النسبية: مقياس لمدى اختلاف توزيع واخد احتمالي عن توزيع احتمالي مرجعي آخر"، وذلك لدمج خرائط الاستجابة هذه وإرجاع خريطة أقوى.

لتقييم أداء الخوارزمية المقترحة، أجريت تجارب على معيارين للملاحقة بالأشعة تحت الحمراء الحرارية: VOT-TIR 2015 وVOT-TIR 2016. تُظهِر النتائج التجريبية أن الخوارزمية حققت نتائج معقولة بالمتانة Robustness والدقة بلغت 0.56.

في الدراسة السابقة عولجت مسألة الملاحقة كمسألة تمييزية حيث معظم طرائق الملاحقة باستخدام الأشعة تحت الحمراء الحرارية Infra-Red Tracking (TIR) تمييزية تعالج مشكلة الملاحقة كمهمة تصنيف (كشف).

ينطلق البحث (Li, 2018) من أن هدف المصنف (التنبؤ بالصنف) غير مقترن بهدف الملاحق (تقدير الموقع). تركز مهمة التصنيف على الاختلاف بين الفئات للكائنات، بينما تتعامل مهمة الملاحقة بشكل أساسي مع الاختلاف داخل الفئة لنفس الكائنات. لذلك اعتبرت الدراسة مشكلة ملاحقة الأشعة تحت الحمراء الحرارية TIR مهمة للتحقق من التشابه، والتي ترتبط بشكل جيد بهدف مهمة الملاحقة. اقترحت الدراسة ملاحق يعمل بالأشعة تحت الحمراء الحرارية TIR عبر شبكة عصبونية تلافيفية Convolutional Neural Network (CNN) هرمية واعية بالمكان، تسمى HSSNet للحصول على كل من السمات المكانية والدلالية لكائن الأشعة تحت الحمراء الحرارية TIR، ثم تم تصميم شبكة عصبونية تلافيفية تدمج الطبقات التلافيفية الهرمية المتعددة. بعد ذلك، تم اقتراح شبكة واعية بالمكان لتعزيز القدرة التمييزية للميزة الهرمية المندمجة. بعد ذلك، يتم تدريب هذه الشبكة من البداية إلى النهاية على مجموعة بيانات كبيرة لكشف الفيديو المرئي لمعرفة التشابه بين الكائنات المقترنة قبل نقل الشبكة إلى مجال TIR. بعد ذلك، تُستخدم الشبكة السيامية المدربة مسبقًا هذه لتقييم التشابه بين قالب الهدف والمرشحين المستهدفين. أخيرًا، يتم تحديد المرشح الأكثر تشابهًا مع الهدف المتعقب. تُظهر النتائج التجريبية المكثفة على معايير VOT-TIR 2015 وVOT-TIR 2016 أن هذه الطريقة المقترحة تحقق أداءً بالدقة يقارب 0.63.

إن اعتماد الكثير من نظم الملاحقة على الكشف قد قاد بالضرورة لتحسين نماذج وخوارزميات الكشف، وإذا أخذنا بعين الاعتبار بعض التطورات في مجال الصورة الحرارية فنلاحظ بأن الدراسة في (Sun, M.,2021) قد أسست نموذجًا متطوراً لكشف الأهداف، I-YOLO، لصور الطرائق السريعة بالأشعة تحت الحمراء استناداً إلى شبكة You Only Look Once (YOLO v3). وقد اعتمد في هذا النموذج على DRUNet وDRUNet لتحسين جودة الصورة بسبب انخفاض نسبة الإشارة إلى الضجيج في الصور الحرارية، وتم استخدام EfficientNet لتحسين بنية العمود الفقري وتعزيز القدرة على إعادة استخدام السمات والتعرف على الأهداف الحرارية لتحسين أداء الكشف، كذلك تبنى أيضًا خوارزمية k-means وتقنية تحسين البيانات استنادًا إلى مجموعة البيانات الفعلية.

يقوم البحث على دراسة تأثير المكونات المختلفة المكررة على أداء النموذج: K-means وDRUNet وEfficientNet وتم إجراء جميع التجارب على مجموعة بيانات FLIR. يتم تنفيذ خوارزمية K-means على نطاق واسع لتغيير حجم المربعات المحددة. وصل مؤشر دقة الكشف Mean Average precision (mAP) إلى 0.55 مما يدل على زيادة بنسبة 0.03 بالنسبة لخط الأساس. يتم استخدام DRUNet لتقليل الضجيج في صورة الأشعة تحت الحمراء وتحسين تأثير تدريب النموذج. تحسن mAP لجميع الفئات بشكل كبير ووصل mAP إلى 0.78، وهو تحسن بنسبة 23%. ومن خلال اختبار الإصدارات المختلفة من EfficientNet، وجدت الدراسة أن EfficientNet-b2 يُظهر أفضل أداء شامل باعتباره العمود الفقري للنموذج. كل ذلك قد ساهم في وصول mAP النهائي إلى 0.89. وبمقارنة نموذج I-YOLO بنماذج الكشف الحديثة الأخرى. تحسنت mAP بنسبة 12.95% و4.35% بناءً على مجموعتي بيانات، مقارنةً بنموذج YOLOv4.

إن الدراسة (Sanyam, 2022) تقدم شرحاً عن Simple Online Realtime Tracking (SORT) الذي يعتبر من أهم منهجيات الملاحقة، حيث تُستخدم مرشحات كالمان والخوارزميات المجرية لتتبع الكائنات، فهو يتكون من أربعة مكونات رئيسية،

- الكشف: ويستخدم كواشف مثل FrRCNN وYOLO وغيرها.
- التقدير: باستخدام نموذج سرعة ثابتة بمرشح كالمان.
- ارتباط البيانات: تُحسب مصفوفة التكلفة عن طريق نسبة التقاطع على الاتحاد IOU ويُحل التعيين على النحو الأمثل باستخدام الخوارزمية المجرية.
- إنشاء وحذف هويات المسار: هذه الوحدة مسؤولة عن إنشاء وحذف المعرفات. يتم إنشاء هويات فريدة وتدميرها وفقًا لـ IOUmin.

يُظهر SORT أداءً ممتازًا من حيث دقة الملاحقة. لكنه يُرجع مسارات تحتوي على عدد كبير من مُبدِّلات المُعرِّفات، ويفشل في حالة الانسداد، ويرجع ذلك إلى مصفوفة الارتباط المُستخدمة. يستخدم DeepSORT مقياس ارتباط أفضل يجمع بين مُوصِفات الحركة والمظهر. بالتالي فإن النموذج المطور DeepSORT هو خوارزمية ملاحقة

لا تعتمد فقط على سرعة وحركة الكائن، بل أيضًا على مظهره، وذلك بالإعتماد على استخدام نهج التعلم المتري لجيب التمام حيث جيب التمام مقياس يساعد النموذج على استعادة هوياته في حالة الانسداد طويل الأمد وفشل تقدير الحركة أيضًا. استخدام هذه الأشياء البسيطة يمكن أن يجعل أداة التتبع أكثر قوة ودقة.

تم استخدام مجموعة فرعية من مجموعة بيانات MOT17، وأظهرت تقنية DeepSORT أداءً بصريًا جيدًا نسبيًا، حيث بلغ متوسط الدقة Multi-Object Tracking Accuracy (MOTA) التي تم الحصول عليها 28.6%، و نسبة الحقاظ على الهوية Identification F1-score (IDF1) 43%.

إن البحث (Gao, X, 2023) هو من الأبحاث القليلة التي ركزت على مشكلة الفشل أثناء التداخل باستخدام خوارزمية DetTrack حيث ركزت على تحسين اكتشاف الكائنات المحجوبة. في البداية يتم استخدام التنبؤ بمسار الحركة لتحسين الحد الأعلى لاكتشاف الكائنات المحجوبة. بعد ذلك، يتم استخدام معلومات الميزة المكانية الزمنية بين الكائن والبيئة المحيطة لتتبع الكائنات المتعددة. أخيرًا، يتم استخدام إطار الفرضية ملاحقة الكائن المحجوب تمامًا.

إن معاملات الضبط تكون كما يلي:

- عتبة اكتشاف النتيجة العالية هي 0.5 أي 50%
- عتبة اكتشاف النتيجة المنخفضة هي 0.3
- الاحتفاظ بمعلومات المسار بالزمن يكافئ 50 Frame، عندما لا يتم مطابقة المسار

وفي حالة تأكيد الاكتشاف، سيتم تجاهل المسار. في التدريب التجريبي، يتم استخدام YOLOX كشبكة أساسية، ويتم استخدام معلمات النموذج المدربة مسبقًا على مجموعة بيانات COCO لتهيئة النموذج. أجريت تجارب مقارنة باستخدام خوارزميات الملاحقة المتقدمة على مجموعات بيانات MOT16 وMOT17 وMOT20. تم استخدام مؤشرات تقييم معيارية لمجموعة البيانات لتقييم جوانب مختلفة من أداء الملاحقة. المؤشرات هي الدقة ومدى الحفاظ على هوية الكائن وتم اجراء تجارب عديدة لاختيار عتبات الكشف العليا والدنيا المناسبة، وقد كان الأداء مثاليًا عندما تم تحديد عتبة أداء الملاحقة المثالي 0.5 لاكتشاف الدرجات العالية و0.3 لاكتشاف الدرجات المنخفضة. حيث وصلت الدقة في هذه الطريقة إلى 73%، ووصل مدى الحفاظ على الهوية IDF1 (Identification F1-score) إلى 70.9%.

يقدم البحث (Ibrahim,2023) دراسة على أهمية استخدام الكاميرات الحرارية في مجالات مختلفة، بما في ذلك أنظمة المراقبة وأنظمة مساعدة السائق المتقدمة (Advanced driver-assistance systems) ADAS، حيث تعمل بشكل أفضل في الإضاءة المنخفضة من كاميرات الضوء المرئي. لكن التحديات في مجال نظام المراقبة أو ADAS المتعلقة بالكاميرات الحرارية هي الانسداد والتقاطع الحراري بين الأشياء ذات المظهر المتشابه أثناء مهام اكتشاف الأشياء أو ملاحقتها، مما قد يؤدي إلى اكتشاف خاطئ وإيجابيات خاطئة وفقدان

الملاحقة. تم تقديم تحليل أداء لخوارزمية You Only Look Once (YOLO) جنبًا إلى جنب مع الملاحقة بالتعلم العميق في الزمن الحقيقي عبر الإنترنت (DeepSORT) ومدى تحسين الملاحقة الحرارية متعددة الكائنات عبر الإنترنت (MOT) القائمة على الفيديو الحراري في مشهد الانسداد والتقاطع الحراري. يتم استخدام YOLO، كواحدة من أحدث الطرائق لمهمة الكشف. بعد ذلك، يتم تعقب الكائن المكتشف من YOLO باستخدام DeepSORT. توضح النتائج أن MOT عبر الإنترنت للصور الحرارية المتسلسلة باستخدام YOLO-DeepSORT حقق درجة MOTA بنسبة 44.2% وIDF1 بنسبة 45.3%. وبعد أن تمت إضافة مثال سلبي في عملية تدريب YOLO لتقليل الاكتشاف الخاطئ، وقد أدى ذلك إلى تحسن مع درجة MOTA إلى نسبة 63.8%، ودرجة IDF1 إلى نسبة 54.6%.

3-1 مشكلة البحث

تعتمد الكثير من خوارزميات ملاحقة الكائنات المتعددة Multi-Object Tracking (MOT) على كشف الكائنات بإطار الصورة الحالي، ثم يتم ربطها مع الكائنات المكتشفة في الإطار السابق لتحقيق تأثير الملاحقة، بحيث تستطيع الخوارزمية تحديد أن الكائن المكتشف في الإطار الحالي هو ذاته المكتشف في الإطار السابق. وتستخدم الكثير من الخوارزميات لذلك نماذج الحركة وطرائق التوقع Prediction لمتابعة وتقدير حركة الكائن من إطار لآخر مثل مرشح كالمان Kalman Filter، وهناك خوارزميات تستخدم أنواعاً من خوارزميات الشبكات العصبونية التي تعتمد على الذاكرة طويلة المدى مثل Long Short-Term Memory (LSTM)، بالتالي تعمل على مراقبة المسار وتقدير حركة الكائن ضمن خوارزميات الكشف والمتابعة. وحديثا بعض الخوارزميات تعتمد على آلية الانتباه الموجودة ضمن المحولات Transformer للربط بين الكائنات في الإطارات المتتالية. يلاحظ أن طرائق الملاحقة المعتمدة على الكشف والربط بين الإطارات عبارة عن طرائق جيدة للملاحقة بالزمن الحقيقي، فهي غير معقدة حسابياً وبرمجياً بخلاف الطرائق الأخرى التي تحتاج لعتاد قوي لتقوم بدورها على الشكل الأمثل.

لقد حلت الطرائق السابقة العديد من المشكلات الصعبة التي تعترض الملاحقة للأهداف المتعددة في لقطات الصور الملونة RGB لكن كانت هناك تحديات تتعلق بشكل أساسي في تداخل الأهداف الملاحقة بحيث كانت هذه المشاكل هي العقد الأساسية التي تعترض طرائق الملاحقة. بالإضافة لذلك فإن فيديوهات الصور الحرارية التي تحوي كائنات متعددة تمثل تحدياً آخر من ناحية الكشف والملاحقة أومن ناحية التداخل.

يمكن للتصوير الحراري بالأشعة تحت الحمراء اكتشاف درجة حرارة الجسم البشري بغض النظر عن ظروف الإضاءة، وبالتالي يمكن استخدام الكاميرا الحرارية العاملة بالأشعة تحت الحمراء للتعرف على النشاط البشري للمراقبة الذكية (Yeom, 2021). إن الصور الحرارية لها عيوب واضحة، مثل الدقة المنخفضة، وانخفاض نسبة الإشارة إلى الضجيج، ونقص معلومات الملمس واللون والحواف غير الواضحة. وبالرغم من حلول الذكاء الاصطناعي محل الكثير من خوارزميات كشف الكائنات وتصنيفها وملاحقتها، إلا أن التعلم العميق لم يحقق نجاحات كبيرة في الصور الحرارية للأسباب الأنفة بالإضافة إلى أن التعلم العميق بحاجة للعديد من مجموعات البيانات

للتدريب والتحقق، فالتطوير في مجموعات البيانات في مجال الصور الحرارية لم يجاري المجال المرئي بل تخلف عنه كثيراً بسبب صعوبات بناء مجموعات البيانات، وهو ما يمثل عقبة حقيقية أمام تطوير كاشف وملاحق كائنات متعددة مخصص للصور الحرارية. ويمكن تلخيص مشكلة البحث بالبنود التالية:

- الكائن الهدف ضمن صورة حرارية: أي مشكلة البصمة الحرارية، حيث البصمة الحرارية غير صلدة ومنخفضة السمات وضعيفة التمييز.
- الكائن الهدف صغير: ويلاحظ هذا عندما يكون الهدف بعيداً، حيث عندما يكون الهدف صغيراً سيعاني أيضاً من ضعف السمات والملامح المهمة التي تستخدم للتعرف عليه.
- الكائنات تتداخل خلال حركتها: وهي المشكلة الأكبر حيث شكل الكائنات المتقارب في الصورة الحرارية يخفض من إمكانية الترابط بين الكائنات خلال الإطارات المتتالية ضمن الفيديو.

4-1 هدف البحث

بناءً على مشكلة البحث تبرز الحاجة لبناء نظام ملاحقة جيد قادر على تمييز الكائنات المتحركة بعد تداخلها ضمن الصورة الحرارية ويكون ذلك بكلفة عتاديه وزمنية صغيرة بحيث تبقى العملية ضمن الزمن الحقيقي.

تعتمد مقاربة الطريقة المقترحة على الدمج بين الطرائق التقليدية مع أساليب الذكاء الاصطناعي، وتركز على:

- استخدام طرائق دقيقة وسريعة للكشف عن الأشخاص في الصورة، ومحاولة تحسين ظروف الكشف بالصورة الحرارية، بالتالي زيادة السمات والخصائص للكائنات الفقيرة بالسمات.
- استخدام الخصائص والسمات الحركية بالإضافة للسمات المحسنة، لضمان ترابط الكائنات المكتشفة ضمن مراحل الملاحقة، وخصوصاً أثناء التداخل.

5-1 أهمية البحث

يعتبر تداخل الكائنات الأهداف في الرؤية الحاسوبية، من أهم المشاكل المعيقة للملاحقة الفعالة في عالم الذكاء الاصطناعي. في الرؤية الليلية تستخدم الكاميرا الحرارية كوسيلة متفردة لمراقبة الكائنات الحية، لكن تواجه الرؤية الحرارية الحاسوبية صعوبة كبيرة في تصنيف الكائنات وملاحقتها، ويعود ذلك إلى فقدانها للألوان والمعالم الداخلية وتشوه أجزائها بالصورة الحرارية مما يحد من القدرة على المعالجة وتمييزها عن بعضها ضمن الصنف الواحد. ومن أهم المعيقات التي تحد من درجة نجاح خوارزمية الملاحقة هي الاحتجاب Occlusion وخصوصاً الناتج عن تداخل هذه الكائنات، بالتالي ضياع الهدف الملاحق أثناء وبعد التداخل، لذلك نقوم هنا بتطوير خوارزمية تمييز للكائنات المتداخلة في الزمن الحقيقي، باعتماد السمات الطيفية الحرارية والحركية للمساهمة بتذليل الصعوبات التي تعترض استخدام الكاميرات الحرارية للملاحقة والتمييز الفعال.

6-1 مسوغات البحث

كثير من الأبحاث تناولت مسألة كفاءة خوارزمية الملاحقة بشكل عام دون معالجة كل مشكلة مؤثرة لوحدها، مما جعل الكثير من الخوارزميات متقاربة بنتائجها، لكن استخدام هذه الخوارزميات على اللقطات الحرارية الليلية سيقابل بالكثير من التحديات إذا لم نقل الفشل. لذلك ارتأينا القيام بدراسة المشكلات التي تعترض الملاحقة وركزنا على تغير شكل الكائن والاحتجاب كونه الأكثر شيوعاً، وحاولنا حل هذه المشاكل بغض النظر عن الخوارزمية المستخدمة وطبقنا ما توصلنا إليه في تطوير منهجية تستخدم أهم الخوارزميات لنكون أمام منهجية محسنة قادرة على حل المشاكل ما أمكن وبذلك يتم التمييز النسبي بين الكائنات المختلفة وذلك بالصورة الحرارية.

7-1 محددات البحث

إن مسألة كفاءة خوارزميات الملاحقة بشكل عام هو عنوان مطاط تندرج تحته العديد من المشاكل المتفرقة، وعند استخدام هذه الخوارزميات على اللقطات الحرارية الليلية سيعني الكثير من التحديات الإضافية. يتعلق ذلك بكثرة الأصناف المراقبة ونوع الخوارزمية وقابلية البث الحراري للصنف المراقب، فالكائنات الحية تبث طاقة حرارية، أما الكائنات الأخرى فهي تكتسب كمية من الحرارة في كثير من الأحيان قريبة من الوسط المحيط بها بالتالي من الصعوبة كشفها. لذلك سيقتصر البحث على الكائنات الحية ونختار صنف مهم منها وهو الإنسان، وستكون الدراسة ضمن مشهد حراري ثابت (الكاميرا غير متحركة)، أما بالنسبة لحركة الأهداف فستكون ضمن سرعات زاوية أقل من 2 درجة في الثانية (يقابل سرعة خطية 1m/s من أجل مسافة 30m)، وقد لعبت زاوية الكاميرا المستخدمة دور في تحديد هذه السرعة وهي سرعات كافية لمراقبة الأشخاص وملاحقتهم في الحالات الشائعة للعديد من التطبيقات، وبالإمكان في حالة تطبيقات تتطلب سرعة زاوية أكبر استخدام كاميرا بفتحة زاوية أكبر وتظل النتائج جيدة.

من الجدير بالذكر أن الخوارزمية المطورة قد ركزت على أشخاص متحركين، ومن الممكن اجراؤها على كائنات متنوعة أخرى سواء تم ذلك بدون تعديل على الخوارزمية، أو بتعديلات بسيطة كما سنرى.

8-1 حدود البحث

الحد المكاني:

جرى تنفيذ البحث في قسم هندسة الالكترونيات والاتصالات في كلية الهندسة الميكانيكية والكهربائية بجامعة دمشق.

الحد التقاني:

تم استخدام حاسوب Asus معالجه انتل i7, 2.6GHz وذاكرة 8 GB من أجل المراحل التجريبية. ومن أجل تجريب الخوارزمية المحسنة تم استخدام منصة ColabResearch للتغلب على سرعة المعالجة بالحواسب،

وهي تؤمن معالجة ضمن المسرع T4GPU وهو كذلك يستعمل Google Drive لتخزين النموذج والبرمجيات بالتالي الاستقلال عن مشاكل استخدام الحواسب الشخصية.

تم استخدام كاميرا حرارية غير مبردة ابعاد حساسها بالبيكسل 384 x 288 P، وزاوية الرؤية 14 درجة.

يلاحظ بأن أبعاد حساس الكاميرا بالإضافة لزاوية الرؤية الصغيرة قد حد من التأكد من إمكانية معالجة السرعات الزاوية الكبيرة نسبياً للهدف.

1-9 منهج البحث وأدواته

من أجل وضع فرضيات وتخمينات المسألة تم الاعتماد على المنهج الوصفي في تحديد المشكلة قيد البحث، وجمع البيانات والمعلومات عنها، كذلك في تحديد مدى نجاح بعض نتائج المعالجة. ولتحليل المشكلة ودراسة الأجزاء واستخراج الحلول تم اعتماد المنهج التحليلي. بالإضافة لذلك تم اعتماد المنهج التجريبي المعتمد على البرمجة، بالتوافق مع المناهج الأخرى سواءً خلال مرحلة توصيف المشكلة واستنباط الحلول أو خلال استعراض النتائج ودراسة مدى تحقيقها للأهداف. أما بالنسبة لأدوات البحث فقد استخدم بالإضافة لبعض مجموعات البيانات، كاميرا حرارية مع حاسوب سريع لتحصيل بعض البيانات الإضافية، والتي حُصلت في جامعة دمشق بتجارب ميدانية بمساعدة طلبة متطوعين، واستخدمت في مختلف التجارب. كذلك تم استخدام مجموعة من البرمجيات بإصداراتها الجديدة، حيث غطى الماتلاب ما تطلب من المحاكاة، وتم الاعتماد على برمجيات بايثون لإجراء التجارب بالإضافة للمحاكاة، حيث يمتلك البايثون سجلا حافلا سواء بالبرمجيات الملائمة أو المكتبات الوفيرة المتوافقة معه في مجالات الذكاء الاصطناعي.

1-10 بنية الاطروحة

تتكون الأطروحة من خمسة فصول رتبت مواضيعها على الشكل التالي:

- يتناول الفصل الأول الإطار العام للبحث ويتضمن: مقدمة، الدراسات السابقة، مشكلة البحث، هدف البحث، أهمية البحث، مسوغات البحث، محددات البحث، حدود البحث، منهج البحث وأدواته، بالإضافة لبنية الأطروحة.
- يتناول الفصلان الثاني والثالث الأسس النظرية التي اعتمدنا عليها لتطوير خوارزميتنا. تم في الفصل الثاني اختيار بعض المبادئ الأساسية في معالجة الصورة الرقمية وتجزئتها وخضع اختيارها لفهم مراحل تطوير الخوارزمية، فتم استعراض مبادئ تشكيل الصور الرقمية الملونة والصور الحرارية، بالإضافة للعمليات والتوابع والمرشحات، ومبادئ تجزئة هذه الصور. أما في الفصل الثالث فقد تم عرض الأسس النظرية لخوارزميات الكشف وملاحقة الكائنات المتعددة.

- في الفصل الرابع، تم طرح مشاكل الملاحقة للكائنات المتعددة، تعترض هذه المشاكل العديد من خوارزميات الملاحقة المطورة، كذلك تم عرض للمشاكل الإضافية التي تعترض الملاحقة في الصورة الحرارية، ثم تم وضع إطار وطرائق الحل وفق منهجية عملنا. وفي نهاية الفصل عُرضت المعايير وطرائق الاختبار التي تضمن جودة العمل في طريقتنا المحسنة.
- يتناول الفصل الخامس التنفيذ العملي والبرمجي والاختبارات المنفذة على الخوارزمية المقترحة لتمييز الكائنات الحرارية المتداخلة ضمن الصورة الحرارية.
- في الختام تعرض الأطروحة بعض التوصيات والآفاق المستقبلية.

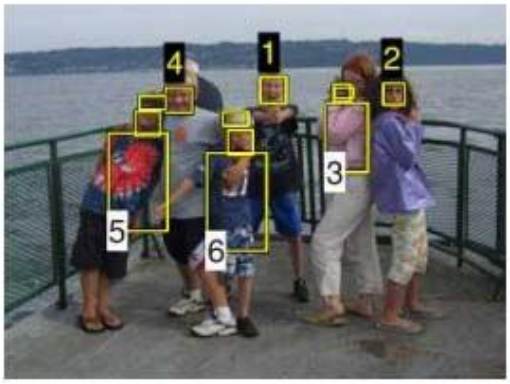
الفصل الثاني

الصورة الرقمية وتجزئتها

Digital image and segmentation

مقدمة:

حينما ننظر إلى صورة مشهد ما، فإن النظام البصري البشري لا يواجه أي مشكلة في تفسير الاختلافات في الشفافية وفصل الكائن بشكل صحيح عن خلفيته، كذلك بالنظر إلى صورة جماعية مؤطرة، يمكن بسهولة عد وتسمية كل الأشخاص الموجودين في الصورة كما في الشكل (2-1) وحتى تخمين مشاعرهم من تعبيرات وجوههم (Szeliski, 2022)، إذا ذهبنا باتجاه الحاسوب وأردنا من النظام الحاسوبي أن يحدد كل ما يمكن أن يحدده النظام البشري، سنلاحظ مدى الصعوبة التي يواجهها، فبالرغم من التقدم الكبير في تقانات معالجة الصورة خلال السنوات الأخيرة إلا أن أفضل ما تم الوصول إليه من نتائج لا يتخطى ما يستطيع طفل في أعوامه الأولى أن يحققه.



الشكل (2-1) القدرة على تمييز الكائنات في الصورة (Szeliski, 2022)

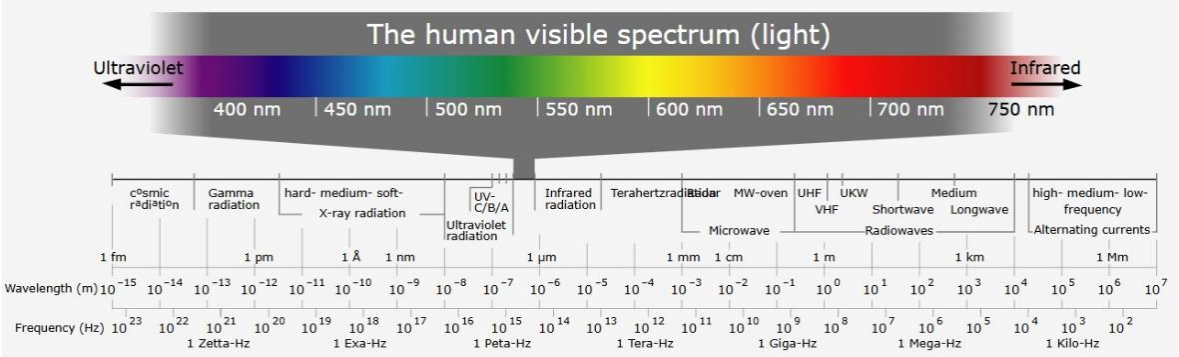
تعود صعوبة بناء نظام رؤية حاسوبي يعتمد الصورة الرقمية إلى أننا نسعى لاستعادة بعض المجهول نظرًا لعدم كفاية المعلومات لتحديد الحل بشكل كامل، لذلك يجب أن نلجأ إلى النماذج القائمة على الفيزياء والاحتمالية، أو التعلم الآلي من خلال مجموعات كبيرة من الأمثلة.

يتم في هذا الفصل عرض ملخص لبعض المبادئ النظرية الهامة، التي تم اعتمادها في هذا البحث للمساعدة في فهم وتحليل المشكلة، وبالتالي بناء خوارزمية الحل، حيث يناقش في تشكيل الصورة الرقمية بالتركيز على الصورة ذات التدرج الرمادي من المصدر الحراري، بالإضافة لعرض العديد من التوابع وتقنيات المعالجة لهذه الصورة، وهي التي استخدمت خلال تطوير طريقتنا في كافة مراحل التحليل والتجريب.

2-1 مصادر الصورة

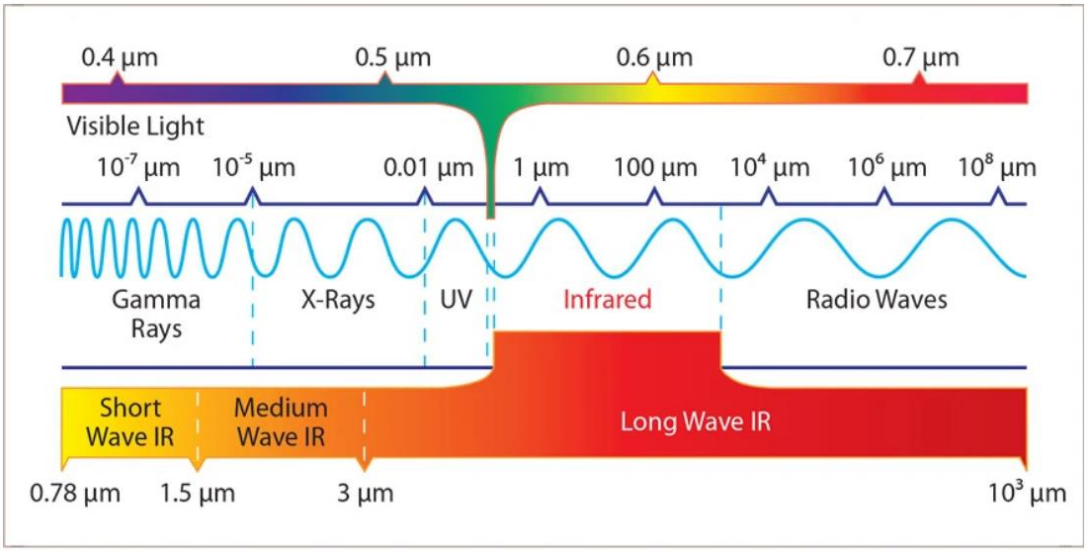
إن المصدر الرئيسي للصور المستخدمة اليوم هو طيف الطاقة الكهرومغناطيسية. يتم إنشاء الصور الاصطناعية، المستخدمة في النمذجة والتصور، بواسطة الحاسوب، تعتبر الصور المستندة إلى الإشعاع من الطيف الكهرومغناطيسي Electromagnetic spectrum هي الأكثر شيوعًا، وخاصة الصور في نطاقات الأشعة السينية والبصرية للطيف. يمكن تصور الموجات الكهرومغناطيسية على أنها موجات جيبية منتشرة ذات أطوال موجية مختلفة، أو يمكن تصورها على أنها تيار من الجسيمات، كل منها يسافر في نمط يشبه الموجة ويتحرك بسرعة

الضوء. يحتوي كل جسيم على كمية معينة (أو حزمة) من الطاقة. يعتبر النطاق المرئي للطيف هو الأكثر شيوعًا في جميع أنشطتنا، ويتفوق التصوير في هذا النطاق على جميع النطاقات الأخرى من حيث اتساع التطبيقات. في الشكل (2-2) يتم عرض الطيف المرئي مكبرًا لتسهيل المحتويات، ويلاحظ أنه يشمل نطاقًا ضيقًا للغاية من الطيف الكهرومغناطيسي الإجمالي. ويمتد هذا الطيف من البنفسجي Violet بطول موجة 0.400 ميكرون إلى الأحمر Red بطول موجة 0.750 ميكرون.



الشكل (2-2) الطيف المرئي ضمن الطيف الكهرومغناطيسي (Frank, 2023)

إذاً الطيف المرئي ليس سوى جزء صغير من الطيف الكهرومغناطيسي بأكمله. يمتد من الحافة الحمراء للطيف المرئي، نوع من الطاقة الإشعاعية غير المرئية، تسمى الأشعة تحت الحمراء (اشعاع حراري)، لها أطوال موجية أطول من الأطوال الموجية للضوء المرئي، ولها خاصية تسخين الأشياء التي تتعرض للإشعاع (Lea, 2011). في الشكل (2-3) نرى أن الأشعة تحت الحمراء مقسمة إلى مناطق ذات أطوال موجية تسمى الأشعة تحت الحمراء القريبة Near-infrared والأشعة تحت الحمراء متوسطة الطول الموجي Mid-wavelength-infrared والأشعة تحت الحمراء البعيدة Far-infrared، وكل منها سميت على اسم قربها النسبي من الطيف المرئي.



الشكل (2-3) الطيف المرئي وطيف الأشعة تحت الحمراء (Lea, 2011)

إن البشر مثل أي كائن حي هم أنفسهم مصدر للإشعاع تحت الأحمر. تشير معظم الأبحاث إلى أن البشر يصدرون إشعاعًا تحت أحمر بطول موجة يقع في المجال [3,50 µm].

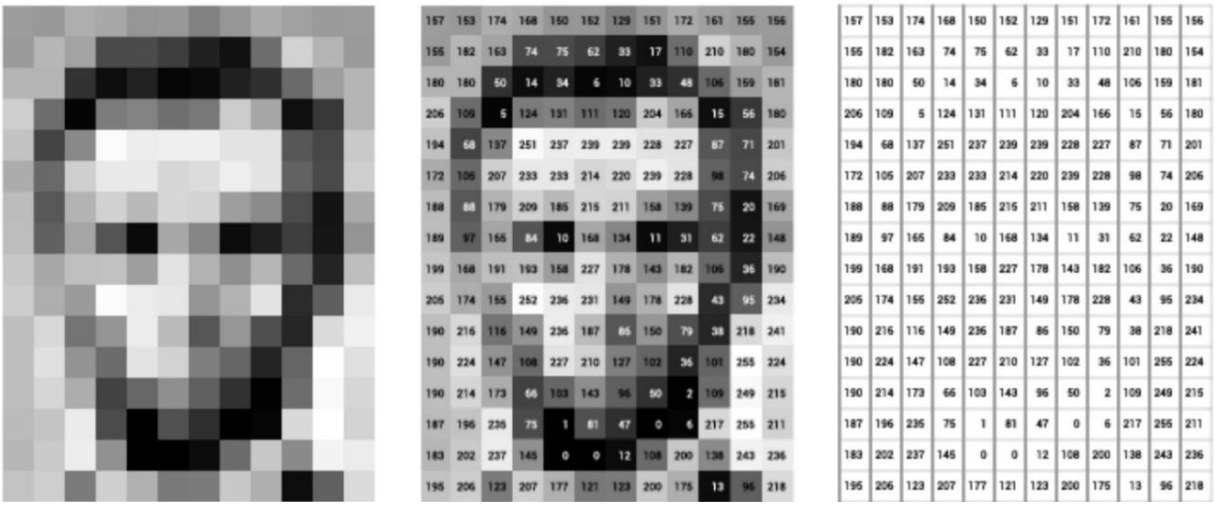
يمكن باستخدام عدد من قوانين الفيزياء، بما في ذلك ثابت بلانك، وثابت بولتزمان، وحسابات درجة الحرارة والضوء، حساب طول الموجةλ_{peak} (Lea, 2011) كما بالعلاقة (2-1).

$$\lambda_{peak} = \frac{2898}{T_k} = 9.5\mu m \qquad (1\text{-}2)$$

حيث T_k هي درجة الحرارة بالكالفن. لذلك نلاحظ أن أجهزة التصوير الحراري المستخدمة للرؤية البشرية الليلية عادة ما تكون حساسة في نطاق 7 إلى 14 µm.

2-2 الصورة الرقمية

يبين الشكل (2-4) مثالاً على تحصيل الصور الرقمية. حيث يقوم نظام التصوير بالتقاط أشعة مصدر الإضاءة الصادرة أو المنعكسة عن الهدف وبعدها يُسقط المشهد على مستوى للحصول على صورة رقمية.



الشكل (2-4) تحصيل الصورة الرقمية (Stanford, 2024)

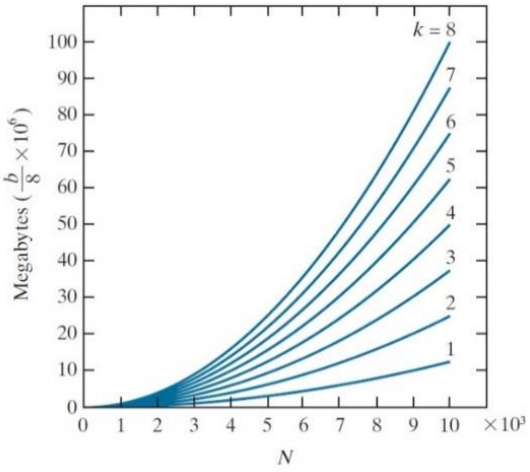
تترتب القيم الموافقة للطاقة المستقبلة من كل بقعة في المنظر ضمن مواقعها في مصفوفة ثنائية تمثل الصورة الرقمية. إذاً يشار إلى الصور بتابع ثنائي الأبعاد من النموذج f (x, y). قيمة f عند الإحداثيات المكانية (x, y) هي كمية قياسية يتم تحديد معناها الفيزيائي من خلال مصدر الصورة، وقيمها متناسبة مع الطاقة التي يشعها المصدر الفيزيائي. ويكون تابع الصورة بالعلاقة (2-2) (Szeliski,2022). حيث N,M أبعاد مصفوفة الصورة أما f(x,y) تابع كمية الطاقة بالنسبة للمنطقة المقابلة للإحداثي x, y .

إن القيمة العظمى للتابع تحد من إمكانيات وحدة المعالجة، فيمكن عند حجز بايت واحد للقيم أن تكون القيم محصورة في المجال [0,255] حيث 0 تمثل القيمة المكافئة للون الأسود والقيمة العظمى 255 مكافئة للون الأبيض،

$$f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix} \quad (2\text{-}2)$$

إن القيمة العظمى للتابع تحد من إمكانيات وحدة المعالجة، فيمكن عند حجز بايت واحد للقيم أن تكون القيم محصورة في المجال [0,255] حيث 0 تمثل القيمة المكافئة للون الأسود والقيمة العظمى 255 مكافئة للون الأبيض، وما بينهما يمثل مقدار تدرج الرمادي (Szeliski,2022)، وعدد هذه التدريجات خاضع للعلاقة 2^K حيث k=8 هي عدد البتات، وفي حالتنا بايت واحد، ويكون عدد البتات التي نحتاجها لتخزين الصورة بالعلاقة (3-2) (Szeliski,2022). يبين الشكل (2-5) عدد البايتات التي تحتاجها الصوة للتخزين حسب أبعادها N وقيمة k

$$B= N. M. k \quad (3\text{-}2)$$



الشكل (2-5) عدد البايتات التي تحتاجها الصوة للتخزين (Szeliski,2022)

يلاحظ التزايد الاسي الذي تحتاجه الصورة من البايتات سواءً للتخزين أو للمعالجة وذلك بتكبير أبعادها مما يؤدي لزيادة زمن المعالجة.

1-2-2 تمييزية الصورة

التمييزية Resolution هي مقياس لأصغر التفاصيل التي يمكن تمييزها في الصورة. ومن الناحية الكمية، يمكن التعبير عن الدقة المكانية بعدة طرائق مثل عدد النقاط لكل وحدة مسافة (نقطة لكل إنش dpi) وهي مقياس لتمييزية الصورة المستخدمة في صناعة الطباعة والنشر(Szeliski,2022)، وفي حالتنا نفضل استخدام مقدار الزاوية الصغرى الذي نستطيع تمييزه ضمن بيكسل واحد، فبفرض زاوية فتحة الكاميرا الافقية A هي 14 درجة وعرض المشهد N هو 340 بيكسل فتعطى التمييزية بالعلاقة (3-2)، (Szeliski,2022):

$$\text{Res} = \frac{\text{A}}{\text{N}} = 0.04^{\text{o}}/\text{P} \quad (3\text{-}2)$$

بالإضافة للتمييزية المكانية، من الشائع الإشارة إلى عدد البيتات المستخدمة لتحديد مقدار الكثافة باسم "تمييزية الكثافة". على سبيل المثال، من الشائع أن نقول إن الصورة التي يتم تقسيم شدتها إلى 256 مستوى لها تمييزية شدة تبلغ 8 بتات.

2-2-2 التوابع الهامة

تعتمد طرائق معالجة الصور على التغيير المباشر بقيم البكسلات في الصورة، تعمل التوابع على بكسلات مفردة من الصورة لمهام مثل التلاعب بالتباين وتحديد عتبة الصورة. ويقوم المرشح المكاني بإجراء عمليات على جوار كل بكسل في الصورة. تشمل أمثلة المرشحات المكانية على تنعيم الصورة وتوضيحها.

معكوس الصورة: يتم الحصول على الصورة السلبية Inverse image لصورة ذات مستويات شدة في النطاق [0, L-1] باستخدام التابع المعطى بالعلاقة (2-4)، (Gonzalez,2018):

$$fs=L-1-fr \qquad (4\text{-}2)$$

حيث fs هو الصورة الناتجة، وfr تمثل الصورة الدخل.

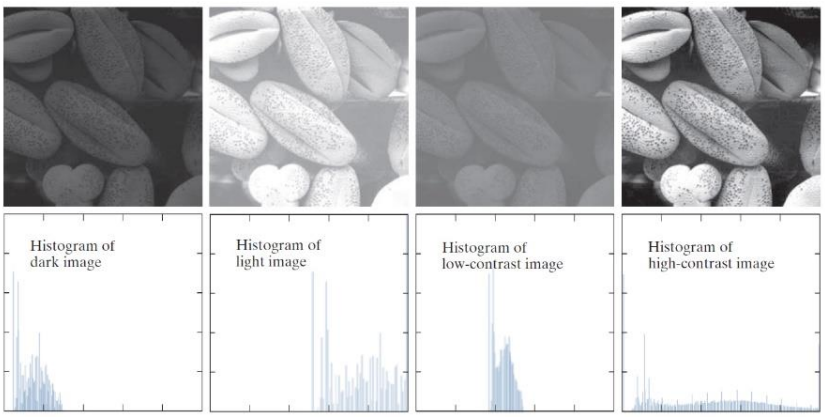
الهيستوغرام: يتم تعريف الهيستوغرام Histogram لـ f بالعلاقة (2-5)، (Gonzalez,2018):

$$h(r_k) = n_k : \ k = 0, 1, 2, \dots, L-1 \qquad (5\text{-}2)$$

حيث n_k هو عدد وحدات البكسل في f بكثافة r_k. ويتم تعريف الهيستوغرام الطبيعي لـ f بالعلاقة (2-6):

$$P(r_k) = \frac{h(r_k)}{M.N} = \frac{n_k}{M.N} \qquad (6\text{-}2)$$

يرتبط شكل الهيستوغرام بمظهر الصورة. على سبيل المثال، يوضح الشكل (2-6) صوراً بأربع خصائص أساسية للكثافة (Gonzalez,2018): الظلام، والضوء، والتباين المنخفض، والتباين العالي.

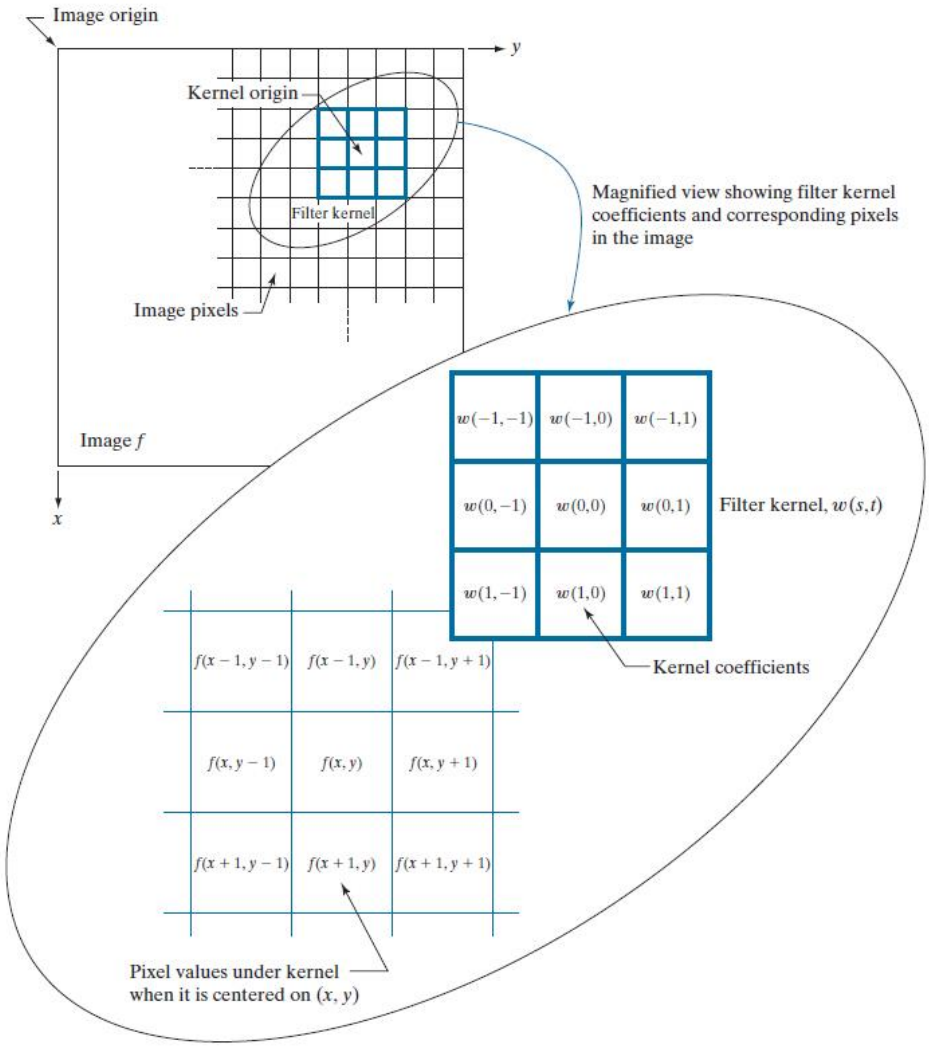


الشكل (2-6) هيستوغرام لصور عائدة لصورة وحيدة بظروف مختلفة (Gonzalez,2018)

كما تظهر الهيستوغرامات الخاصة بالصورة. نلاحظ في الصورة المظلمة أن خطوط الهيستوغرام الأكثر اكتظاظاً تتركز في الطرف السفلي (الغامق) من مقياس الكثافة. وبالمثل، فإن الخطوط الأكثر اكتظاظاً في الصورة الفاتحة منحازة نحو الطرف الأعلى من المقياس. تحتوي الصورة ذات التباين المنخفض على هيستوغرام ضيق يقع نحو منتصف مقياس الكثافة، بالنسبة للصورة، فإن هذا يعني مظهرًا رماديًا باهتًا. أخيرًا، نرى أن مكونات الهيستوغرام للصورة ذات التباين العالي تغطي نطاقاً واسعاً من مقياس الكثافة، مع وجود عدد قليل من الخطوط أعلى بكثير من الخطوط الأخرى.

3-2 المرشحات المكانية الخطية

يقوم المرشح المكاني الخطي linear spatial filter بإجراء عملية مجموع حاصل الضرب بين صورة f وقالب المرشح، w. والقالب هو مصفوفة يحدد حجمها جوار العملية، وتحدد معاملاتها طبيعة المرشح كما في الشكل (2-7).



الشكل (2-7) قالب المرشح (Gonzalez, 2018)

فيما يلي العلاقة (2-7) (Gonzalez,2018) التي تعطي قيمة بيكسل الصورة (x,y) بعد الترشيح بقالب w بأبعاد 3 x 3:

$$
\begin{aligned}
g(x,y) = \\
& w(-1,-1)f(x-1,y-1) + w(-1,0)f(x-1,y) + w(-1,+1)f(x-1,y+1) \\
& +w(\ 0,-1)f(x,y-1) \quad + w(0,0)f(x,y) \quad + w(0,+1)f(x,y+1) \\
& +w(+1,-1)f(x+1,y-1) + w(+1,0)f(x+1,y) + w(+1,+1)f(x+1,y+1)
\end{aligned} \tag{7-2}
$$

بشكل عام معادلة المرشح تعطى بالعلاقة (2-8):

$$g(x,y) = \sum_{s=-a}^{+a} \sum_{t=-b}^{+b} w(s,t)f(x+s,y+t) \tag{8-2}$$

في أي نقطة (x, y) في الصورة، تكون استجابة المرشح، g(x, y)، هي مجموع حاصل ضرب معاملات القالب أو النواة Kernel وبكسلات الصورة التي تحيط بها النواة.

2-3-1 الترابط والتلاف المكاني

يتم توضيح الترابط Correlation المكاني ووصفه رياضياً بالعلاقة (2-9)، (Gonzalez,2018):

$$(w \otimes f)(x,y) = \sum_{s=-a}^{+a} \sum_{t=-b}^{+b} w(s,t)f(x+s,y+t) \tag{9-2}$$

يتكون الترابط من تحريك النواة فوق الصورة، وحساب مجموع القيم في كل موقع. آليات التلاف Convolution المكاني هي نفسها، باستثناء أن نواة الترابط تدور بمقدار 180 درجة. ويعطى الالتفاف بالعلاقة (2-10):

$$(w * f)(x,y) = \sum_{s=-a}^{+a} \sum_{t=-b}^{+b} w(s,t)f(x+s,y+t) \tag{10-2}$$

2-3-2 مرشح التمرير المنخفض

كما ورد سابقاً يتكون المرشح المكاني الخطي من التفاف صورة باستخدام قالب المرشح. ويتم تحديد وظيفة هذا المرشح عن طريق تحديد قيم معينة للقالب أو النواة، ومن هذه المرشحات Low Pass Filter (LPF) مرشح التمرير المنخفض أو مرشح التنعيم، ويكون قالبه كما في الشكل (2-8).

$$\frac{1}{9}\times\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$$

الشكل (2-8) قالب مرشح التنعيم (Gonzalez, 2018)

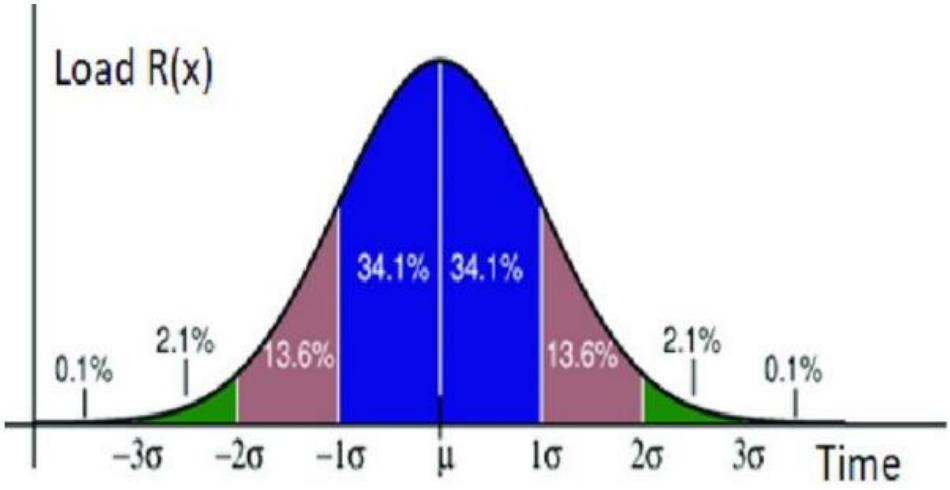
وهنا سيكون الناتج من تطبيقه هو صورة منعمة، كل بيكسل فيها ناتج عن محصلة متوسط Average هذه النقطة مع الجوار. ويستخدم هذا المرشح لتنعيم الصورة والتخلص من الضجيج فيها. تجدر الإشارة إلى أنه يتم تحديد درجة الضجيج المراد معالجته من خلال حجم النواة وقيم معاملاتها.

2-3-3 نوى مرشح غاوص منخفض التردد

نظرًا لبساطتها، فإن المرشحات الصندوقية مناسبة للتجريب السريع. وهي مفيدة أيضاً عندما يكون من المرغوب فيه تقليل تأثير التنعيم على الحواف. تتمتع المرشحات الغاوصية Gaussian بنفس المزايا الحسابية مثل المرشحات المتوسطة، ولكنها تتمتع بمجموعة من الخصائص الإضافية التي تجعلها مثالية لمعالجة الصور. وتعطى قيمة التابع الغاوصي بالعلاقة (2-11)، (Rudrabhatla, 2020)

$$w(s,t) = G(s,t) = \frac{1}{2\pi\sigma^2} e^{-\frac{s^2+t^2}{2\sigma}} \qquad (11\text{-}2)$$

ويبين الشكل (2-9) تابع التوزيع الغاوصي، والشكل (2-10) يمثل مرشح أو قالب تنعيم غاوصي



الشكل (2-9) تابع التوزيع الغاوصي (Rudrabhatla, 2020)

w=1/16*	1	2	1
	2	4	2
	1	2	1

الشكل (2-10) قالب المرشح الغاوصي 3 x 3 (Gonzalez, 2018)

في هذا القالب أعطيت أهمية (تثقيل أكبر) للخانة المركزية وبدرجة أقل للخانات الأبعد وهكذا بحيث يتماشى مع المعادلة السابقة.

2-3-4 التنعيم التكيفي

هناك طريقة مرغوبة أكثر وهي التنعيم الانتقائي، أي التنعيم في منطقة الضجيج فقط (الانتقال اللوني الكبير في شدة الاضاءة)، وليس التنعيم في المنطقة الخالية من الضجيج. بهذه الطريقة يحتمل أن يتم التقليل من تأثير التمويه، يطلق عليها التصفية التكيفية، ويتم ذلك عندما يكون الفرق بين الشدة القصوى والحد الأدنى من الكثافة للمنطقة المحلية أكبر من عتبة معينة T،

تعطى قيمة التباين في المنطقة R هو dR بالعلاقة (2-12)، (Gonzalez, 2018).

$$dR = \max(R) - \min(R) > T \qquad (12\text{-}2)$$

2-3-5 مرشح التمرير المرتفع (المرشحات الحدية)

تسلط المرشحات الحدية الضوء على التحولات في الكثافة. نظرًا لأن حساب المتوسط يشبه التكامل، أي أنه يمكن تحقيق مرشح الحدية من خلال الاشتقاق المكاني. وهو يعطي أهمية للتغيرات الكبيرة في الكثافة، مثل الحواف. غالباً ما يُشار إلى التنعيم باسم الترشيح منخفض التمرير(High Pass Filter (HPF، أي السماح فقط بانتقالات المستوى الصغيرة وهو مصطلح مستعار من معالجة المجال الترددي. غالباً ما يُشار إلى الحدية باسم الترشيح عالي التمرير. في هذه الحالة، يتم تمرير الترددات العالية (المسؤولة عن التفاصيل الدقيقة أي انتقالات المستوى الكبيرة)، بينما يتم إضعاف الترددات المنخفضة أو حذفها. تعتمد مرشحات الحدية على المشتقات من الدرجة الأولى والثانية على العلاقة (2-13)، (Gonzalez, 2018):

$$\frac{\partial f}{\partial x} = f(x+1) - f(x) \qquad (13\text{-}2)$$

ومن أجل المشتق من الدرجة الثانية (المشتق اللابلاسي Laplacian) ويعطى بالعلاقة (2-14):

$$\frac{\partial^2 f}{\partial^2 x} = f(x+1) + f(x-1) - 2f(x) \qquad (14\text{-}2)$$

ومن أجل البعدين x,y تصبح العلاقة (2-15):

$$\nabla f = \frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial^2 y} \qquad (15\text{-}2)$$

يمثل الشكل (2-11) عدة أشكال للمرشح اللابلاسي

0	1	0	1	1	1	0	−1	0	−1	−1	−1
1	−4	1	1	−8	1	−1	4	−1	−1	8	−1
0	1	0	1	1	1	0	−1	0	−1	−1	−1

الشكل (2-11) المرشح اللابلاسي (Hardan, 2018)

2-4 تجزئة الصورة

تعتبر تجزئة الصورة Image segmentation من أهم المعالجات على الصورة التي تسمح باستخلاص الكائنات وفصلها عن الخلفية (Parker, 2010)، ويمكن تعريف تجزئة الصورة رياضياً كما يلي:

إذا كانت R تمثل المنطقة المكانية بالكامل التي تشغلها صورة. يمكننا أن ننظر إلى تجزئة الصورة كعملية

1- تقسم R إلى n منطقة فرعية، R1, R2,.. ,Rn وهي تعطى بالعلاقة (2-16)، (Parker, 2010):

$$U_{i=1}^{n} R_i = R \qquad (16\text{-}2)$$

2- إن المجموعة R_i هي مجموعة متصلة

3- من أجل $i \neq j$ فإن: $R_i \cap R_j = \Phi$

4- نقاط المجموعة Ri لها الكثافة ذاتها: $Q(R_i) = True$

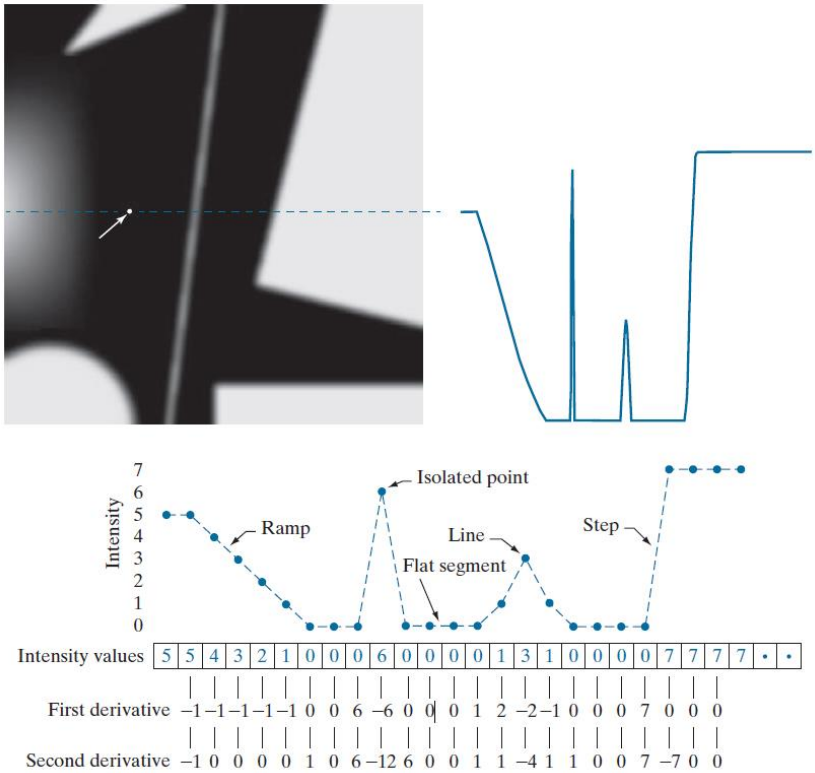
5- نقاط المجموعات المختلفة ليست لها ذات الكثافة: $(R_i \cup R_j) = False$

إن المشكلة الأساسية في التجزئة هي تقسيم الصورة إلى مناطق تلبي الشروط السابقة. تعتمد خوارزميات التجزئة للصور أحادية اللون بشكل عام على واحدة من فئتين أساسيتين تتعاملان مع خصائص قيم الكثافة: عدم الاستمرارية والتشابه. في الفئة الأولى، نفترض أن حدود المناطق مختلفة بشكل كافٍ عن بعضها البعض، وعن الخلفية، للسماح باكتشاف الحدود بناءً على عدم الاستمرارية المحلية في الكثافة. التجزئة القائمة على الحافة هي النهج الرئيسي

المستخدم في هذه الفئة. تعتمد طرائق التجزئة القائمة على المنطقة في الفئة الثانية على تقسيم الصورة إلى مناطق متشابهة وفقًا لمجموعة من المعايير المحددة مسبقًا.

2-4-1 كشف النقاط والخطوط والحواف

تعتمد طرائق الكشف على اكتشاف التغيرات الحادة والمحلية في الكثافة. إن وحدات البيكسل على الحافة هي وحدات بكسل تتغير عندها شدة نقاط الصورة بشكل مفاجئ، يلاحظ في الشكل (2-12) مجموعات من وحدات البيكسل على الحافة المتصلة. وكما رأينا سابقاً فإن المشتقات من الدرجة الأولى أو الثانية ستكون مناسبة للكشف.



الشكل (2-12) المشتق الأول والثاني (Hardan, 2018)

يمكن تنفيذ المشتقات الأولى في معالجة الصور باستخدام مقدار التدرج gradient. يتم تعريف تدرج الصورة f عند الإحداثيات (x,y) على أنه مصفوفة عمودية ثنائي الأبعاد وتعطى بالعلاقة (2-17)، (Hardan, 2018):

$$\nabla f = \text{grad}(f) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \qquad (17\text{-}2)$$

ويكون مطال التدرج معطى بالعلاقة (2-18):

$$M(x,y) = \|\nabla f\| = \text{mag}(\nabla f) = \sqrt{{g_x}^2 + {g_y}^2} \qquad (18\text{-}2)$$

يبين الشكل (2-13) استخدام القوالب لكشف الحواف (Hande, 2021).

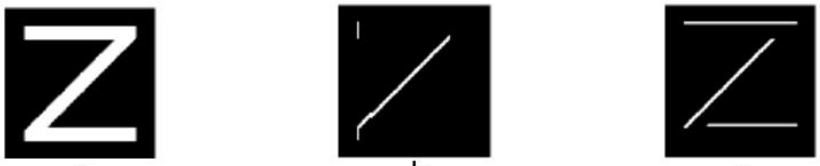
القالب الأيسر في الشكل (13-2-a) لكشف الحواف الأفقية والقالب الأيمن في الشكل (13-2-a) يستخدم للحواف الشاقولية، أما في الشكل (14-2) فيبين نتائج تطبيق القوالب على صورة أصلية a، في الصورة التالية b ناتج تطبيق قالب كشف الحواف الشاقولية ثم الناتج c ناتج تطبيق قالب كشف الحواف الأفقية.

−1	−2	−1
0	0	0
1	2	1

−1	0	1
−2	0	2
−1	0	1

a b

الشكل (13-2) قالب Sobel لكشف الحواف (Hande, 2021)



a b c

الشكل (14-2) ناتج كشف الحواف (Kalipayun, 2018)

2-4-2 تابع العتبة

نظرًا لخصائصه البديهية، وسهولة تنفيذه، وسرعته الحسابية، يتمتع تابع تحديد العتبة للصورة بمكانة مركزية في تطبيقات تقسيم الصور. يتم تعريف تابع العتبة وفق العلاقة (19-2):

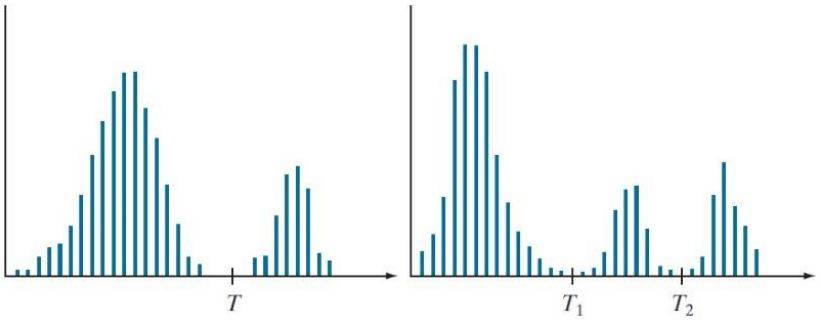
$$g(x,y) = \begin{cases} 1 & \text{if} \quad f(x,y) > T \\ 0 & \text{if} \quad f(x,y) \le T \end{cases} \qquad (19\text{-}2)$$

حيث T عبارة عن عتبة تمثل مستوى لوني معين في تدرج الرماديات، وقد تم استخدام القيمة المنطقية 1/0 للناتج والتي ستحدد ناتج التابع بحيث أن قيمة الـ 1 المنطقية للمناطق الفاتحة وشدة اضاءتها أكبر من قيمة العتبة T، ويكون الناتج هنا تقسيم الصورة لقسمين.

يمكن أن تكون العتبة متغيرة مثلا عتبتان لتحديد أكثر من قسمين وتصبح العلاقة (20-2)، (Gonzalez,2018):

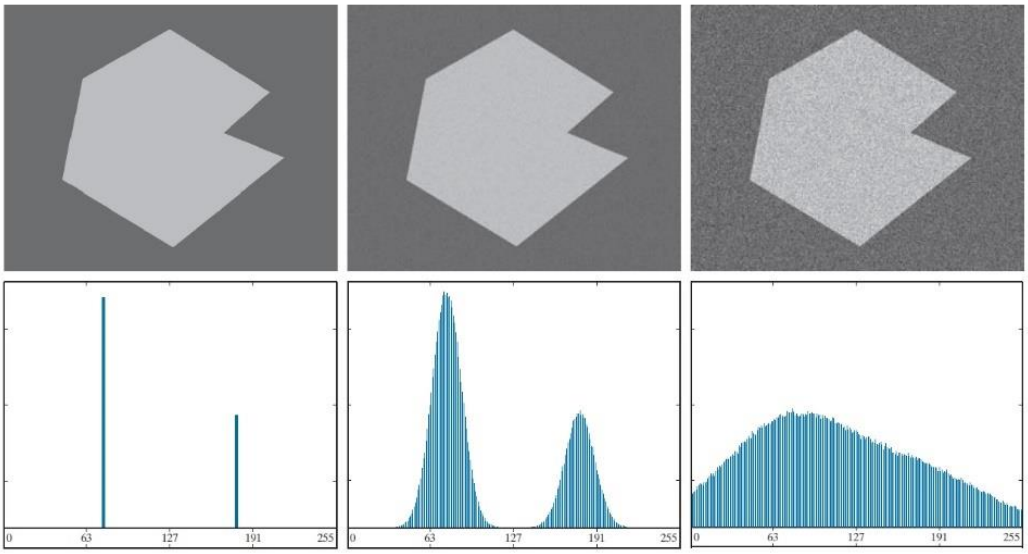
$$g(x,y) = \begin{cases} a & \text{if} \quad f(x,y) > T_2 \\ b & \text{if} \quad T_1 \le f(x,y) \le T_2 \\ c & \text{if} \quad f(x,y) \le T_1 \end{cases} \qquad (20\text{-}2)$$

ويبين الشكل (15-2) هيستوغرام الصور التي نريد اجراء التجزئة عليها



الشكل (2-15) تابع العتبة في الهيستوغرام (Gonzalez,2018)

بناءً على ما سبق، يمكننا أن نستنتج بشكل حدسي أن نجاح التجزئة مرتبط بشكل مباشر بعرض وعمق المنطقة التي تفصل بين أنماط الهيستوغرام، فإن العوامل الرئيسية التي تؤثر على النجاح هو الاختيار الصحيح لقيم T. ويرينا الشكل (2-16) تأثير الضجيج على اختيار T فبالصورة اليسرى الشكل (2-16-a) الصورة خالية من الضجيج والصورة التالية (b) صورة مضجة بضجيج غاوصي بمتوسط 0 وانحراف معياري يبلغ عشر مستويات لونية، أما الثالثة (c) فهي صورة مضجة بضجيج غاوصي بمتوسط 0 وانحراف معياري يبلغ 50 مستوى لوني.



a b c

الشكل (2-16) استخدام الهيستوغرام في التجزئة (Gonzalez,2018)

5-2 العمليات المورفولوجية

تشير كلمة المورفولوجي Morphology عادةً إلى فرع من فروع علم الأحياء يتعامل مع شكل وبنية الحيوانات والنباتات. تستخدم الكلمة ذاتها في سياق علم الأشكال الرياضية كأداة لاستخراج مكونات الصورة المفيدة في تمثيل ووصف شكل المنطقة، مثل الحدود والهياكل والغلاف. ويتم الاهتمام أيضاً بالتقنيات الشكلية للمعالجة المسبقة أو اللاحقة (Gonzalez,2018)، و(Parker, 2010)، كالترشيح الشكلي والتخفيف والتقليم. والفقرات التالية تبين أهم فوائد العمليات المورفولوجية

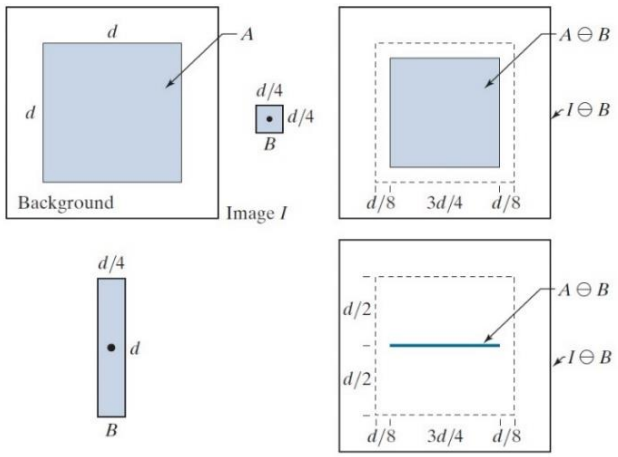
1-5-2 الحت

يعطى الحت Erosion بالعلاقة (2-21)، (Gonzalez,2018):

$$A \ominus B = \left\{ z \middle| (B)_z \subseteq A \right\} \tag{21-2}$$

حيث A هي مجموعة من وحدات البيكسل الأمامية، وB هي عنصر هيكلي، وz هي قيم المقدمة.

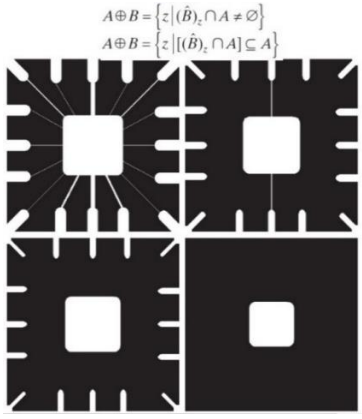
أي الناتج هو مجموعة جميع النقاط z بحيث تكون B، المترجمة بواسطة z، موجودة في A. وبكلمات أخرى، تشير هذه المعادلة إلى أن تآكل A بواسطة B. يبين الشكل (2-17) الناتج بحيث تم حت A بمقدار أبعاد نصف B



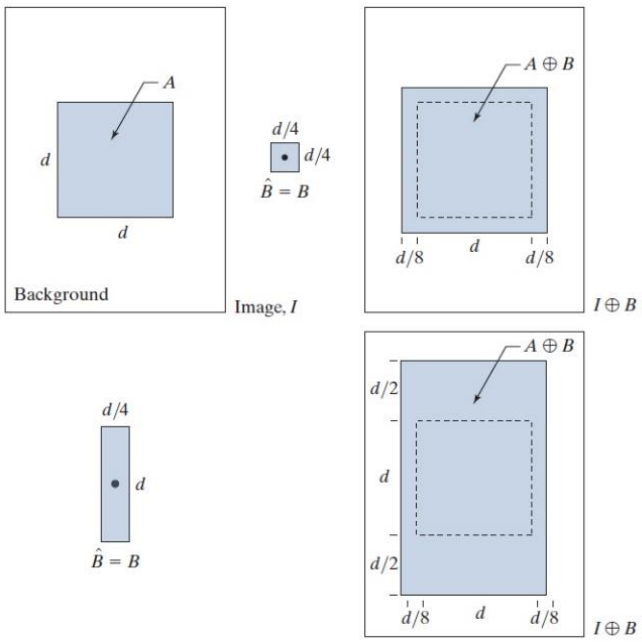
الشكل (2-17) الحت (Gonzalez,2018)

2-5-2 التمددDilation

يؤدي التابع لتمدد المنطقة المعالجة Dilation بحيث يقضى على المناطق البينية الرفيعة (Gonzalez,2018)، ويبين الشكل (2-17) عمل التمدد، بينما يبين الشكل (2-19) ناتج التمدد.

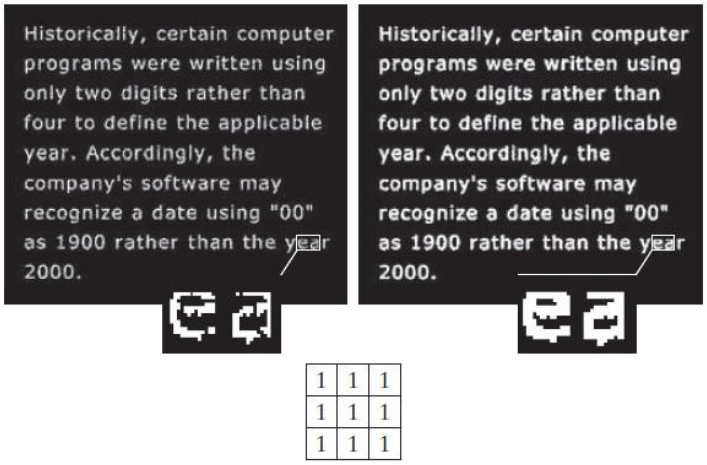


الشكل (2-18) مثال عن التمدد (Gonzalez,2018)

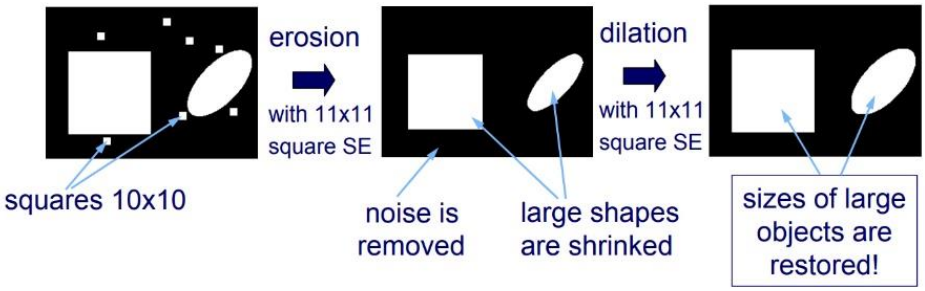


الشكل (2-19) ناتج التمدد (Gonzalez,2018)

يبين الشكل (2-20) ترميم الأحرف المتآكلة باستخدام تقنية التمدد، أما الشكل (2-21) فيبين اجراء الحت ثم التمدد لحذف الاجسام الصغيرة والبقاء ضمن حجم الكائن الأصلي فيما يسمى عملية الفتح.



الشكل (2-20) اصلاح الأحرف المتآكلة بالتمدد (Gonzalez,2018)



الشكل (2-21) الحت والتمدد لإزالة الأجزاء الصغيرة (Tarabalka,2022)

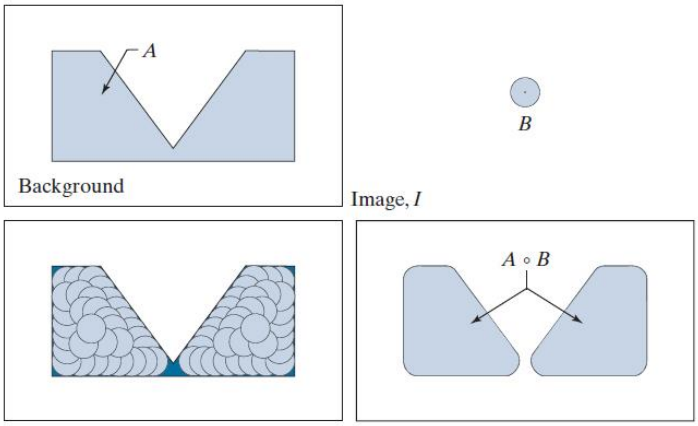
3-5-2 الفتح والاغلاق

تعطى علاقتي الفتح والاغلاق Opening & Closing بالعلاقات (2-22)، (Gonzalez,2018):

$$A \circ B = (A \ominus B) \oplus B$$
$$A \bullet B = (A \oplus B) \ominus B \qquad (22\text{-}2)$$

تستخدم تقنية الفتح لفصل كائنين ملتصقين كما في الشكل (2-22).

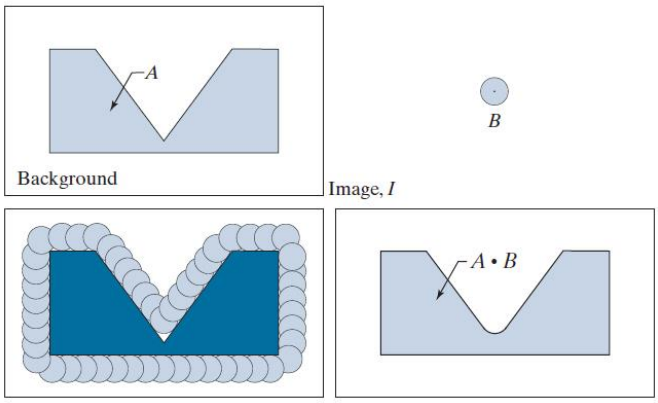
$$A \circ B = \bigcup \{(B)_z \mid (B)_z \subseteq A\}$$



الشكل (2-22) تقنية الفتح (Gonzalez,2018)

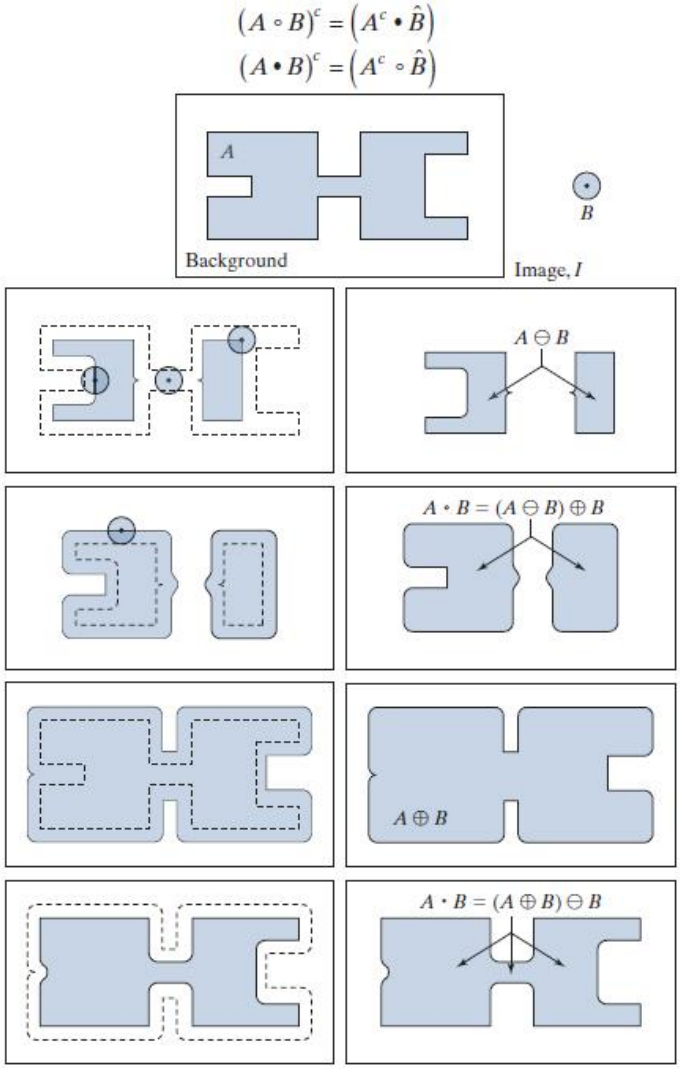
ويري الشكل (2-23) نتائج تطبيق تقنية الاغلاق على ذات الصورة.

$$A \bullet B = \left[\bigcup \{(B)_z \mid (B)_z \cap A = \varnothing\}\right]^c$$



الشكل (2-23) تقنية الاغلاق (Gonzalez,2018)

بالتالي يمكن استخدام هذه التقنية لوصل أجزاء متقاربة لكائن. يعرض الشكل (2-24) بعض التطبيقات الهامة لهذه التقنيات.



الشكل (2-24) بعض الاصلاحات التي تقدمها تقنيات الفتح والإغلاق (Gonzalez,2018)

خلاصة الفصل الثاني:

تم في هذا الفصل عرض للمبادئ النظرية لتشكيل الصورة الرقمية بالتركيز على الصورة ذات التدرج الرمادي من المصدر الحراري، بالإضافة لعرض العديد من التوابع وتقنيات المعالجة لهذه الصورة التي استخدمت خلال تطوير الطريقة المقترحة في كافة مراحل التحليل والتجريب.

الفصل الثالث

بنى وخوارزميات ملاحقة الكائنات المتعددة

Multi-Object Tracking Structures and Algorithms

مقدمة:

في هذا الفصل يتم عرض مبادئ وخوارزميات هامة تخدم موضوع الاطروحة سواء بشكل مباشر أو بشكل غير مباشر، ويقسم الموضوع إلى ستة أقسام تتعلق بكشف الكائنات ومن ثم ملاحقتها.

تتمكن الملاحقة الحرارية بالأشعة تحت الحمراء من ملاحقة الهدف في الظلام الدامس، بالتالي العمل في مجالات حساسة مثل حماية الخصوصية (Liu, 2017) بينما يفشل التتبع البصري دائمًا في حالة الرؤية السيئة. لكن بالرغم من ذلك تواجه جودة الملاحقة الحرارية صعوبات جمة، تتعلق بسمات الكائن الحراري التي تؤثر بشكل كبير على جودة الملاحقة، لذلك في بداية هذا الفصل تم عرض للسمات العامة المهمة في الكائنات والتي يعتمد عليها كاشف الكائنات، وقدمنا بعض الخوارزميات التي تكشف هذه السمات.

بعد مناقشة السمات يتم عرض ملخصاً عن مبادئ الذكاء الاصطناعي الذي يعتبر من أهم التقانات الحالية التي تقدم حلولا جيدة لكافة المشاكل المعاصرة وبالأخص مشاكل الصورة والفيديو، حيث تعد خوارزميات الذكاء الاصطناعي في ملاحقة الكائنات من أهم الخوارزميات المستخدمة في مجال الرؤية الحاسوبية والتي تهدف بشكل عام إلى تتبع الكائنات منذ بداية ظهورها في الفيديو أو على عدسة الكاميرا لحين اختفائها.

بعد دراسة ملخصة عن بعض السمات تمت دراسة أخرى عن نماذج الحركة ومرشح كالمن الشهير، وهو تقانة هامة تم استخدامها بفعالية في آلية وطرائق الحل ضمن النظام المقترح.

في نهاية الفصل يتم شرح المبادئ النظرية لخوارزميات الكشف والملاحقة التي تم الاعتماد عليها في المنهج المقترح.

3-1 السمات وطرائق تحديدها

عندما يكون المراد تحديد كائنات داخل مشهد ما فهناك حاجة لبعض الموصِّفات التي تصف هذا الكائن، ومجموع الموصِّفات الموجودة في هذا الكائن تشكل مجموعة متميزة لا تدل على صنف آخر، تدعى هذه الموصِّفات بالسمات. يتم استخدام السمات لإنشاء نموذج الذي يصنف الأشياء في الصورة، ومن أهم السمات الحواف، الزوايا، الشكل، اللون، الحجم الاتجاه، الخ... ولابد من الإشارة أن السمات يمكن استخدامها لأكثر من تحديد أصناف الكائنات، مثلاً يمكن استخدامها لتحديد وتمييز الكائنات من بعضه ضمن ذات الصنف.

3-1-1 خوارزميات كشف الحواف

هناك عدد من التعريفات الممكنة للسمة، كل منها قابل للتطبيق في ظروف محددة مختلفة. أحد التعريفات الأكثر شيوعًا والأكثر عمومية هو حافة الخطوة المثالية الموضحة في الشكل (3-1) الصورة اليسرى a. لأنه من السهل اكتشافها: في حالة عدم وجود ضجيج، فإن أي تغيير كبير في المستوى الرمادي يشير إلى حافة.

a b

الشكل (3-1) الحافة تمثل انتقال بالمستوى اللوني (Malche, 2024)

يعد اكتشاف الحواف Edge Detection تقنية لمعالجة الصور لإيجاد حدود الكائنات داخل الصور. إنه يعمل عن طريق الكشف عن الانقطاعات في السطوع. قد يكون هذا مفيدًا جدًا في استخراج معلومات مفيدة من الصورة لأن معظم معلومات الشكل تكون محاطة بالحواف. تعمل طرائق الكشف الكلاسيكية عن الحواف من خلال الكشف عن الانقطاعات في خوارزمية الكشف بالاشتقاق، ونظرًا لأن الصور ثنائية الأبعاد، فمن المهم مراعاة تغييرات المستوى في العديد من الاتجاهات. لهذا السبب، يتم استخدام المشتقات الجزئية للصورة، فيما يتعلق بالاتجاهات الرئيسية x و y. يمكن الحصول على تقدير لاتجاه الحافة الفعلي باستخدام المشتقات في x وy كمكونات للاتجاه الفعلي على طول المحاور، وحساب مجموع المتجه. يحدث أن يكون العامل المعني هو التدرج في الشدة، كل ذلك تم شرحه في فقرة "المرشحات الحدية". سيكون حجم الحافة عدداً حقيقياً، وعادةً ما يتم تحويله إلى عدد صحيح عن طريق التقريب. يُقال إن أي بكسل له تدرج يتجاوز قيمة عتبة محددة هو بكسل حافة، والبعض الآخر ليس كذلك. من الناحية الفنية، سيقوم كاشف الحواف بالإبلاغ عن بكسلات الحواف فقط.

3-1-1-1 خوارزمية الاعتماد على القوالب Sobel

هناك مجموعة واسعة من أدوات الكشف عن الحواف القائمة على القوالب. تم اختيار اثنتين كمثالين هنا، وذلك ببساطة لأنهما يوفران أفضل مجموعات بكسلات الحواف أثناء استخدام قالب صغير. إن أكثر خوارزميات الكشف عن الحواف شيوعًا هي خوارزمية الكشف عن الحواف Sobol، كما هو مبين في الشكلين (3-2) و(3-3).

الشكل (3-2) كشف الحواف (Tian, 2021)

الشكل (3-3) كشف الحواف بخوارزمية Sobel (Yuheng, 2017)

يتكون معامل الكشف سوبل من نواة تلافيفية 3*3. نواة بسيطة Gx ونواة مستديرة 90 درجة Gy. يتم إجراء قياسات منفصلة عن طريق تطبيق كل من النواة بشكل منفصل على الصورة (Yuheng, 2017). وتعطى مصفوفة نواة سوبل (Yuheng, 2017) كما يلي:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \qquad G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

عند وجود حافة مائلة سينتج لدينا تغيير بالمعاملين Gx وGy وذلك سيعني اتجاه الحرف، وعندها يمكن حساب متوسط النتائج. لكل بكسل في الصورة، ويجب تحديد المقدار الناتج.

سيحسب مقدار التدرج وزاويته من العلاقات (3-1)، (Yuheng, 2017):

$$G = \sqrt[2]{{G_x}^2 + {G_y}^2} \qquad \theta = \arctan(\frac{G_y}{G_x}) \qquad (1\text{-}3)$$

3-1-1-2 خوارزمية الاعتماد على القوالب Kirsch

المثال الثاني لاستخدام القوالب هو المثال الذي وصفه Kirsch، وتم اختياره كمثال هنا لأن هذه النماذج لها دافع مختلف عن دافع Sobel. بالنسبة للحالة 3 x 3، تكون القوالب كما يلي:

$$K0 = \begin{matrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{matrix} \quad K1 = \begin{matrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{matrix} \quad K2 = \begin{matrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{matrix} \quad K3 = \begin{matrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{matrix}$$

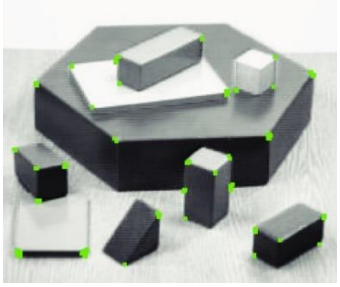
$$K4 = \begin{matrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{matrix} \quad K5 = \begin{matrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{matrix} \quad K6 = \begin{matrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{matrix} \quad K7 = \begin{matrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{matrix}$$

هذه الأقنعة هي محاولة لنمذجة نوع تغيير مستوى الرمادي الذي يُرى بالقرب من حافة ذات اتجاهات مختلفة، بدلاً من تقريبها إلى التدرج اللوني. يوجد قناع واحد لكل اتجاه من اتجاهات البوصلة الثمانية.

على سبيل المثال، تشير الاستجابة الكبيرة للقناع K0 إلى حافة عمودية (تدرج أفقي) عند البيكسل المقابل لمركز القناع. يبدو أن مشغل Kirsch هو الأفضل بين مشغلي القوالب، على الرغم من أن الاثنين قريبان جدًا. يتفوق كل من مشغلي القالب على مشغلي المشتقات البسيطة، خاصة مع زيادة الضجيج.

3-1-2 بعض خوارزميات كشف الزوايا

الزاوية هي النقطة التي يقف جوارها المحلي في اتجاهين مهيمنين ومختلفين. بمعنى آخر، يمكن تفسير الزاوية على أنها تقاطع حافتين، حيث تكون الحافة تغييرًا مفاجئًا في إضاءة الصورة. الزوايا هي السمات المهمة في الصورة، ويطلق عليها عمومًا كنقاط اهتمام ثابتة للترجمة، والتناوب، والإضاءة. إن الزوايا هي مجموعة فرعية من النقاط المهمة في الصورة، والتي يمكن تعريفها ببساطة على أنها أي نقطة في الصورة يتغير عندها الإشارة بشكل كبير ثنائي الأبعاد (Juranek, 2018).



الشكل (3-4) كشف الزوايا (Juranek, 2018)

تظهر أهمية كشف الزاوية Corner detection في الكشف عن نقاط الاهتمام المميزة في اكتشاف المناطق المتشابهة حتى بتغير إضاءة الصورة أو ميلانها. فكشف الزاوية مهم في توصيف الأشكال حيث يصعب توصيف كل الأشكال في الصور بطرائق أخرى بينما تساهم مناطق الزوايا التي تعتبر مميزة.

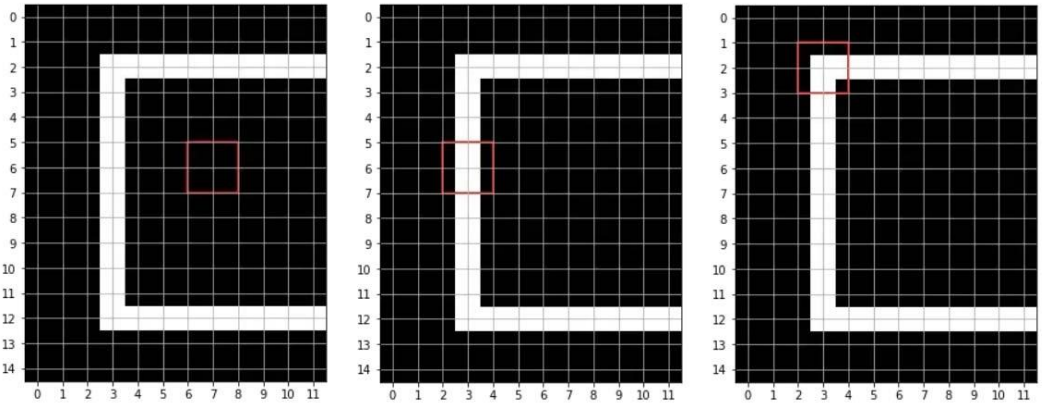
1-2-1-3 خوارزمية Moravec

تعتبر خوارزمية Moravec من أقدم الخوارزميات لكشف الزاوية. لنعتبر المصفوفة المبينة في الشكل (3-5).

```
img = np.array([
            [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 100, 100, 100, 100, 100, 100, 100, 100, 100],
            [0, 0, 0, 100, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 100, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 100, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 100, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 100, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 100, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 100, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 100, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 100, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 100, 100, 100, 100, 100, 100, 100, 100, 100],
            [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0],
            [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0],
])
```

الشكل (3-5) مصفوفة لاختبار Moravec (Thapa, 2022)

وهي ممثلة للصورة الموجودة في الشكل (3-6). عندما نزيح صورة التصحيح بالاتجاهات الثمانية، يمكننا حساب مجموع المسافة المربعة بين الصورة وصورة القناع المزاح كفرق بين الصورتين.



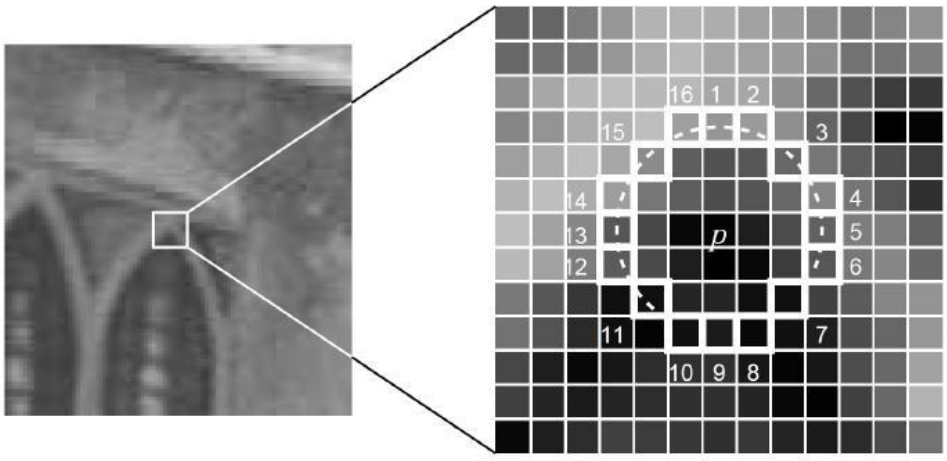
الشكل (3-6) حالات اختبار المنطقة المنبسطة والحافة الشاقولية والافقية (Thapa, 2022)

سنلاحظ ثلاث حالات:

1- القناع في منطقة متجانسة الإضاءة (لا يوجد فرق كبير).
2- القناع في منطقة حافة (سيؤدي التحول على طول الحافة إلى تغيير طفيف، لكن التحول بشكل عمودي على الحافة سيؤدي إلى تغيير كبير).
3- القناع في منطقة زاوية (ستؤدي جميع الإزاحات إلى تغيير كبير).

3-1-2-2 خوارزمية كاشف الزوايا السريع

تتميز خوارزمية كاشف الزوايا السريع Fast corner detector بسرعة الكشف لأن فكرتها بسيطة (Medium, 2019). بأخذ النقطة المعتبرة P في الشكل (3-7)، والبكسلات الأخرى في دائرة ذات نصف قطر معين وخصائص بكسل المركز. عندما يصل الفرق إلى معيار معين، تعتبر النقطة نقطة ركن.



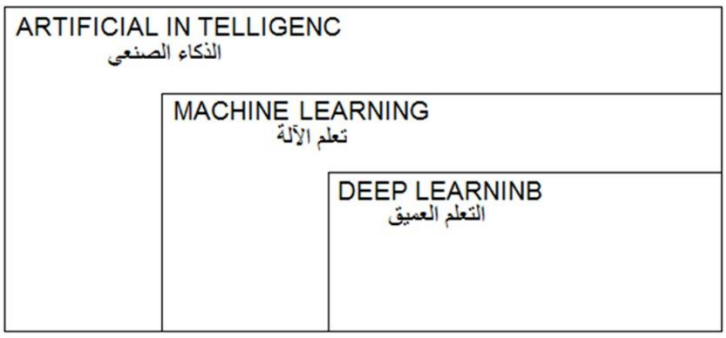
الشكل (3-7) مبدأ الكشف السريع للزاوية (Medium, 2019).

بعد الاختبار، وجد أنه عندما يكون نصف قطر الدائرة المحدد 3، يمكن موازنة نتيجة الكشف والكفاءة. كما هو موضح في الشكل (3-7). تتم مقارنة النقطة p مع 16 بكسل في الدائرة التي يبلغ نصف قطرها 3 لمقارنة فرق المستويات الرمادية، إذا كانت القيمة المطلقة لاختلاف التدرج الرمادي أكبر من عتبة معينة، فإن عدد النقاط التي فيها العلاقة محددة، يتجاوز العدد المحدد تعتبر النقطة نقطة ركن. هنا الرقم المحدد هو 12.

إن الخوارزمية تعتمد فكرة بسيطة وكمية صغيرة من الحساب، وهي مناسبة لتطبيقات الكشف في الوقت الحقيقي.

2-3 الذكاء الاصطناعي

إن الذكاء الاصطناعي Artificial Intelligence (AI) هو أحد المحركات الرئيسية لتغيير الحياة الاجتماعية والاقتصادية. ومن جهة أخرى هو يساهم في تقدم أحدث التقنيات في العديد من مجالات الدراسة، كأدوات مفيدة للأبحاث الرائدة (Jiang, 2022). يشير مفهوم الذكاء الاصطناعي إلى محاكاة الذكاء البشري في الآلات المبرمجة للتفكير والتصرف مثل البشر. تم تصميم أنظمة الذكاء الاصطناعي لأداء المهام التي تتطلب عادةً الذكاء البشري، مثل حل المشكلات واتخاذ القرار والتعلم. يمكن تحقيق الذكاء الاصطناعي من خلال تقنيات مختلفة. تسمح هذه التقنيات للآلات بمعالجة كميات كبيرة من البيانات والتعلم منها واتخاذ القرارات بناءً على الأنماط المستفادة. ينمو الذكاء الاصطناعي ويتطور بسرعة، مع تطبيقات في العديد من الصناعات، بما في ذلك الرعاية الصحية والتمويل والترفيه. ويضم الذكاء الاصطناعي تعلم الآلة والتعلم العميق كما هو مبين بالشكل (3-8)



الشكل (3-8) التعلم العميق في الذكاء الاصطناعي (الباحث)

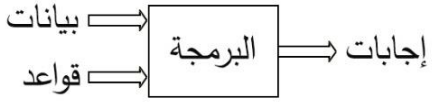
يمكن تشبيه عملية الذكاء الاصطناعي وتعلم الآلة بالمولود الجديد الذي لا يستطيع تعلم أو عمل أي شي بمفرده إذا لم يعلمه والداه ويدربانه وينقلان له المعرفة ويمكنانه من التعرف على ما هو خطأ وما هو صواب، وبذلك تتعلم الآلة وتصبح قادرة على اتخاذ القرارات وإعطاء الاستنتاجات واقتراح الحلول.

1-2-3 تعلم الآلة

نعرف أن الآلة هي مجرد عتاد قابل للبرمجة ولكن في عصر الذكاء الاصطناعي أصبحت قابلة للتعلم أيضًا بطريقة تمكنها من إصدار أحكام وقرارات مشابهة للبشر من خلال اتباع طريقة معينة في البرمجة تسمى تعلم

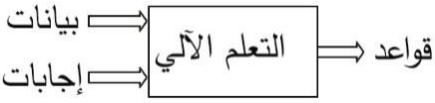
الآلة Machine Learning (ML) وهي مصطلح مرافق للذكاء الاصطناعي يُمكِّن الآلة من التعلم من مجموعات البيانات بتطبيق خوارزميات ونماذج وأنماط مسبقة البناء عليها نعطيها لها مع البيانات لتستنتج بذلك منها المعلومات دون أن تتم برمجتها وتعليمها بشكل صريح وبذلك تتعلم الآلة. إذاً تعلم الآلة أو ما يعرف اختصارًا MLهو مصطلح تقني يعني استخدام مجموعة من التقنيات والأدوات التي تساعد أجهزة الحاسوب والآلات الذكية عمومًا على التعلم والتكيف من تلقاء نفسها (Johnsen, 2024)

في أسلوب البرمجة الكلاسيكية، يُدخل البشر القواعد ضمن برامج حاسوبية ويحددون بدقة البيانات التي يجب معالجتها وفقًا لهذه القواعد كما في الشكل (3-9) ويكون الخرج إجابات محددة ناتجة عن عمليات المعالجة.



الشكل (3-9) البرمجة التقليدية (الباحث)

في نموذج البرمجة الجديد هذا الذي سمي تعلم الآلة أصبح بمقدور البشر إدخال البيانات للحاسوب بالإضافة إلى الإجابات المتوقعة وفقًا للبيانات كما في الشكل (3-10)، ويكون الخرج هو القواعد التي تم استنتاجها بشكل برنامج أو ما يُسمى بالنموذج Model ثم يمكن بعد ذلك تطبيق هذه القواعد على البيانات الجديدة لإنتاج الإجابات.



الشكل (3-10) التعلم الآلي (الباحث)

بمعنى آخر نظام التعلم الآلي يُدرّب بدلًا من برمجته، حيث تقدّم له العديد من عينات البيانات المتعلقة بالمهمة المطلوبة، ليكتشف بنية إحصائية في هذه العينات تربط المدخلات بالمخرجات (البيانات بالإجابات) وتسمح في النهاية للنظام بالوصول إلى قواعد لأتمتة المهمة. فعلى سبيل المثال إذا عرضنا على الحاسوب مجموعات ثنائية من الأرقام كالتالي (2,4) و (3,6) و (4,9) ثم طلبنا منه بناء على هذه المعطيات أن يتنبأ بالرقم الذي يجب أن يأتي مع الرقم 5 فسوف يحتاج في بداية الأمر إلى إيجاد المنطق بين مجموعات البيانات التي مررناها له وتطبيق نفس المنطق للتنبؤ بالرقم الجديد، عملية العثور على هذا المنطق يسمى تعلم الآلة وبعد إيجاد هذا المنطق سنتمكن من تطبيقه للتنبؤ بكل رقم جديد.

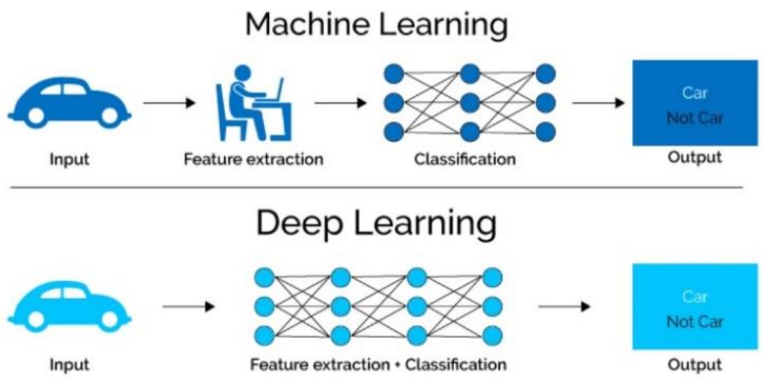
3-2-2 التعلم العميق

إن التعلم العميق Deep Learning (DL) نوع من الذكاء الصناعي وشكل خاص من أشكال تعلم الآلة كما في الشكل (3-11). يبدأ سير تعلم الآلة مع وجود "السمات ذات الصلة" التي يتم استخراجها يدويا من الصور.

ثم يتم استخدام السمات لإنشاء نموذج الذي يصنف الأشياء في الصورة. أما مع التعلم العميق، يتم استخراج السمات ذات الصلة تلقائيا من الصور. ويوضح الشكل (3-11) الفرق بين التعلم العميق وتعلم الآلة العادي.

يستخدم التعليم العميق هياكل خوارزمية محددة تُسمى الشبكات العصبونية، مصممة على غرار الدماغ البشري. تحاول طرائق التعلُّم العميق أتمتة المهام الأكثر تعقيدًا التي تتطلب عادةً الذكاء البشري (Seseri, 2-23).

يمكن تطبيق هذه التقنية على نوع معين من النماذج يسمى الشبكة العصبونية الاصطناعية، يتم فيها عملية تدريب النموذج - باستخدام الأمثلة المقدمة للنموذج أثناء التدريب. تكمن قوة التعلم العميق في أنه يمكن استخدامه بعد ذلك لإجراء تنبؤات أو تصنيفات على بيانات جديدة لم يسبق رؤيتها من قبل (Seseri, 2023).



الشكل (3-11) الفرق بين تعلم الآلة والتعلم العميق (Seseri, 2023, Source: Towards Data Science)

إن خوارزميات تعلم الآلة التقليدية ليست مفيدة أثناء العمل مع البيانات ذات الأبعاد العالية، حيث يكون لدينا عدد كبير من المدخلات والمخرجات. أما التعلم العميق، فهو ذو كفاءة عالية في هذه الحالات.

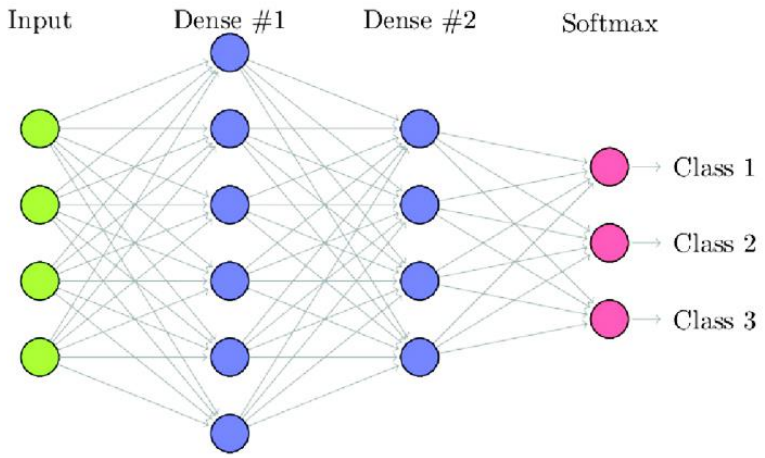
3-3 خوارزميات كشف الكائنات

لقد حظي اكتشاف الكائنات، باعتباره إحدى أكثر المشكلات الأساسية والتحديات في مجال الرؤية الحاسوبية، باهتمام كبير في السنوات الأخيرة. وعلى مدى العقدين الماضيين، شهدنا تطورًا تقنيًا سريعًا في اكتشاف الكائنات وتأثيره العميق على مجال الرؤية الحاسوبية بالكامل (Zou, 2023). الهدف من اكتشاف الكائنات هو اكتشاف جميع حالات الكائنات من فئة واحدة أو عدة فئات معروفة، مثل الأشخاص أو السيارات أو الوجوه في صورة ما. عادةً ما يوجد عدد صغير فقط من الكائنات في الصورة، ولكن هناك عدد كبير جدًا من المواقع والمقاييس المحتملة التي يمكن أن تحدث فيها والتي تحتاج إلى استكشاف بطريقة ما. يتم الإبلاغ عن كل اكتشاف ببعض أشكال معلومات الموضع. يمكن أن يكون هذا بسيطًا مثل موقع الكائن أو الموقع والمقياس أو المربع المحدد أو قناع التجزئة. في مواقف أخرى تكون معلومات الوضع أكثر تفصيلاً. على سبيل المثال، قد يحسب كاشف الوجه مواقع العينين والأنف والفم، بالإضافة إلى المربع المحدد للوجه. ومع الزيادة التدريجية في تطور خوارزميات التعلم العميق في مجال اكتشاف الكائنات، لوحظ تحسن كبير في أداء نماذج اكتشاف الكائنات (Kaur, 2023).

وتعتبر الشبكات العصبونية التلافيفية Convolutional Neural Network (CNN) رائدة في مجالات اكتشاف الكائنات، تجزئة الصور، تصنيف الصور، التعرف على الكلام، ومعالجة اللغة الطبيعية.

3-3-1 الشبكات العصبونية التلافيفية CNN

يتم تعريف الشبكة العصبونية Neural network على أنها نموذج للتعلم الآلي، يتخذ قرارات بطريقة تشبه الدماغ البشري، من خلال استخدام عمليات تحاكي الطريقة التي تعمل بها الخلايا العصبونية الحيوية معاً لتحديد الظواهر، ووزن الخيارات، والوصول إلى استنتاجات، فالشبكة العصبونية هي نموذج حوسبة مستوحى من نظام العصبونات في الدماغ البشري، وتُعتبر جزءًا أساسيًا من مجال الذكاء الاصطناعي وتهدف إلى تمثيل ومحاكاة عملية التعلّم والتعامل مع المعلومات بطريقة مُشابهة للدماغ البشري. تبنى الشبكة العصبونية من مجموعة من العناصر المعروفة باسم العقد (العصبونات) والتي ترتبط ببعضها عن طريق الروابط. وتتكوّن الشبكة العصبونيّة من طبقة إدخال Input layer، لاستقبال المعلومات، وطبقات مُتعددة من العقد الداخليّة Hidden layers، التي تقوم بمعالجة البيانات، وأخيرًا طبقة المُخرجات Output layer، التي تُعطي النتائج النهائيّة كما في الشكل (3-12). تعمل الشبكة العصبونية عن طريق ادخال المعلومات إلى العقد في الطبقة الأولى، ثم يتم تمرير هذه المعلومات من خلال العقد في الطبقات الداخليّة وفق قواعد مُحدّدة. يتم تعديل قوة الروابط بين العقد بناءً على الخبرة والتعلُّم من البيانات المتاحة، وهو الأمر الذي يُعرف بتدريب الشبكة العصبونية، حيث تتكيَّف تدريجيًا لأداء مُهمة مُحدَّدة من خلال تعديل الأوزان والقيم المُرتبطة بالعقد. تتميّز الشبكات العصبونية بقدرتها على التعامل مع البيانات المُعقّدة والتعلُّم من الأمثلة، مما يجعلها أداة فعّالة في حل المشكلات التي تكون صعبة للغاية بالنسبة للأنظمة الكلاسيكية.



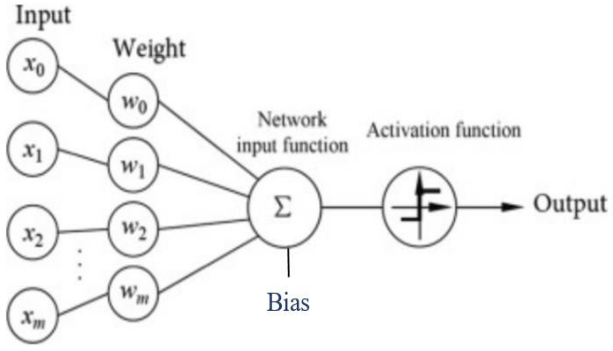
الشكل (3-12) بنية الشبكة العصبوية (Aabidi,2023)

يتم العمل في الشبكة العصبونية كما يلي:

- في الشكل (3-12) يتم تغذية المعلومات من طبقة الادخال Input لكل خلية عصبونية إلى الطبقة المخفية الأولى#Dense1، فترتبط المداخل كل على حدة بالخلايا العصبونية في الطبقة التالية من خلال القنوات.

- يتم تعيين قيمة عددية تُعرف بالوزن لكل قناة من هذه القنوات.
- يتم ضرب المدخلات بالأوزان المقابلة ثم يتم تغذية هذا المجموع المرجح كمدخل للطبقات المخفية.
- يتم تمرير الإخراج من الطبقات المخفية من خلال وظيفة التنشيط التي ستحدد ما إذا كان سيتم تنشيط الخلايا العصبونية المعينة أم لا.
- تنقل الخلايا العصبونية المنشطة البيانات إلى الطبقات المخفية التالية Dense#2 وهكذا بالنسبة للطبقات المخفية الأخرى Hidden= {Dense#1, Dense#2}. بهذه الطريقة، يتم نشر البيانات عبر الشبكة، وهذا ما يعرف بالانتشار الأمامي.
- في طبقة الإخراج Output، تتنبأ الخلية العصبونية ذات القيمة الأعلى بالإخراج. هذه المخرجات تكون على صيغة احتمالات لكل من الأصناف Class.
- تتم مقارنة المخرجات المتوقعة مع المخرجات الفعلية للحصول على الخطأ. ثم يتم نقل هذه المعلومات مرة أخرى عبر الشبكة، وتعرف العملية باسم Back propagation.
- بناءً على هذه المعلومات، يتم تعديل الأوزان. تتم هذه الدورة من الانتشار الأمامي والخلفي عدة مرات على مدخلات متعددة حتى تتنبأ الشبكة بالإخراج بشكل صحيح في معظم الحالات. هذا ينهي عملية تدريب الشبكة العصبونية. قد يزداد الوقت المستغرق لتدريب الشبكة العصبونية في بعض الحالات.

إن الشكل (3-13) يمثل Perceptron وهو أول نموذج لشبكة عصبونية اصطناعية يعتمد على نموذج العصبون (Liang, 2020)، حيث x_i هي مجموعة المدخلات، w_i هي الأوزان، b هي الانحياز، وActivation function هي أي وظيفة تنشيط.

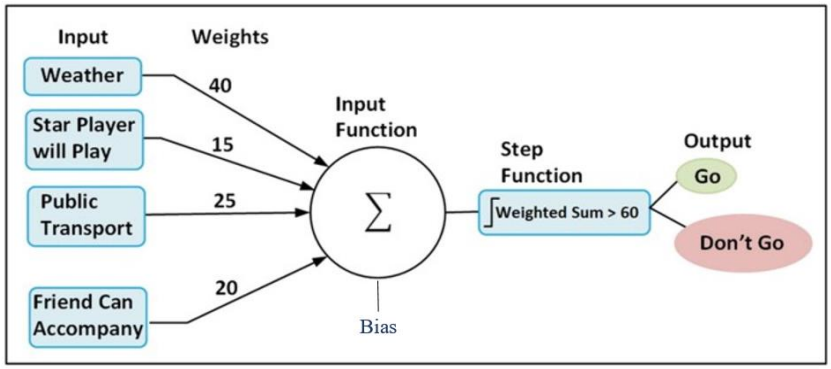


الشكل (3-13) نموذج أولي لشبكة عصبونية اصطناعية (Liang, 2020)

وتكون علاقة الخرج (3-2):

$$Output = b + (\sum_{i=0}^{m} w_i x_i).\, Activation\ function \qquad (2\text{-}3)$$

إن توابع التفعيل متعددة، وفي المثال التالي يستخدم تابع Step Function كتابع عتبة يفرز الخرج الناجح Go كخرج فوق 60، بحيث بعد التفعيل يعطي النتائج الثنائية Go/ Don't Go.



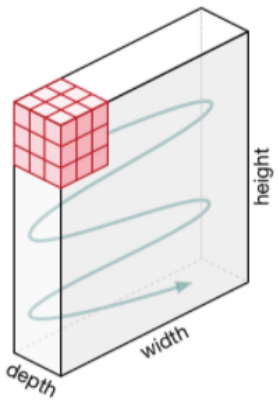
الشكل (3-14) مثال على تابع تفعيل (Ganguly, 2021)

ولمعالجة الصور بالشبكات العصبونية يجب:

- تقسيم بيانات الصور إى قسمين القسم الأول مجموعة تدريب يدرب النموذج عليها، والقسم الثاني هو مجموعة اختبار ويتم اختبار كفاءة النموذج عليها، ويجب أن تكون أعداد الصور للنموذج كبيرة وذلك للحصول على نتائج أفضل.
- أن تكون مجموعة بيانات الصور عالية الجودة للحصول على معلومات أكثر وضوحًا، ولكن لمعالجتها قد تحتاج إلى شبكات عصبونية أعمق.
- يتم تحويل صور RGB إلى درجات رمادية في بعض الخوارزميات قبل إدخالها في شبكة عصبونية.

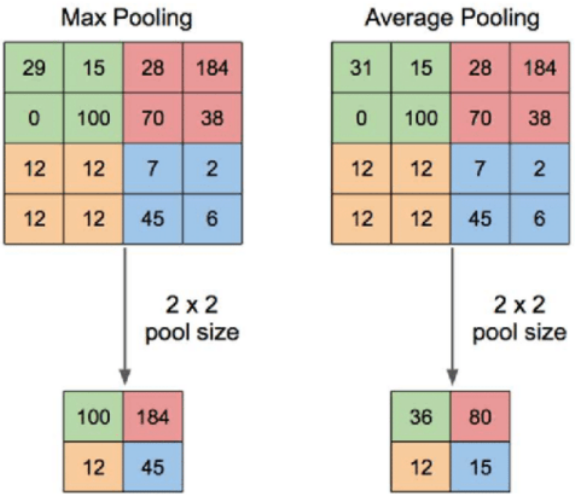
تتكون الشبكة العصبونية التلافيفية Convolution Neural network (CNN)، باختصار، من ثلاث طبقات:

- الطبقة التلافيفية (CONV): هي لبنة البناء الأساسية لـ CNN، وهي مسؤولة عن تنفيذ عملية الالتفاف، ويسمى العنصر المتضمن في تنفيذ عملية الالتفاف في هذه الطبقة النواة / المرشح (المصفوفة) وذلك كما في الشكل (3-15). تقوم النواة بإجراء تحولات أفقية ورأسية بناءً على معدل الخطوة حتى اتمام الصورة الكاملة (Manganiello, 2021).



الشكل (3-15) عملية الالتفاف في الشبكات العصبونية (Manganiello, 2021).

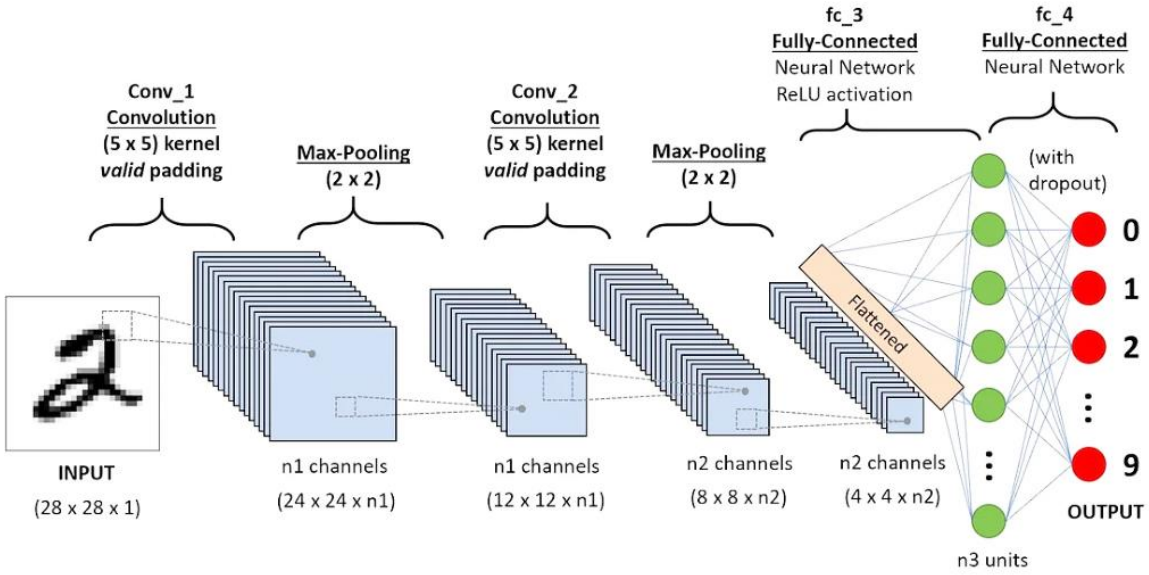
- طبقة التجميع POOL: هذه الطبقة مسؤولة عن تقليل الأبعاد. يساعد على تقليل الإمكانات الحاسوبية المطلوبة لمعالجة البيانات. هناك نوعان من أنواع التجميع كما في الشكل (3-16): Max Pooling أو Min Pooling و Average Pooling. يُرجع Max pooling القيمة القصوى من المنطقة التي تغطيها النواة على الصورة. يُرجع متوسط التجميع متوسط جميع القيم الموجودة في الصورة المغطى بالنواة.



الشكل (3-16) التجميع حسب العدد الأعظمي/ المتوسط (Yani, 2019)

- طبقة متصلة بالكامل Fully Connected Layer (FC): تأخذ هذه الطبقة شكل شعاع أو صف، حيث يتم توصيل كل مدخلات الصف بجميع الخلايا العصبونية الناتجة بالطبقة السابقة.

يبين الشكل (3-17) بنية شبكة عصبونية تلافيفية بكامل مراحلها، تقوم بكشف الأرقام



الشكل (3-17) شبكة عصبونية لكشف الأحرف (Saha, 2023)

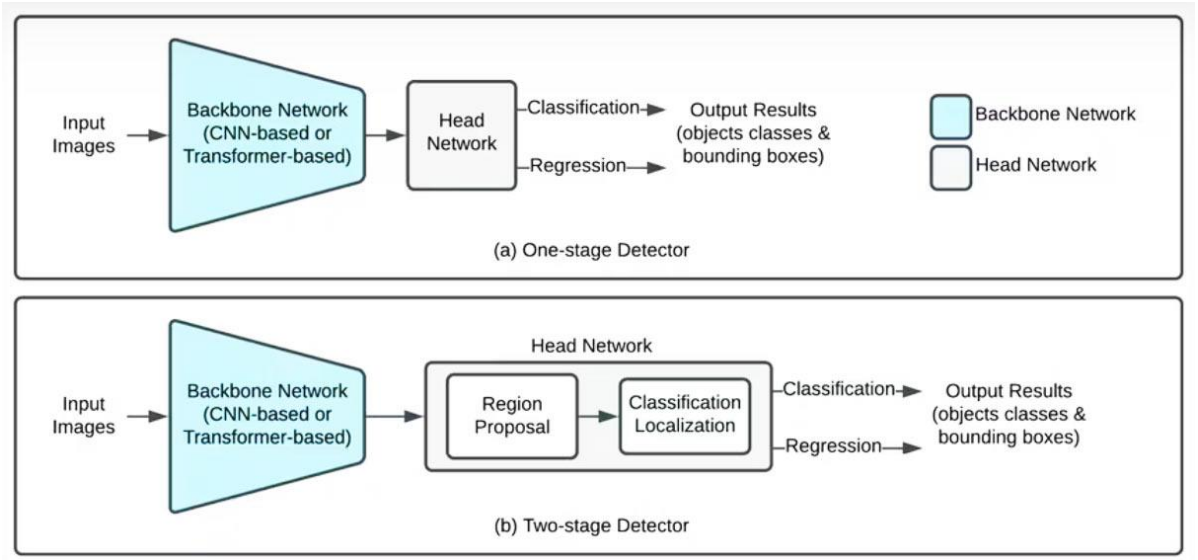
تُستخدم شبكات CNN على نطاق واسع في تصنيف الصور حيث يتم تمرير كل صورة عبر سلسلة الطبقات للحصول على قيمة احتمالية جيدة.

2-3-3 خوارزمية الكشف YOLO

هناك فئتان رئيسيتان من خوارزميات الكشف عن الكائنات: خوارزميات كشف مرحلة واحدة، وخوارزميات كشف بمرحلتين كما في الشكل (3-18). والطريقتان تقدمان نتائج جيدة. ومع ذلك، فإن كل منهما تقدم فوائد مختلفة، حيث تتمتع الأولى بوقت استدلال أقل والأخيرة بدقة أفضل (Encord1, 2025).

تعمل خوارزميات الكشف ذات المرحلة الواحدة بشكل أسرع لأنها تعتمد على شبكة واحدة لإنتاج المطلوب. تتخطى هذه النماذج خطوات استخراج السمات الوسيطة، مثل اقتراح المنطقة. تستخدم صورة الإدخال الخام لتحديد الكائنات وإنشاء إحداثيات مستطيل الكشف المحيط Bounding Box. من أفضل الأمثلة على خوارزميات الكشف ذات المرحلة الواحدة هي خوارزمية You only look once (YOLO). حيث تستطيع الوصول للمطلوب بمرور واحد للصورة.

تتألف أجهزة الكشف ذات المرحلتين، مثل Fast R-CNN، من شبكتين. الأولى هي شبكة اقتراح المنطقة RPN التي تحلل الصورة وتستخرج المناطق المحتملة للكائنات، والشبكة الثانية هي مستخرج سمات قائم على CNN يحلل المناطق المقترحة. يقوم الأخير بتحديد الكائنات الموجودة ويخرج إحداثيات المستطيل المحيط بها.



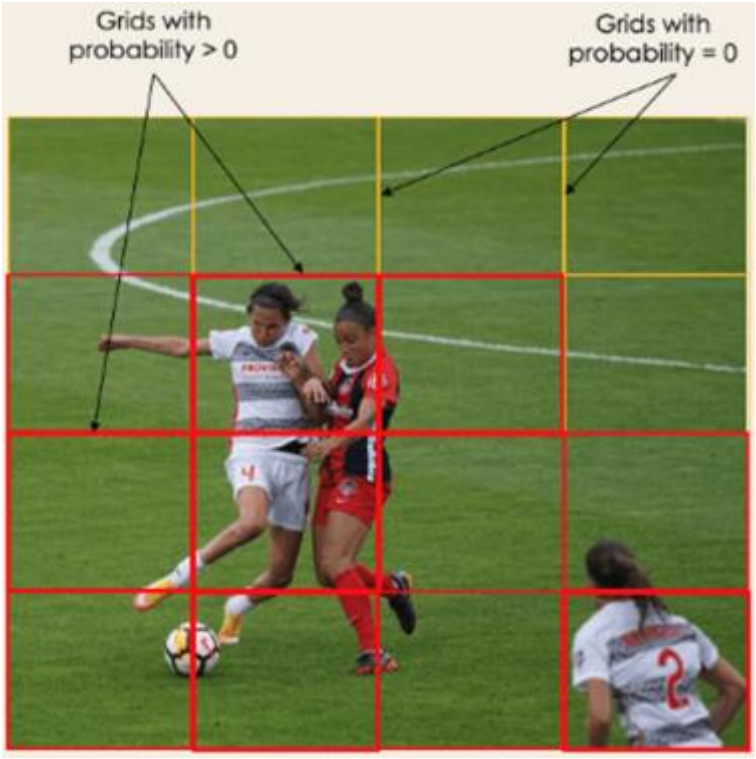
الشكل (3-18) خوارزميات الكشف بمرحلة واحدا وخوارزميات الكشف بمرحلتين (Kang.2025).

تحقق الخوارزميات المتطورة التي تستخدم الشبكة العصبونية التلافيفية (Convolutional Neural Network) CNN تقدماً كبيراً في مجال اكتشاف الكائنات، وذلك بسبب كفاءتها في تصنيف الأشياء، ولكن منذ تطوير الكاشف You Only Look Once (YOLO) (أنت تنظر مرة واحدة فقط) بدأ الباحثون ينظرون اليها كأداة كشف سريعة

جداً (Mishra, 2022) حيث استطاعت الكشف بمعدل 30 إطار في الثانية بمعامل mAP=57.9% على بيانات COCO الاختبارية. تعتبر YOLO حالياً أشهر أدوات الكشف عن الكائنات في الصور الرقمية وقد مرت بمراحل عديدة منذ نشأتها في عام 2015، بعدها نما YOLO وتطور بسرعة كأداة كشف سريعة للكائنات ليصل إلى الإصدار الأخير من YOLO-V8 في 2023. وهو يعتبر من أفضل أدوات الكشف (Vijayakumar, 2024)، (Tao, 2024)، (Shin, 2024) حيث يتميز بصفات الكشف في الزمن الحقيقي، عالي الجودة، وأداء عالٍ للتصنيف، على أساس معايير حسابية محدودة ولكنها فعالة.

الفكرة الأساسية وراء YOLO هي تقسيم الصورة المدخلة إلى عدة صور جزئية، والتنبؤ باحتمالية وجود كائن وإحداثيات المستطيل المحيط بالكائن لكل جزء. تعتمد الطريقة بالمجمل على الشبكات العصبونية التلافيفية CNN لاستخراج السمات من الصورة. يمكن تقسيم عملية YOLO إلى عدة خطوات (Kotai, 2022):

1. يتم تقسيم الصورة إلى عدة صور جزئية، وكل خلية مسؤولة عن التنبؤ بمجموعة من المستطيلات المحيطة واحتمالات الفئة كما في الشكل (3-19). تبدأ هذه الخطوة الأولى بتقسيم الصورة الأصلية إلى خلايا شبكية $N \times N$ متساوية الشكل، حيث N في حالتنا هو 4. كل خلية في الشبكة مسؤولة عن تحديد فئة الكائن الذي تغطيه والتنبؤ بها، بالإضافة إلى قيمة الاحتمالية/الثقة. مخرجات الشبكة عبارة عن مجموعة من مستطيلات الكشف المحيطة واحتمالات الفئة لكل جزء. المربعات المحاطة بالأحمر تمثل مربعات باحتمالية كائن داخلها.



الشكل (3-19) تقسيم الصورة إلى شبكة من الخلايا (Kotai, 2022)

2.الخطوة التالية هي تحديد المربعات المحيطة التي تتوافق مع المستطيلات التي تسلط الضوء على جميع الكائنات الموجودة في الصورة. تحدد YOLO سمات هذه المربعات المحيطة في العلاقة (3-4) للشعاع y ، (Kotai, 2022):

$$Y = [pc, bx, by, bh, bw, c1, c2] \quad (3\text{-}3)$$

حيث Y: هو تمثيل الشعاع النهائي لكل مربع محيط، وهذا مهم بشكل خاص خلال مرحلة التدريب.

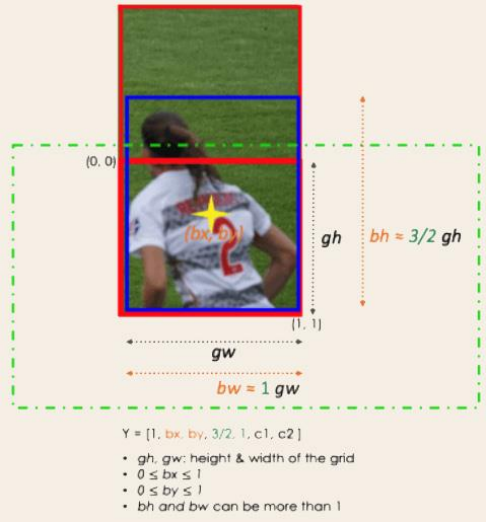
Pc: يتوافق مع درجة احتمال الشبكة التي تحتوي على كائن. إن جميع الشبكات باللون الأحمر سيكون لها درجة احتمالية أعلى من الصفر. أما باقي المربعات فإن احتمالية كل خلية صفراء هي صفر.

bx، by: هي إحداثيات x وy لمركز المربع المحيط.

bh، bw: يمثلان ارتفاع وعرض الصندوق المحيط.

c1 و c2: فئتي اللاعب والكرة. يمكننا الحصول على العديد من الفئات التي تتطلبها حالة الاستخدام.

يتم بعد ذلك تصفية المربعات المحيطة لإزالة المربعات المتداخلة واختيار المربع ذو الاحتمالية الأعلى كما في الشكل (3-20).



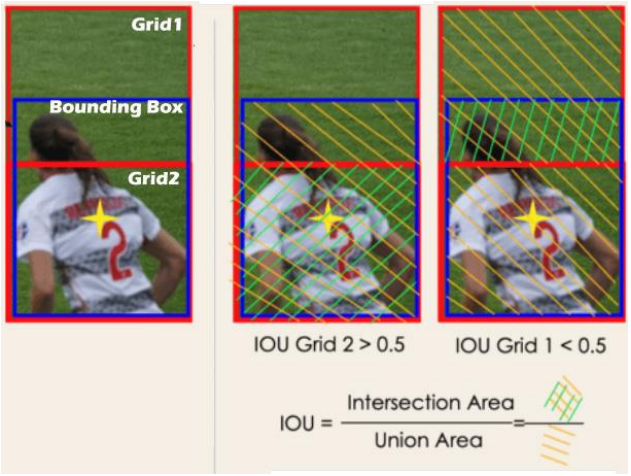
الشكل (3-20) معاملات المتجه Y (Kotai, 2022)

3. حساب نسبة التقاطع على الاجتماع Intersection Over Unions (IOU)

في معظم الأحيان، يمكن أن يحتوي كائن واحد في صورة ما على عدة مربعات شبكية مرشحة للتنبؤ، على الرغم من أنها ليست جميعها ذات صلة. الهدف من IOUϵ[0,1] هو التخلص من مربعات الشبكة هذه للاحتفاظ بالمربعات ذات الصلة فقط. يحدد المستخدم حد اختيار IOU الخاص به، والذي يمكن أن يكون على سبيل المثال 0.5، ثم يقوم YOLO بحساب IOU لكل خلية شبكة وهي منطقة التقاطع

مقسومة على منطقة الاتحاد. وأخيرًا، فإنه يتجاهل التنبؤ بخلايا الشبكة التي لها عتبة IOU<=0.5 ويأخذ في الاعتبار الخلايا التي لها عتبة IOU>0.5.

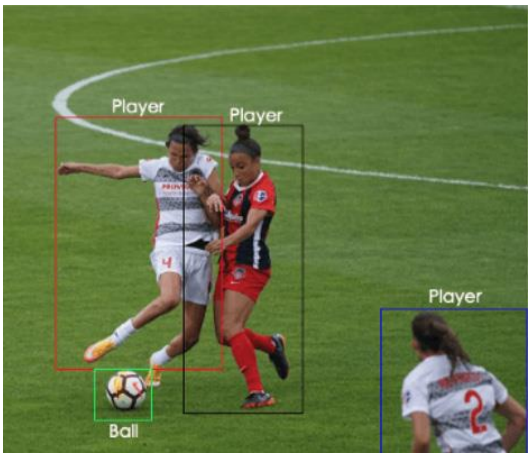
يوجد في الشكل (3-21) رسم توضيحي لتطبيق عملية تحديد الشبكة على الكائن، يمكننا أن نلاحظ أن الكائن كان يحتوي في الأصل على مرشحين للشبكة، تم اختيار "الشبكة 2 فقط في النهاية.



الشكل (3-21) اختيار الشبكة ذات النسبة الأعلى لـ IOU (Kotai, 2022)

بتكرار العمل فإن الناتج النهائي عبارة عن مجموعة من مستطيلات الكشف المحيطة المتوقعة وتسميات الفئات لكل كائن في الصورة كما في الشكل (3-22). كما ذكر سابقاً فإن IOU تقيس مدى تشابه المربعات المحيطة، لذلك يمكن تعيين عتبة لتجاهل جميع المربعات المحيطة المتوقعة التي تكون قيمة IOU الخاصة بها منخفضة جدًا، سيكون هذا هو عتبة IOU.

في YOLO معامل آخر مهم وهو عتبة الثقة، حيث درجة الثقة هي القيمة التي تمثل ثقة النموذج في تنبؤاته. إذا قمنا بتعيين حد لهذا المعامل، يمكننا إزالة جميع التنبؤات التي لم يكن النموذج متأكدًا منها.



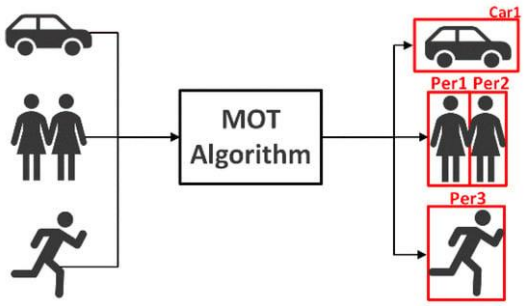
الشكل (3-22) الناتج النهائي لخوارزمية YOLO (Kotai, 2022)

إحدى المزايا الرئيسية لـ YOLO هي أنه يعالج الصورة بأكملها في مسار واحد، مما يجعلها أسرع وأكثر كفاءة من أدوات كشف الكائنات ذات المرحلتين.

4-3 خوارزميات ملاحقة الكائنات المتعددة MOT

تعد ملاحقة الكائنات المتعددة Multiple Object Tracking (MOT) واحدة من أهم تطبيقات الرؤية الحاسوبية والتي تهدف بشكل عام إلى ملاحقة الكائنات منذ بداية ظهورها في الفيديو أو على عدسة الكاميرا لحين اختفائها وبالتالي إمكانية تسجيل معلومات حول هذه الكائنات ومعرفة سلوكها (Bashar, 2022)، ولهذه الخوارزمية تطبيقات كثيرة ومتنوعة منها المراقبة الأمنية والسيارة ذاتية القيادة والتصوير الطبي.

من الناحية التقنية تبدأ عملية ملاحقة الكائنات باكتشافها أولاً، أي تحديد موضع هذه الكائنات في الصورة عن طريق وضع مستطيلات محيطة بها Bounding box وتحديد نوعها (Car, Person, ….) كما في الشكل (3-23)، ثم تقوم خوارزمية الملاحقة بتعيين رقم معرف ID خاص بكل كائن ضمن النوع الواحد (1,2, …)، لتبدأ عندها عملية الملاحقة لكل كائن تم تعيين رقم معرف له (Bhuiyan, 2023).



الشكل (3-23) مخرجات MOT. رقم يحدد كائنًا معينًا في الفيديو (Bhuiyan, 2023).

يتطلب إجراء عملية ملاحقة الكائنات في البداية كما ذكرنا سابقاً تحديد موضع الكائنات في الصورة لذلك يجب استخدام خوارزمية الكشف عن الكائنات Object Detection بشكل مرافق لخوارزمية ملاحقة الكائنات.

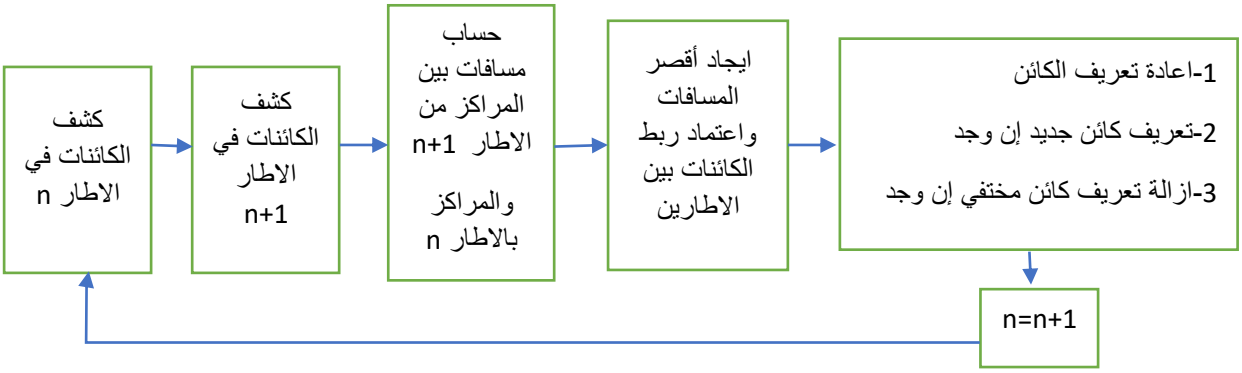
بعد كشف الكائن تبدأ عملية الملاحقة له طالما هو في حالة ظهور وتنتهي هذه العملية عند اختفاء الكائن.

هناك نوعان رئيسيان لخوارزميات الملاحقة:

1. ملاحقة الكائنات في مقطع فيديو: يتم تطبيق خوارزمية ملاحقة الكائنات في هذا النوع على فيديو مسجل مسبقاً حيث يتم اكتشاف الكائنات الموجودة في إطارات الفيديو وتتبع انتقالها عبر هذه الإطارات.
2. ملاحقة الكائنات في الزمن الحقيقي: يتم اكتشاف الكائنات وتتبعها مباشرة وفي الزمن الحقيقي عن طريق معالجة الإطارات المتدفقة من كاميرا المراقبة، حيث هذا النوع يمثل تحدياً أكبر فهو يتطلب موارد حاسوبية قوية تساعد على المعالجة في الزمن الحقيقي دون حدوث تأخير زمني.

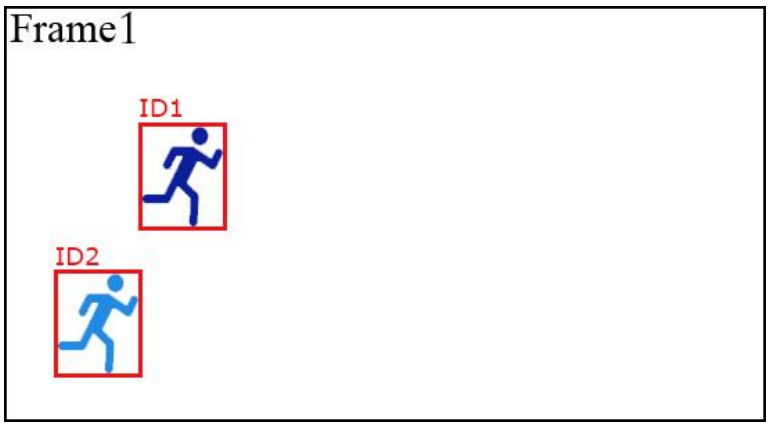
1-4-3 المبدأ العام لعمل خوارزمية ملاحقة الكائنات MOT

بفرض لدينا مقطع فيديو مسجل، ونريد تطبيق خوارزمية الملاحقة عليه. يرينا الشكل (3-24) مخطط مسلكي مبسط لخوارزمية كشف الكائنات المتعددة



الشكل (3-24) مخطط مسلكي مبسط لخوارزمية كشف الكائنات المتعددة (الباحث)

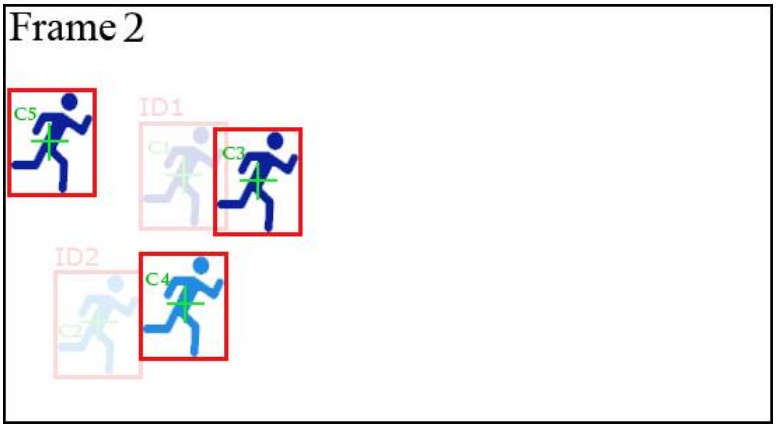
في البداية تتم قراءة الإطار Frame الأول من الفيديو ويطبق عليه خوارزمية الكشف عن الكائنات وحساب مركز كل كائن أي مركز مستطيل الإحاطة Bounding box ثم يتم إعطاء معرف ID خاص بكل كائن، الآن بالنظر إلى الشكل (3-25) يتوضح لدينا وجود كائنين في الإطار الأول الكائن ذو المعرف 1 و الكائن ذو المعرف 2.



الشكل (3-25) الكشف وتعريف الكائنات (الباحث)

ان كشف الكائن الهدف يمكن من ايجاد احداثيات مركزه C.

بعد التقاط الإطار الثاني تنفذ خوارزمية الكشف عن الكائنات من جديد كما في الشكل (3-26)

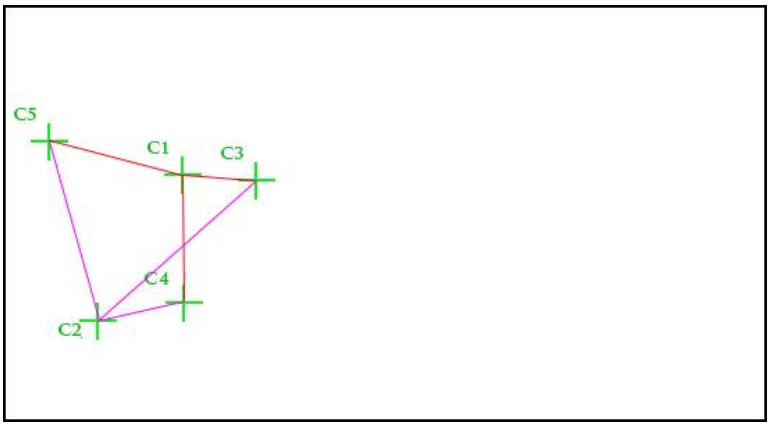


الشكل (3-26) الكشف في الإطار اللاحق (الباحث)

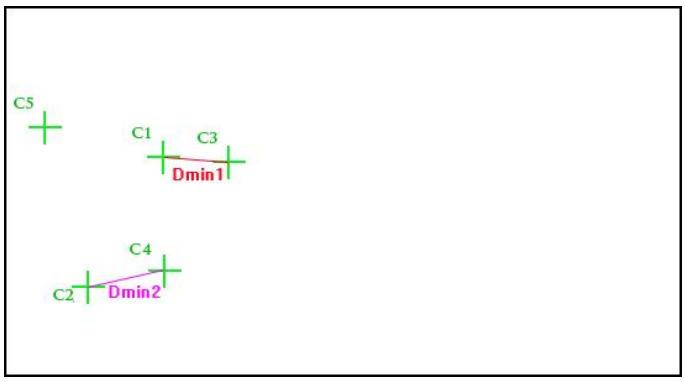
في هذا الإطار لابدّ أن نتذكر بأن الكائنات تتحرك خلال الفيديو بالتالي تتغير مواضع هذه الكائنات بالنسبة للإطار السابق. بالنظر إلى الشكل (3-26) نلاحظ أن النقاط C1, C2 تمثل المواضع القديمة (المراكز) للكائنات في الإطار الأول، وقد مثلت الكائنات القديمة بالرسوم الفاتحة، أما النقاط C3, C4 تمثل المواضع الجديدة (المراكز) للكائنات في الإطار الثاني. كذلك يلاحظ ظهور كائن جديد تم كشفه بالموضع C5.

نلاحظ أن الكائنات الثلاثة الموجودة في الإطار الثاني هي جديدة، لذلك يجب أن تتم عملية كشف الترابط بين كائنات الإطار الثاني مع كائنات الإطار الأول كما يلي:

- يتم حساب المسافات بين موضع كل كائن من الإطار الأول مع مواضع الكائنات في الإطار الثاني، فبالنسبة للكائن ذو المعرّف 1 في الإطار الأول يتم حساب المسافات بينه وبين الكائنات الثلاثة المكتشفة في الإطار الثاني (المسافات الحمراء) وكذلك بالنسبة للكائن ذو المعرف 2 في الإطار الأول يتم حساب المسافات بينه وبين الكائنات الثلاثة المكتشفة في الإطار الثاني (المسافات البنفسجية) الشكل (3-27).
- البحث عن أصغر مسافة بين كل كائن من الكائنات الموجودة في الإطار الأول مع الكائنات الموجودة في الإطار الثاني. كما هو موضح في الشكل (3-28).



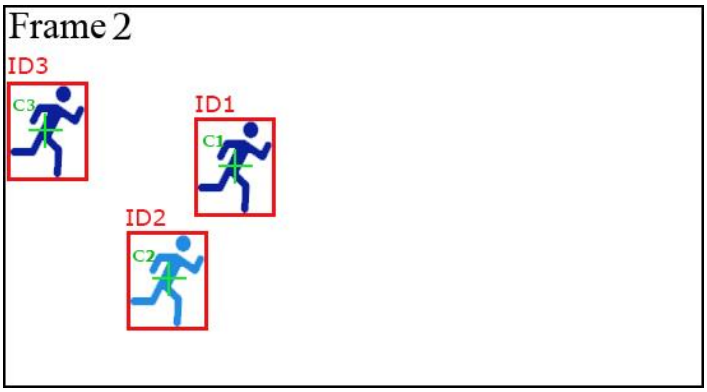
الشكل (3-27) البحث عن ترابط الكائنات (الباحث)



الشكل (3-28) تحديد الكائنات المترابطة (الباحث)

بالنسبة للكائن ذو المعرّف 1 المسافة بينه وبين الكائن ذو المركز C3 أصغر ما يمكن Dmin1 بالتالي يعتبر الكائن ذو المركز C3 مرتبط به أي هو نفسه ويتم تعريفه وإعطاؤه نفس المعرف أي .1 وبالنسبة للكائن ذو المعرّف 2 المسافة بينه وبين الكائن ذو المركز C4 أصغر ما يمكن Dmin2 بالتالي يعتبر الكائن ذو المركز C4 مرتبط به أي هو نفسه ويتم تعريفه وإعطاؤه نفس المعرف أي 2.

- بالنسبة للكائن المتبقي ذو المركز C5 لم يتم مطابقته مع كائنات الإطار الأول بالتالي من المؤكد أنه كائن جديد ظهر بالإطار الثاني بالتالي يتم تعريفه وإعطاؤه معرف جديد هو 3، ويتم إضافته الى قائمة الكائنات المُلاحقة كما في الشكل (3-29). بعد ذلك يتم تكرار الخطوات السابقة أجل كل إطار.



الشكل (3-29) التعريف في الإطار اللاحق (الباحث)

- من أجل كل كائن ملاحق من الممكن عند إطار معين أن يختفي ظهوره، بالتالي في هذه الحالة يجب حذف هذا الكائن من قائمة الكائنات المُلاحقة.

يبين الشرح السابق طريقة الملاحقة بشكل مبسط وخاصة فيما يتعلق بظهور كائن جديد، ولكن عملياً يجب عدم التثبيت والاضافة للقائمة حتى التأكد بمرور الوقت من أن هذا الكائن جديد، يعطى هذا الوقت بعدد محدد من الإطارات هي العتبة N1 من خلالها يمكن تحديد إن كان الكائن قد تم التثبت من ظهوره أم لا. كذلك الأمر إذا اختفى الكائن ننتظر لمرور N2 إطار، وإذا لم يظهر الكائن في إطارات الفيديو بالتالي يتم إلغاء ملاحقته وحذفه.

يجب عند الملاحقة تحديد بعض المعاملات المهمة:

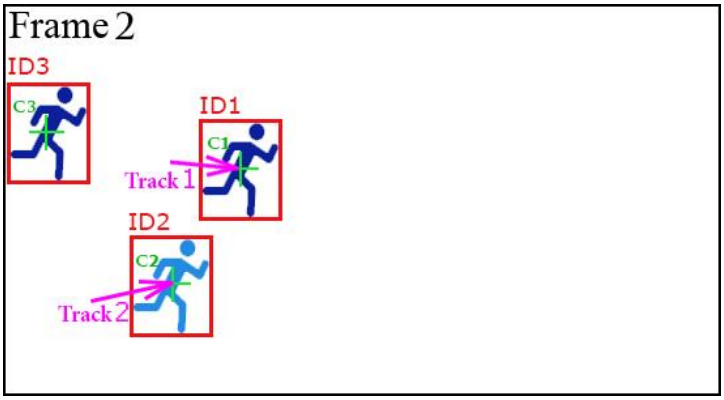
معامل الاختفاء: هو عدد الإطارات الأعظمي الذي يعتبر الكائن بعده مختفياً. إن اختفاء الكائن يمكن أن يحدث بالحالات التالية:

- خروج الكائن المتتبع من المشهد.
- وجود ضجيج مؤثر على عملية الكشف.
- تداخل الكائنات وهو ما نركز عليه في هذا البحث.

معامل المسافة الأعظمية: Max Disappear: أقصى مسافة مستخدمة للربط بين الكائنات أثناء عملية الملاحقة.

وهو معامل مهم جداً، حيث يؤثر على اتخاذ القرار بالربط، ويجب الموازنة بين أن يكون كبيرا لدرجة أنه لا يؤثر على عملية الملاحقة بالسرعة الكبيرة للكائن، أو كونه صغيراً بحيث لا يرتبط كائن بكائن مجاور.

إذاً تقوم خوارزمية MOT تلقائيًا بتحديد عدد الكائنات في كل إطار من الفيديو وربطها بالمسارات الموجودة إذا كانت تنتمي إلى نفس الكائن؛ وإلا، فسيتم إنشاء مسارات جديدة للكائنات المتبقية التي لا يمكن مطابقتها مع أي من المسارات الموجودة، الشكل (3-30). وتعد الأساليب القائمة على التعلم العميق هي السائدة حاليًا.



الشكل (3-30) تحديد كائن جديد في الملاحقة (الباحث)

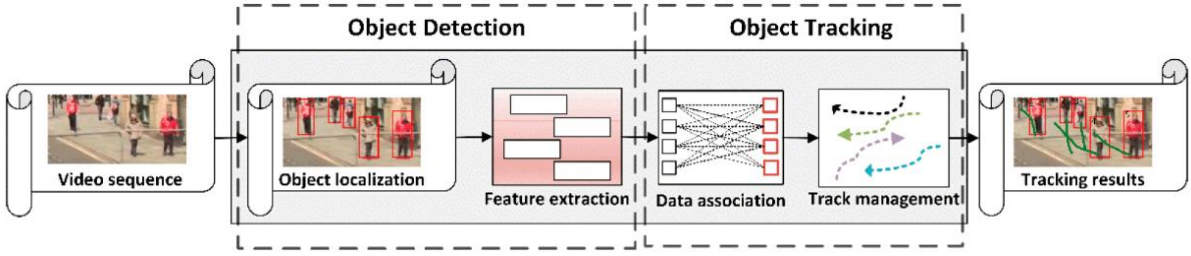
3-4-2 أنواع خوارزميات ملاحقة الكائنات المعتمدة على التعلم العميق

تقسم خوارزمية ملاحقة الكائنات المعتمدة على التعلم العميق إلى ثلاثة أنواع (Ma, 2024) (Bhuiyan, 2023) وذلك حسب بنيتها:

- الملاحقة بالكشف، Tracking By Detection (TBD).
- الكشف والملاحقة المشترك، Joint Detection and Tracking (JDT).
- الملاحقة المعتمدة على المحولات، Transformers.

3-4-2-1 التعلم العميق القائم على خوارزميات TBD

نظرًا لقدرات التعلم العميق القوية في تمثيل السمات، فقد حققت خوارزميات اكتشاف الكائنات مؤخرًا أداءً أفضل. فظهرت تقنية الملاحقة القائمة على الكشف بسرعة كمعيار للتتبع التلقائي للأشياء المتعددة. لإكمال مهمة ملاحقة الكائنات، تتعرف خوارزمية MOT القائمة على TBD أولاً على الكائن في كل إطار من تسلسل الفيديو، وتستخرج سمات الكائن، ثم تربطها وفقًا لتطابق بيانات السمات (Guo, 2022). كما في الشكل (3-31).



الشكل (3-31) تمثيل لخوارزمية TBD (Guo, 2022)

تستخدم كثير من خوارزميات الملاحقة كاشف YOLO لسماته الجيدة، وقد تمكنت YOLO الحديثة من إيجاد حل وسط بين دقة اكتشاف الكائنات والسرعة ، مما مكن من استخدامها كوحدة كشف لملاحقة الكائنات المتعددة.

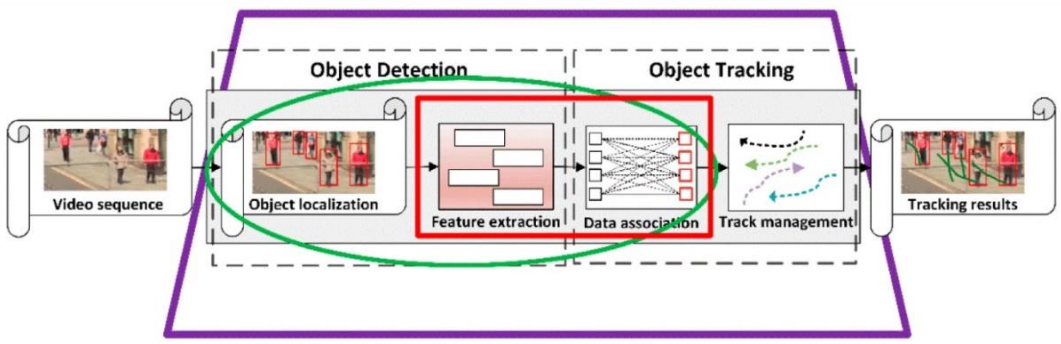
3-4-2-2 التعلم العميق القائم على خوارزميات JDT

بشكل عام، سيذهب التطوير هنا في أحد ثلاث اتجاهات:

أولاً، سيتم تحويل شبكة الكشف ودمجها في مهمة الملاحقة بحيث يمكن لنموذج الشبكة المصمم تعلم احتمالية ارتباط الكائن بين إطارات التسلسل.

ثانيًا، سيتم تنفيذ دمج الوحدات الفرعية في خوارزمية اكتشاف الكائنات أو دمج السمات لتحقيق مهام ملاحقة الكائنات المتعددة.

وثالثًا، سيكون التركيز على دمج الخوارزميات الممتازة في مجال ملاحقة الكائن الفردي. يتم اختيار عدداً قليل من الوحدات الفرعية للدمج، والشكل (3-32) لتوضيح كيفية عمل طريقة JDT.

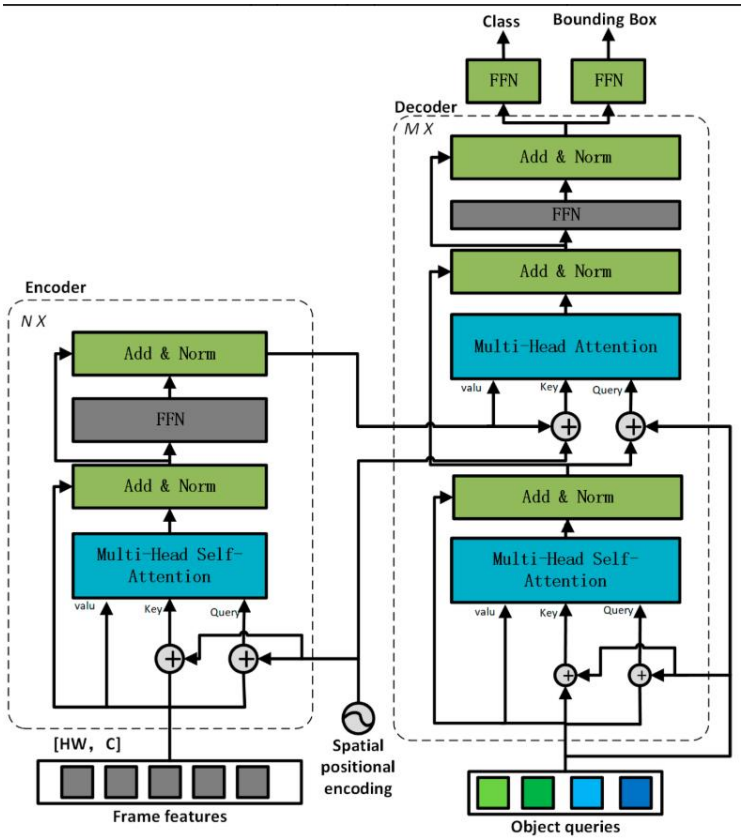


الشكل (3-32) تمثيل لخوارزمية JDT (Guo, 2022)

تعتمد طريقة JDT على دمج الوحدات الفرعية TBD. يحيط شبه المنحرف الأرجواني بوحدة التعرف على الكائنات وتتبع الكائنات ويقوم بدمج الوحدات، وهو النوع الأول من بنية خوارزمية JDT. يحيط المنحنى المستطيل الأحمر بوحدات استخراج السمات وربط البيانات ويدمجها، وهو النوع الثاني من بنية خوارزمية JDT. يحيط المنحنى الإهليلجي الأخضر بوحدات تحديد موقع الكائنات واستخراج السمات وربط البيانات ويدمجها. هذا هو النوع الثالث من إطار عمل خوارزمية JDT.

3-2-4-3 خوارزمية MOT القائمة على المحول Transformer

لقد حققت النماذج القائمة على المحولات تقدمًا كبيرًا في معالجة اللغة الطبيعية Natural (NLP) Language Processing. تم تقديم طرائق الانتباه لأول مرة للتطبيق على الترجمة الآلية، وحققت أداء قوي بمعالجة اللغة الطبيعية (Devi, 2023). إن استخدام الأبحاث الحديثة المحول في مجال الرؤية الحاسوبية، مستوحاة من النجاح الهائل لتصميم المحول. وعلى النقيض من CNN، فإن آلية الانتباه الذاتي للمحول ليست مقيدة بالتفاعلات المحلية؛ بدلاً من ذلك، يمكنه تجميع البيانات من جميع المدخلات وتحقيق التوازي السريع. كما تجعل الذاكرة المثالية للمحول وقوة الحوسبة منه أكثر ملاءمة لمعالجة التسلسلات الطويلة. يبين الشكل (3-33) بنية خوارزمية المحول.



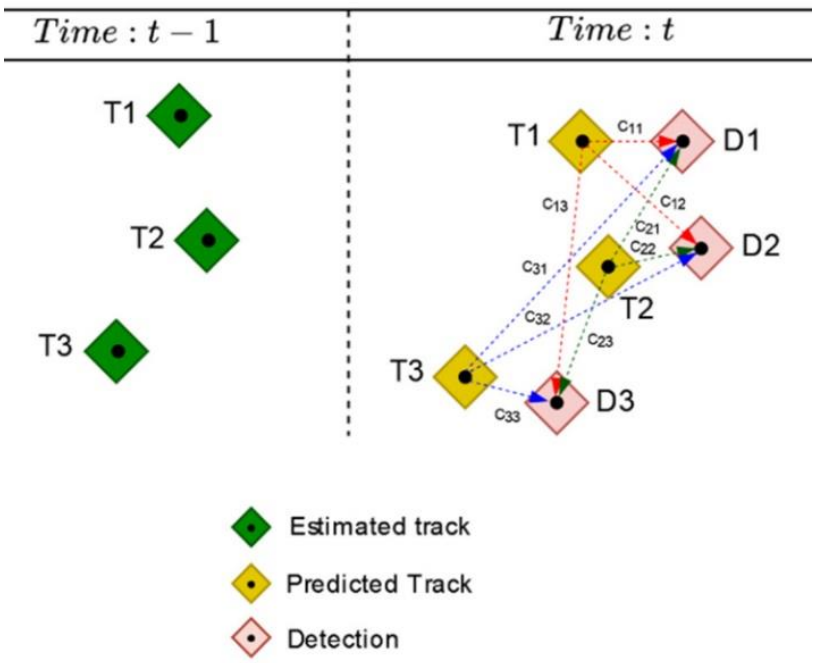
الشكل (3-33) بنية خوارزمية المحول (Devi, 2023).

تعتبر التكلفة الحاسوبية بالعتاد والزمن من اهم المآخذ على استخدام المحولات، لكن مع التطور التكنولوجي السريع بعتاديات الحاسب من الراجح أن يسود استخدامها في المستقبل بمختلف التطبيقات.

5-3 نماذج الحركة

يسمح لنا MOT بملاحقة حركات ومسارات كائنات متعددة بمرور الوقت وربطها بهوياتها الخاصة. يمكن تحليل هذه المعلومات للحصول على رؤى حول سلوكياتهم، والتنبؤ بحركاتهم المستقبلية، وتحديد الأنماط في تحركاتهم. يمكن أن يوفر دمج MOT مع اكتشاف الكائنات نتائج أكثر دقة وموثوقية مقارنة باكتشاف الكائنات بمفردها، وخاصة في السيناريوهات المزدحمة والمعقدة حيث قد تحجب الكائنات بعضها البعض، أو تتحرك داخل وخارج الإطار، أو تغير مظهرها. من خلال دمج MOT مع اكتشاف الكائنات، يمكننا الحصول على فهم أكثر شمولاً للمشهد وتحسين أداء مهام الرؤية الحاسوبية المختلفة. تتشابه المكونات المشاركة في ملاحقة الكائنات المتعددة MOT مع تلك الموجودة في ملاحقة الكائن الفردي، بما في ذلك الملاحظة والتنبؤ والتقدير. ومع ذلك، فإن MOT تقدم مشكلة أكثر تعقيداً، لأنها تنطوي على تتبع العديد من الأجسام في وقت واحد، وتتطلب ربط المسارات الموجودة بشكل صحيح بالاكتشافات الجديدة المقابلة لها.

بعض خوارزميات MOT تستخدم نماذج الحركة من أجل التنبؤ وتقدير الحالة التالية للكائن قيد الملاحقة وذلك للمساهمة بقدر أكبر في معالجة العديد من الصعوبات سواءً بالملاحقة العادية أو بالملاحقة أثناء وجود حالات الاحتجاب والاختفاء، مما يوفر بتكاليف الخوارزمية بشكل عام (Sadli, 2023). يبين الشكل (3-34) ثلاثة كائنات ملاحقة وثلاثة مسارات مقدرة في الوقت $t-1$، بالإضافة إلى ثلاث اكتشافات جديدة وثلاثة مسارات متوقعة في الوقت t. ويقوم نموذج الحركة باستخدام هذه العناصر لتقدير الحالة اللاحقة.



الشكل (3-34) تقدير الحالة الجديدة للكائن أثناء عملية الملاحقة (Sadli, 2023)

1-5-3 النموذج الديناميكي

عند وجود هدف متحرك على خط مستقيم، يمكن حساب موضع الهدف المستقبلي بسهولة باستخدام معادلة الحركة (4-3)، (Becker, 2023):

$$x = \frac{1}{2}at^2 + v_0t + x_0 \qquad (4\text{-}3)$$

حيث x فاصلة المتحرك، t الزمن، v_0 السرعة الابتدائية، a التسارع، x_0 الفاصلة الابتدائية

ومن أجل حركة بالفراغ تعطى بالعلاقات (5-3):

$$x = \frac{1}{2}a_xt^2 + v_{x0}t + x_0$$

$$y = \frac{1}{2}a_yt^2 + v_{y0}t + y_0 \qquad (5\text{-}3)$$

$$z = \frac{1}{2}a_zt^2 + v_{z0}t + z_0$$

تسمى: $x, y, z, v_{x0}, v_{y0}, v_{z0}, a_x, a_y, a_z$ بحالة النظام المدروس وتستخدم x, y, z الحالية كمداخل لخوارزمية التنبؤ، بينما خرج الخوارزمية هو الحالة المستقبلية $x0, y0, z0$.

تعرف المعادلات المعطية أعلاه بالنموذج الديناميكي، وإذا كانت حالة الهدف والنموذج الديناميكي معروفين، فيمكن إنجاز التنبؤ بحالة الهدف اللاحقة المتوقعة بسهولة (Becker, 2023).

في الواقع، لا تكون القياسات دقيقة تمامًا. فهي تحتوي على أخطاء عشوائية أو عدم يقين يمكن أن يؤثر على دقة حالة الهدف المتوقعة. يعتمد حجم الأخطاء على عوامل مختلفة، مثل معايرة نظام القياس ونسبة الإشارة إلى الضجيج. بالإضافة إلى ذلك، لا تتماشى حركة الهدف دائمًا مع معادلات الحركة بسبب عدة عوامل. يؤدي هذا الاختلال بين معادلات الحركة وحركة الهدف الفعلية إلى خطأ أو عدم يقين في النموذج الديناميكي.

لتحسين دقة الملاحقة، من الضروري استخدام خوارزمية تنبؤ تأخذ في الاعتبار عدم اليقين في العملية والقياس.

من أجل متحرك بسرعة منتظمة يمكن كتابة العلاقة (6-3)، (Becker, 2023):

$$\dot{x} = v = \frac{\Delta x}{\Delta t} \qquad (6\text{-}3)$$

يمكن كتابة علاقة الفاصلة والسرعة في اللحظة التالية n+1 كما في المعادلات (3-7):

$$x_{n+1} = x_n + \Delta t . \dot{x}_n \quad (7\text{-}3)$$

$$\dot{x}_{n+1} = \dot{x}_n \quad (8\text{-}3)$$

يُطلق على المعادلات (3-7) (3-8) أعلاه اسم معادلات استقراء الحالة State Extrapolation Equation "وتسمى أيضًا معادلة الانتقال Transition Equation أو معادلة التنبؤ Prediction Equation". تقوم المعادلات هنا باستقراء الحالة التالية (التنبؤ). وهي تعتمد على ديناميكيات النظام وتختلف من مثال إلى آخر.

2-5-3 مرشح $\alpha - \beta$

بالاستناد إلى معادلات الاستقراء يمكن توقع الموضع عند اللحظة n بالعلاقة (3-9)، (Becker, 2023):

$$\hat{x}_{n,n-1} = \hat{x}_{n-1,n-1} + \Delta t . \hat{\dot{x}}_{n-1,n-1} \quad (9\text{-}3)$$

وبالنسبة لتوقع السرعة بالمعادلة (3-10):

$$\hat{\dot{x}}_{n,n-1} = \hat{\dot{x}}_{n-1,n-1} \quad (10\text{-}3)$$

بفرض أن القيمة المقاسة لـ x هي z_n فتكون معادلة تحديث الحالة للسرعة (3-11):

$$\hat{\dot{x}}_{n,n} = \hat{\dot{x}}_{n,n-1} + \beta \left(\frac{z_n - \hat{x}_{n,n-1}}{\Delta t} \right) \quad (11\text{-}3)$$

تعتمد قيمة β على دقة القياس، واختيارها سيؤثر على عدد عمليات القياس والتحديث للوصول إلى الحالة المقبولة.

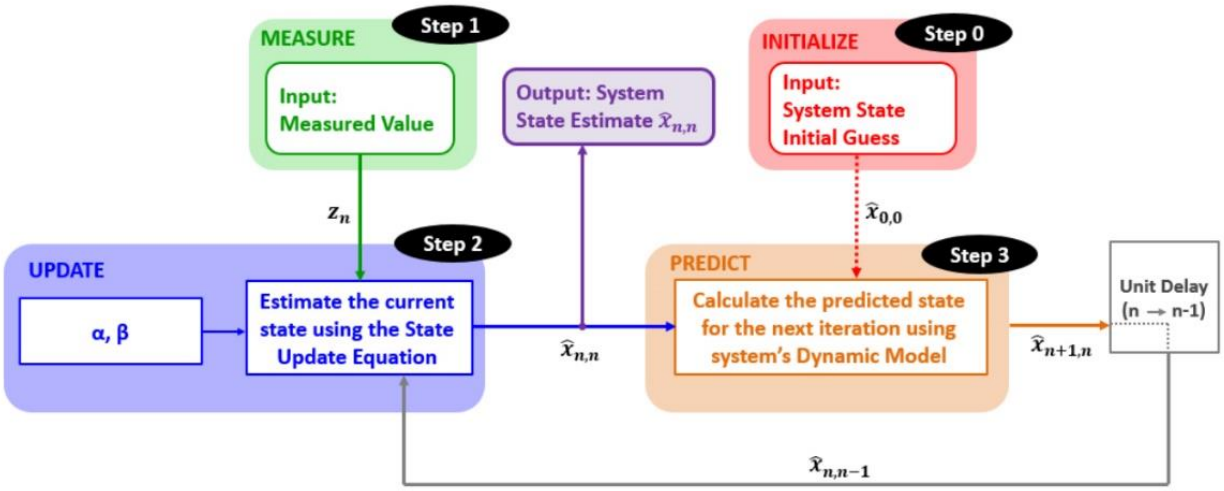
أما بالنسبة لتوقع الفاصلة فيمكن الكتابة بطريقة مماثلة ونحصل على المعادلة (3-12):

$$\hat{x}_{n,n} = \hat{x}_{n,n-1} + \alpha (z_n - \hat{x}_{n,n-1}) \quad (12\text{-}3)$$

الشكل (3-35) تقدير الحالة (Becker, 2023).

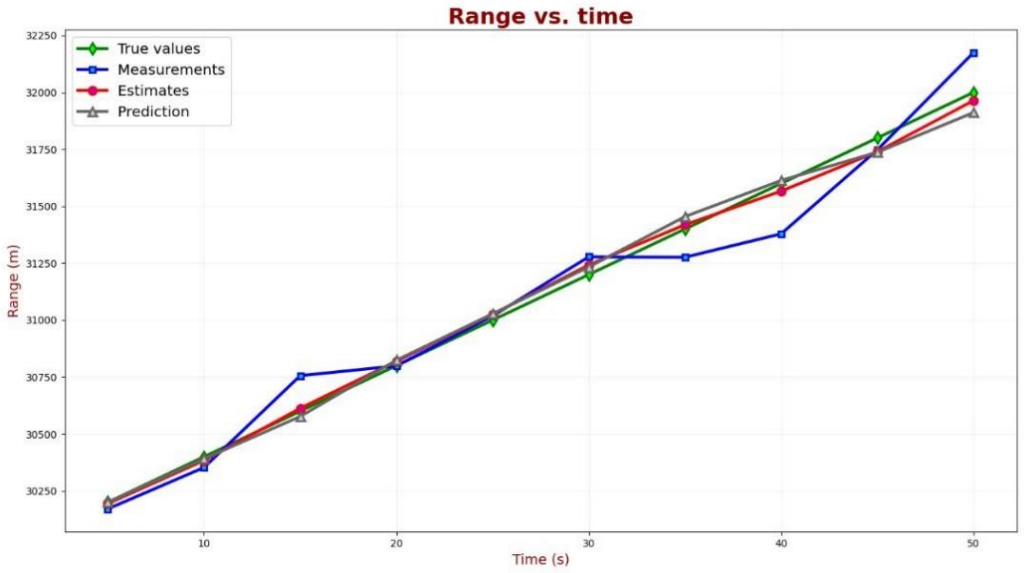
تمثل المعادلات السابقة معادلات تحديث الحالة للمرشح $\alpha - \beta$ State Update Equation

يوضح الشكل (3-36) خوارزمية التقدير المستخدمة هنا



الشكل (3-36) خوارزمية التقدير (Becker, 2023).

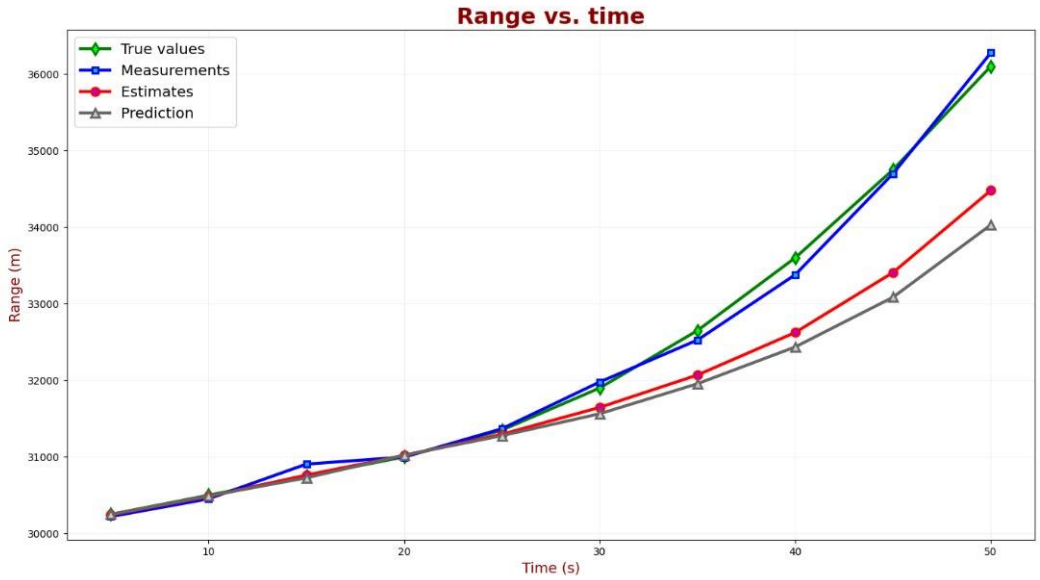
عند تطبيق خوارزمية المرشح على متحرك بسرعة منتظمة سيلاحظ مقدار التنعيم في قيم التقدير بالنسبة للقيم المقاسة، الشكل(3-37)، المنحني الأخضر يمثل القيم الحقيقية، القيم المقاسة تمثل بالأزرق، القيم المقدرة تمثل بالأحمر، والقيم المتوقعة تمثل بالرمادي.



الشكل(3-37) مخطط القيم الحقيقية والقيم المقدرة من قبل مرشح كالمن (Becker, 2023)

تم استخدام قيمة α و β منخفضة، ونحصل بالتالي على تنعيم جيد، لكن سيكون رد فعل المرشح لتغيرات السرعة المستهدفة بطيء، أما إذا استخدمنا تجهيزات عالية الدقة، فإننا نفضل قيمة α عالية و β التي تتبع القياسات. في هذه الحالة، سوف يستجيب المرشح بسرعة لتغيير سرعة الهدف.

في حالة وجود التسارع كما في الشكل (3-38)، سنلاحظ مشكلة حقيقية كما في الشكل (3-37)

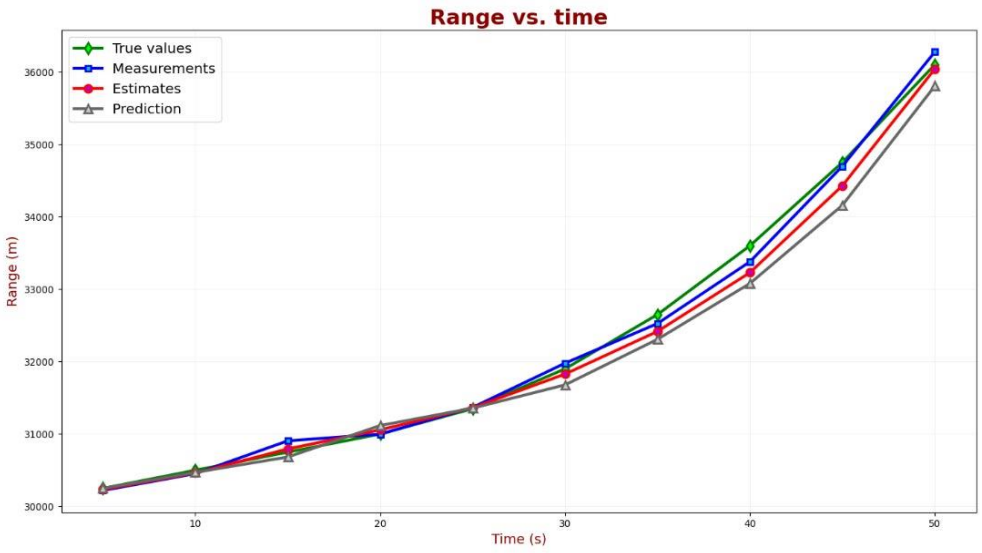


الشكل (3-38) تقدير الموقع بتغيير السرعة (Becker, 2023)

يلاحظ في الأشكال (3-37) و (3-38) أن القيمة المقدرة تبتعد بالتدريج عن القيمة الحقيقية، وينتج بخطأ التأخر. يظهر خطأ التأخر أثناء فترة التسارع، وبعد فترة التسارع، يقوم المرشح بإغلاق الفجوة ويتقارب نحو القيمة الحقيقية. ومع ذلك، يمكن أن يؤدي خطأ التأخر إلى خسارة الهدف. الخطأ المتأخر غير مقبول في الكثير من التطبيقات.

3-5-3 مرشحات α–β–γ

يمثل γ تسارع المتحرك، وبإضافة هذا المعامل لدراستنا السابقة، يمكن لمرشح α–β–γ مع معادلات النموذج الديناميكي التي تتضمن التسارع تتبع الهدف بتسارع ثابت وإزالة خطأ التأخر، كما في الاشكال (3-39).



الشكل (3-39) تقدير المسافة للمرشح α–β–γ (Becker, 2023)

ولكن في حالة وجود هدف مناور أي يغير الاتجاه فجأة عن طريق إجراء مناورة. في مثل هذه الحالات، ينتج مرشح α–β–γ ذو المعاملات الثابتة α–β–γ أخطاء في التقدير، وفي بعض الحالات، يفقد المسار المستهدف.

من أهم مرشحات α–β–(γ)، مرشح كالمن Kalman Filter، من المهم بهذه المرشحات هو اختيار معاملات الترجيح α–β–γ. تستخدم بعض أنواع المرشحات معاملات ترجيح ثابتة؛ ويحسب آخرون معاملات الترجيح لكل تكرار (دورة) للمرشح. إن اختيار الثوابت α وβ وγ أمر بالغ الأهمية للتشغيل السليم لخوارزمية التقدير.

4-5-3 مرشح كالمن

1-4-5-3 بعض المفاهيم الهامة

التباين Variance: التباين هو مقياس لمقدار انتشار مجموعة البيانات من موقع متوسطها.

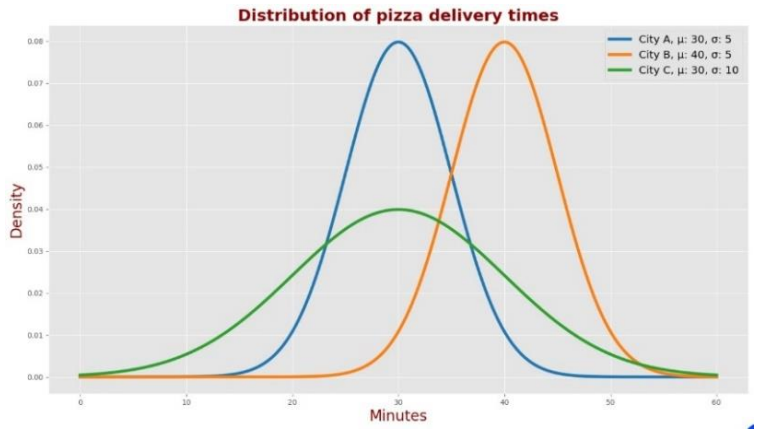
الانحراف المعياري Standard Deviation: الانحراف المعياري هو الجذر التربيعي للتباين، ويرمز إلى الانحراف المعياري بالحرف اليوناني σ، والمتوسط μ. عندما نقوم بتقدير التباين، فإن معادلة حساب التباين لـ N عينة معطاة بالعلاقة (13-3)، (Becker, 2023):

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - \mu)^2 \qquad (13\text{-}3)$$

التوزيع الطبيعي: يتبين أن العديد من الظواهر الطبيعية تتبع التوزيع الطبيعي. ويُوصَف التوزيع الطبيعي، المعروف أيضًا باسم التوزيع الغاوصي بالمعادلة (14-3)، (Becker, 2023):

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad (14\text{-}3)$$

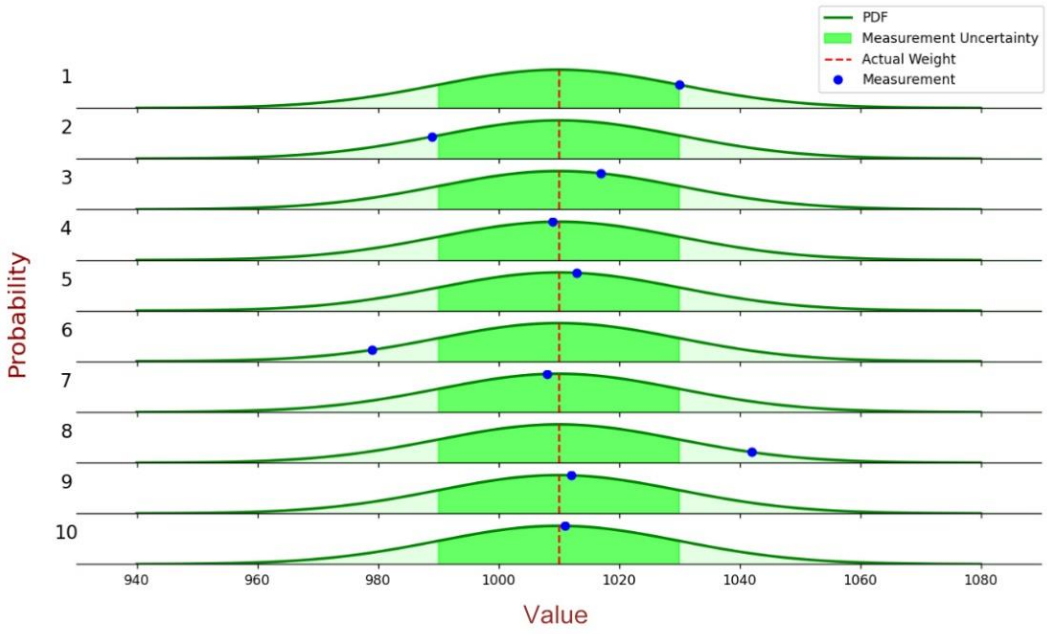
يوضح الرسم البياني في الشكل (3-40) التالي التوزيع الاحتمالي أو تابع الكثافة الاحتمالي Probability Density Function (PDF) لوقت توصيل البيتزا في ثلاث مدن: المدينة A، والمدينة B، والمدينة C.



الشكل (3-40) التوزيع الغاوصي (Becker, 2023)

2-4-5-3 معادلات مرشح كالمن

يعتمد مرشح كالمن على خمس معادلات، منها معادلة تحديث الحالة ومعادلة النموذج الديناميكي. يعامل مرشح كالمان القياسات وتقدير الحالة الحالية وتقدير الحالة التالية (التنبؤات) كمتغيرات عشوائية موزعة بشكل طبيعي. يتم وصف المتغير العشوائي بالمتوسط والتباين. في مرشح كالمن يُعتبر التقدير والقياس كمتحول عشوائي. إن الفرق بين التقديرات والقيم الحقيقية هو خطأ التقدير. ويقل خطأ التقدير كلما أجرينا قياسات إضافية، بينما تقترب القيمة المقدرة من القيمة الحقيقية. لا نعرف خطأ التقدير، ولكن يمكن تقدير عدم اليقين في الحالة، ويشار إلى تباين تقدير الحالة بـ p وبالنسبة لأخطاء القياس فهي الاختلافات بين القياسات والقيم الحقيقية، ولأن أخطاء القياس عشوائية، توصف بالتباين وهي مربع الانحراف المعياري (σ) لأخطاء القياس، (عدم اليقين في القياس) Measurement uncertainty. نرمز إلى تباين القياس بـ r ويمكن توفير تباين أخطاء القياس من قبل مواصفات جهاز القياس، أو حسابه، أو استنتاجه تجريبيًا بالمعايرة، وهو يتعلق بالكثير من المعاملات مثل أخطاء الكاميرا، نسبة الإشارة إلى الضجيج Signal to Noise Ratio (SNR)، وعتبات التجزئة الخ. يوضح الشكل (3-41) مثال لعشرة قياسات لمقدار فيزيائي.



الشكل (3-41) مثال لعشرة قياسات (Becker, 2023)

- تصف النقاط الزرقاء القياسات.
- يتم وصف القيم الحقيقية بواسطة الخط الأحمر المنقط.
- يصف المنحني الأخضر تابع الكثافة الاحتمالية للقياس.
- المنطقة الخضراء العريضة هي الانحراف المعياري σ للقياس، أي أن هناك احتمال بنسبة 68.26% أن تكون قيمة القياس ضمن هذه المنطقة. حيث 7 من أصل 10 قياسات تقع ضمن حدود σ

إن استقراء عدم اليقين التقديري هو: Δt^x ومعطى بالعلاقة (15-3)، (Becker, 2023):

$$p^x{}_{n+1} = p^x{}_{n,n} + \Delta t^2 . p^v{}_{n,n} \quad (15\text{-}3)$$

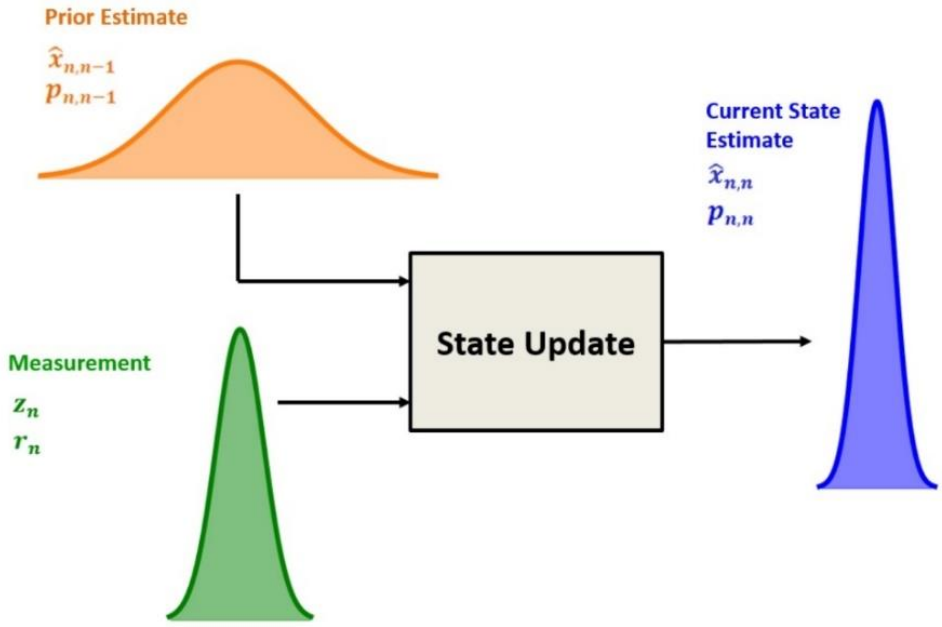
بافتراض نموذج سرعة ثابت

$$p^x{}_{n+1} = p^v{}_{n,n} \quad (16\text{-}3)$$

حيث: p^x هو تباين تقدير الموضع، p^v هو تباين تقدير السرعة

تُسمى معادلة استقراء عدم اليقين في التقدير بمعادلة استقراء التباين، وهي معادلة مرشح كالمان الثالثة. لتقدير الحالة الحالية للنظام، نجمع بين متغيرين عشوائيين كما في (الشكل 3-42):

- تقدير الحالة السابقة (تقدير الحالة الحالية المتوقع في الحالة السابقة)
- القياس



الشكل (3-42) تقدير الحالة (Becker, 2023)

مرشح كالمان هو مرشح مثالي يجمع بين تقدير الحالة السابقة والقياس بطريقة تقلل من عدم اليقين في تقدير الحالة الحالية، وتقدير الحالة السابقة معطى بالعلاقة (17-3)، (Becker, 2023):

$$\hat{x}_{n,n} = w_1 z_n + w_2 \hat{x}_{n,n-1} \quad (17\text{-}3)$$

حيث w_1, w_2 هما أوزان القياس وتقدير الحالة السابقة وتكون $w_1 + w_2 = 1$. وتكتب المعادلة السابقة بالعلاقة (15-3)، وبالنسبة للتباين بالعلاقة (16-3)، وباعتبار مصطلح ربح المرشح k_n هو التباين في التقدير $p_{n,n-1}$ مقسوم على مجموع التباين في التقدير والتباين في القياس r_n ويعطى بالعلاقة (17-3):

$$\hat{x}_{n,n} = w_1 z_n + (1 - w_1)\hat{x}_{n,n-1} \quad (18\text{-}3)$$

$$p_{n,n} = {w_1}^2 r_n + (1 - w_1)^2 \ p_{n,n-1} \quad (19\text{-}3)$$

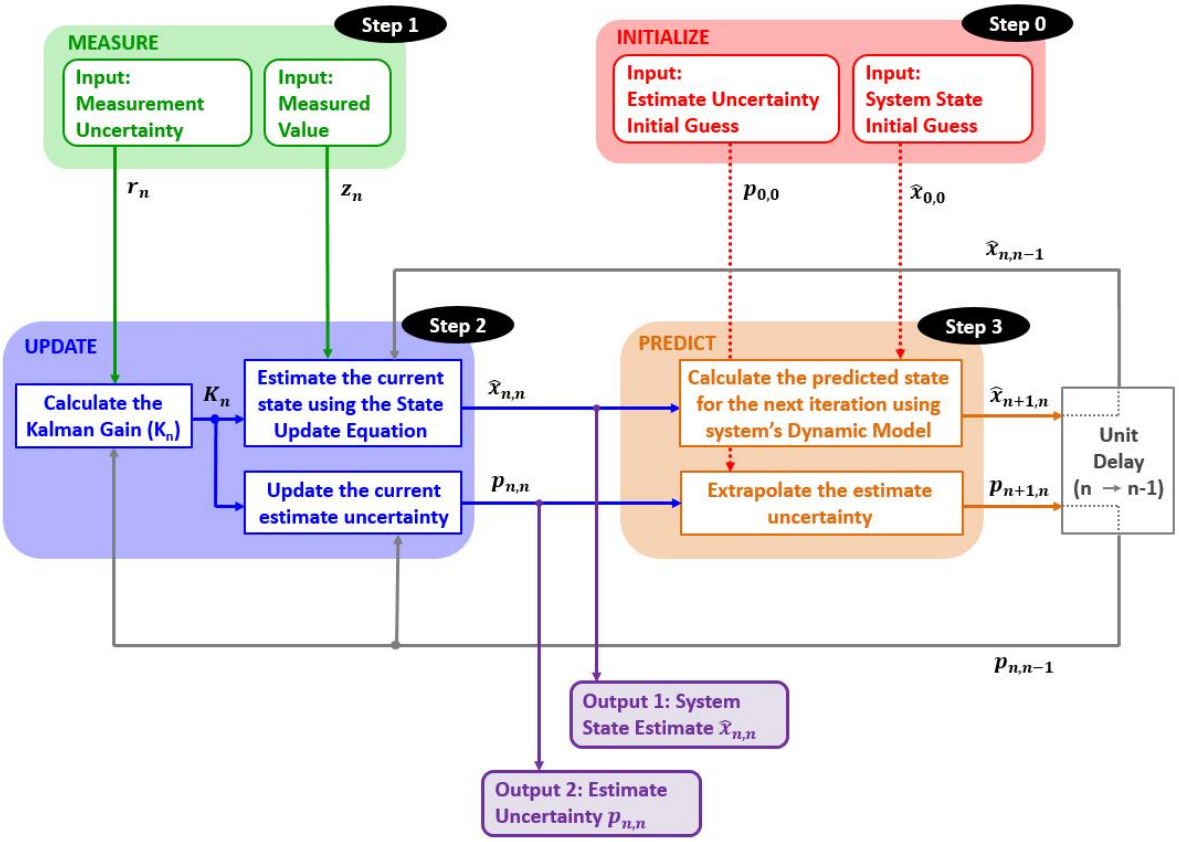
$$k_n = w_n = \frac{p_{n,n-1}}{p_{n,n-1}+r_n} \quad (20\text{-}3)$$

يمكن كتابة معادلات مرشح كالمن الكاملة كما في الجدول (3-1)

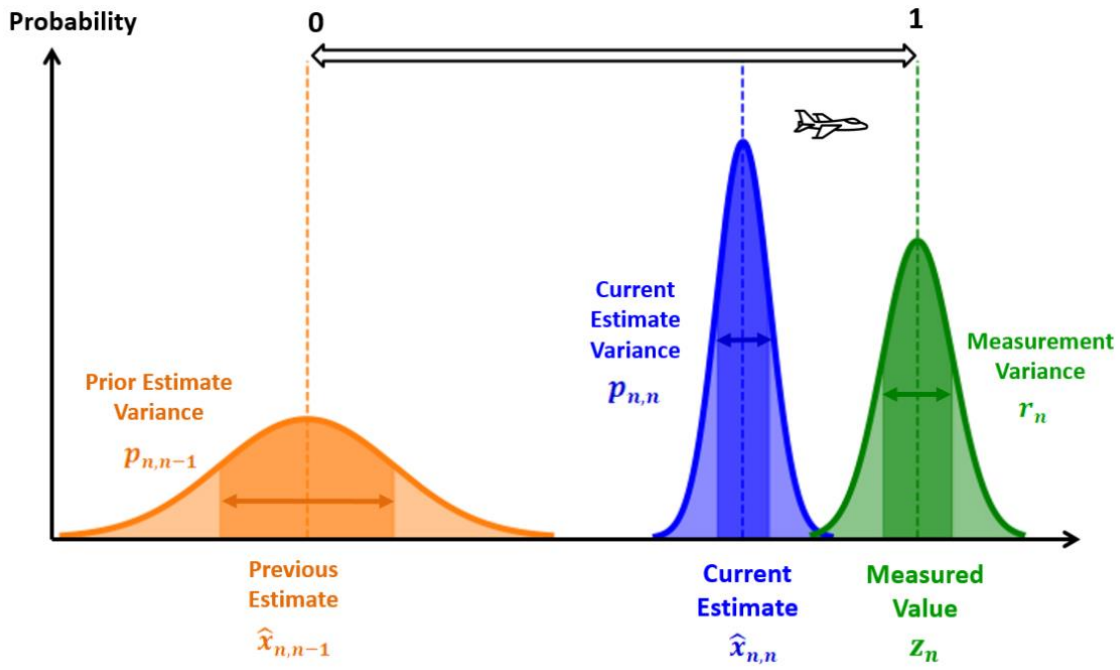
الجدول (3-1) معادلات مرشح كالمن (Becker, 2023)

	Equation	Equation Name
State Update تحديث الحالة	$\hat{x}_{n,n} = \hat{x}_{n,n-1} + K_n\,(z_n - \hat{x}_{n,n-1})$	State Update تحديث الحالة
	$p_{n,n} = (1 - K_n)\,p_{n,n-1}$	Covariance Update تحديث التغاير
	$K_n = \frac{p_{n,n-1}}{p_{n,n-1}+r_n}$	Kalman Gain ربح المرشح
State Predict التنبؤ بالحالة	$\hat{x}_{n+1,n} = \hat{x}_{n,n}$ (For constant dynamics)	State Extrapolation استقراء الحالة
	$\hat{x}_{n+1,n} = \hat{x}_{n,n} + \Delta t\hat{\dot{x}}_{n,n}$ $\hat{\dot{x}}_{n+1,n} = \hat{\dot{x}}_{n,n}$ (For constant velocity dynamics)	
	$p_{n+1,n} = p_{n,n}$ (For constant dynamics)	Covariance Extrapolation استقراء التغاير
	$p^x_{n+1,n} = p^x_{n,n} + \Delta t^2 \cdot p^v_{n,n}$ $p^v_{n+1,n} = p^v_{n,n}$ (For constant velocity dynamics)	

يقدم الشكل (3-43) وصفًا تفصيليًا لمخطط كتلة مرشح كالمان، بينما يبين الشكل (3-44) مخطط الكثافة الاحتمالية لكل من القيم المقاسة والقيم الحالية المخمنة والقيم المقدرة سابقاً.



الشكل (3-43) المخطط الصندوقي لمرشح كالمان (Becker, 2023)



الشكل (3-44) الكثافة الاحتمالية لكل من القيم المقاسة والقيم الحالية المخمنة والقيم المقدرة سابقاً (Becker, 2023)

6-3 خوارزميات الملاحقة

تهدف خوارزمية الملاحقة إلى ملاحقة كائنات متعددة موجودة في إطار. يتم تدريب خوارزميات ملاحقة الكائنات المتعددة على كمية كبيرة من البيانات، على عكس خوارزميات الملاحقة التقليدية. لذلك، ثبت أنها أكثر دقة لأنها يمكن أن تلاحق كائنات متعددة ومن فئات مختلفة في نفس الوقت مع الحفاظ على سرعة عالية. ومن هذه الخوارزميات DeepSORT وJDE وCenterTrack وهي خوارزميات قوية للغاية وتتعامل مع معظم التحديات التي تواجه أجهزة التعقب. إن DeepSORT هي خوارزمية تستخدم في الرؤية الحاسوبية لملاحقة الكائنات المتعددة وهي تطوير لخوارزمية التتبع البسيط عبر الإنترنت في الوقت الفعلي Simple Online Realtime (SORT) Tracking. يقدم DeepSORT التعلم العميق في خوارزمية SORT عن طريق إضافة واصف المظهر لتقليل تبديلات الهوية التعريفية، وبالتالي جعل الملاحقة أكثر كفاءة. لفهم DeepSORT، نوضح أولاً كيف تعمل خوارزمية SORT.

1-6-3 خوارزمية الملاحقة SORT

يتكون SORT من 4 مكونات رئيسية وهي كما يلي:

الكشف: هذه هي الخطوة الأولى في وحدة الملاحقة. في هذه الخطوة، يتم اكتشاف الكائنات الموجودة في الإطار والمراد تتبعها. ثم يتم تمرير هذه الاكتشافات إلى الخطوة التالية. يتم استخدام خوارزميات الكشف مثل YOLO بشكل متكرر.

التقدير: في هذه الخطوة، يتم نشر الاكتشافات من الإطار الحالي إلى الإطار التالي، وهو تقدير موضع الهدف في الإطار التالي باستخدام نموذج السرعة الثابتة. عندما يتم ربط الكشف بهدف، يتم استخدام مربع التحديد المكتشف لتحديث حالة الهدف حيث يتم حل مكونات السرعة بشكل مثالي عبر إطار عمل مرشح كالمن.

ارتباط البيانات: لدينا الآن مربع تحديد الهدف ومربع التحديد المكتشف. لذا، يتم حساب مصفوفة التكلفة كمسافة تقاطع فوق الاتحاد (IOU) بين كل كشف وجميع مربعات التحديد المتوقعة من الأهداف الموجودة. يتم حل التعيين بشكل مثالي باستخدام الخوارزمية المجرية. إذا كانت IOU للكشف والهدف أقل من قيمة عتبة معينة تسمى IOUmin، فسيتم رفض هذا التعيين.

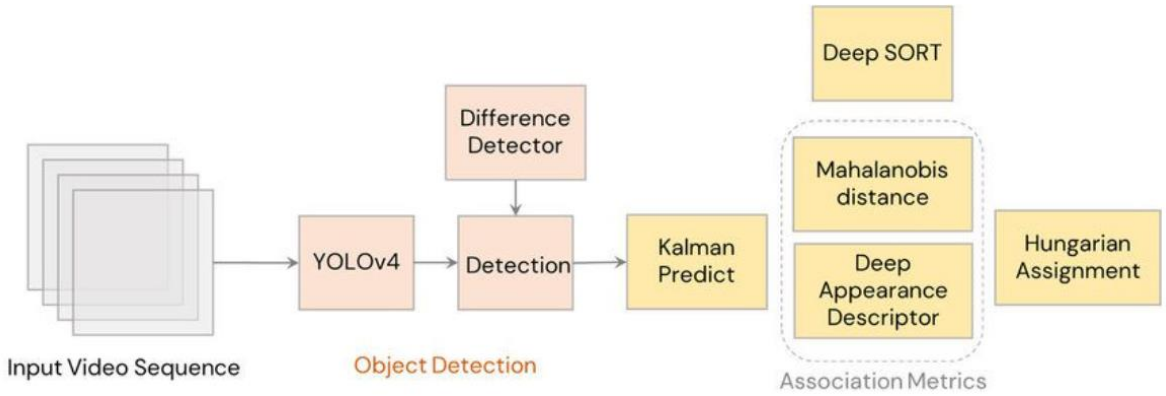
إنشاء وحذف هويات المسار: هذه الوحدة مسؤولة عن إنشاء وحذف المعرفات. يتم إنشاء هويات فريدة وتدميرها وفقًا لـ IOUmin. إذا كان التداخل بين الكشف والهدف أقل من IOUmin، فهذا يدل على الكائن غير المتعقب. يتم إنهاء المسارات إذا لم يتم اكتشافها لإطارات TLost، يمكنك تحديد مقدار الإطار الذي يجب أن يكون لـ TLost. إذا ظهر كائن مرة أخرى، فسيتم استئناف الملاحقة ضمنيًا تحت هوية جديدة.

يمكن تتبع الكائنات بنجاح باستخدام خوارزميات SORT التي تتفوق على العديد من الخوارزميات الحديثة. يمنحنا الكاشف اكتشافات، وتمنحنا مرشحات كالمن مسارات، وتقوم الخوارزمية المجرية بربط البيانات.

3-6-2 خوارزمية DeepSORT

يعمل SORT بشكل جيد للغاية من حيث دقة الملاحقة. لكن SORT يعيد المسارات بعدد كبير من تبديلات المعرف ويفشل في حالة الاحتجاب. ويرجع هذا إلى مصفوفة الارتباط المستخدمة. يستخدم DeepSORT مقياس ارتباط أفضل يجمع بين كل من موصفات الحركة والمظهر. يمكن تعريف DeepSORT على أنه خوارزمية الملاحقة التي تلاحق الكائنات بناءً على المظهر بالاضافة لسرعة وحركة هذه الكائن.

للأغراض المذكورة أعلاه، يتم التدريب بتضمين ميزة تمييزية جيدة دون اتصال بالإنترنت قبل تنفيذ الملاحقة. يتم تدريب الشبكة على مجموعة بيانات إعادة تحديد هوية الشخص على نطاق واسع مما يجعلها مناسبة لسياق التتبع. لتدريب نموذج مقياس الارتباط العميق في DeepSORT، يتم استخدام نهج التعلم القياسي لجيب التمام Cosine metric learning approach. وفقاً لـ DeepSORT، تُأخذ مسافة جيب التمام في الاعتبار معلومات المظهر المفيدة بشكل خاص لاستعادة الهويات بعد الاحتجابات طويلة الأمد عندما تكون الحركة أقل تمييزاً. هذا يعني أن مسافة جيب التمام هي مقياس يساعد النموذج على استعادة الهويات في حالة الاحتجابات طويلة الأمد وفشل تقدير الحركة أيضًا. يمكن أن يؤدي استخدام هذه الأشياء البسيطة إلى جعل الملاحق أكثر قوة ودقة.



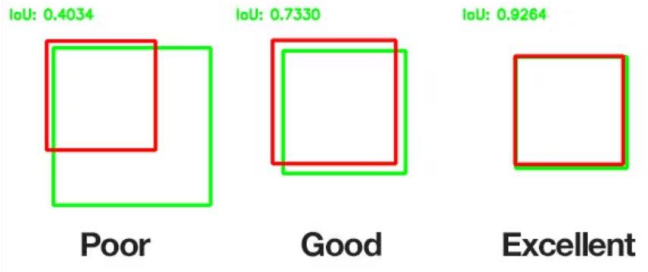
الشكل (3-45) بنية خوارزمية DeepSORT (Borah, 2022)

3-6-3 معايير الاختبار لخوارزمية ملاحقة الكائنات

تنتج خوارزميات الملاحقة تنبؤين رئيسيين: هوية الكائن (العلامة) والموقع (إحداثيات المربع المحيط). يتم تحديد دقة هذه النماذج من خلال تقييم كل من هذه التنبؤات وتحليل مدى قدرتها على تحديد موقع الكائن.

تساعد المقاييس مثل الدقة Accuracy, Precision والتذكر Recall ودرجة F1-score في تقييم قدرة النموذج على تصنيف الكائن الموجود. في حين توفر الدقة صورة عامة، فإن الدقة والتذكر يحكمان على النموذج بناءً على حدوث نتائج إيجابية وسلبية خاطئة (Encord, 2025).

تُستخدم مقاييس مثل التقاطع فوق الاتحاد (IoU) لدقة التوطين. تحسب IoU مدى تطابق المربع المحيط المتوقع مع قيمته الحقيقية الأساسية. تعني القيمة الأعلى تقاطعًا أعلى، وبالتالي دقة أعلى كما في الشكل(3-46).



الشكل (3-46) معيار التقاطع فوق الاتحاد (COWTON, 2019)

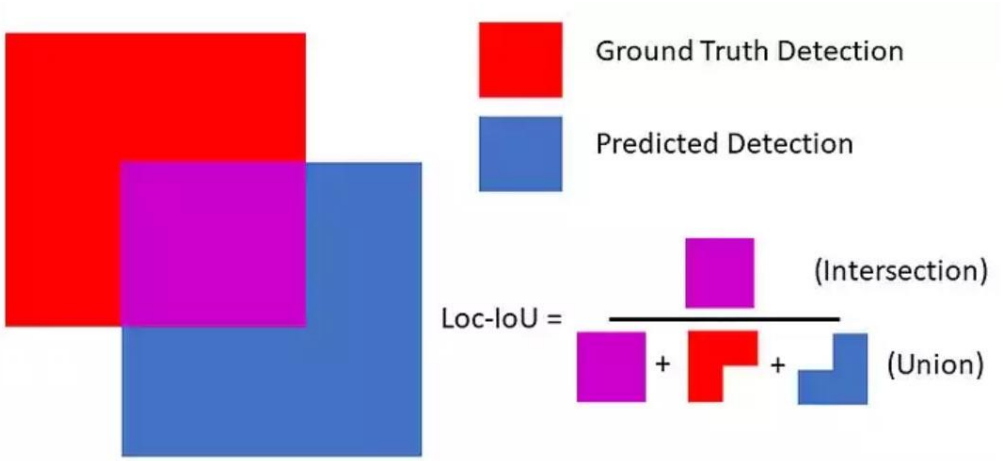
درجة الثقة: يتم حساب درجة الثقة كمعيار للتقييم. تُظهر درجة الثقة هذه احتمالية اكتشاف الصورة بشكل صحيح بواسطة الخوارزمية ويتم إعطاؤها كنسبة مئوية (Shaw, 2021). يتم أخذ الدرجات على متوسط الدقة المتوسطة عند عتبات مختلفة لـ IoU (التقاطع على الاتحاد). يتم حساب الدقة عند قيمة العتبة اعتمادًا على الإيجابيات الحقيقية (TP) والإيجابيات الكاذبة (FP) والسلبيات الكاذبة (FN) والتي تعد نتيجة لمقارنة الكائن المتوقع بالكائن الحقيقي الأساسي

- يتم حساب TP عندما يتم مطابقة مربع حدود الحقيقة الأساسية مع مربع حدود التنبؤ (تتوافق مع IoU).
- يتم حساب FP عندما لا يكون للكائن المتوقع أي ارتباط بكائن الحقيقة الأساسية.
- يتم حساب FN عندما لا يكون لكائن الحقيقة الأساسية أي ارتباط بالكائن المتوقع. يتم إعطاء درجة الثقة كمتوسط لجميع درجات الدقة لجميع العتبات وفق العلاقة (3-21) على النحو التالي (Shaw, 2021):

$$Confidence = Average\ precision = \frac{1}{|Thresholds|}\sum_T \frac{TP}{TP+FP+FN} \qquad (21\text{-}3)$$

مقياس الأداء MOTA: يعتبر MOTA من المقاييس الشائعة لأداء الملاحقة. نتيجة مطابقة الخوارزمية للمسارات المتوقعة مع الأشياء الحقيقية الأساسية في كل إطار، تزيد درجات IoU الإجمالية. وينتج عن هذا TP, FP, FN. تقيس دقة الاكتشاف (DetA) مدى نجاح أداة التتبع في تحديد موقع الكائنات في كل إطار باستخدام عتبات (IoU) عادةً العلاقة (3-22) والشكل (3-47). وهي تقيس بشكل أساسي الدقة المكانية للاكتشافات.

$$DetA = \frac{TP}{Tp+FP+FN} \qquad (22\text{-}3)$$



الشكل (3-47) الدقة المحلية (Luiten, 2021)

يقدم MOTA مفهوم تتبع الهوية لمقاييس اكتشاف الكائنات. وهو يشتمل على تبديلات الهوية (IDSW)، والتي تحدث عندما يتم تعيين كائن واحد من الكائنات الحقيقية الأساسية (GT) لتوقعات مسارات مختلفة بمرور الوقت.

يتضمن حساب MOTA التبعية الزمنية. يتم حساب IDSW عندما يتطابق من GT هدف i مع المسار j في الإطار الحالي ولكن تم مطابقته مع مسار مختلف k حيث (k≠j) في الإطار السابق.

يتم حساب مقياس MOTA عبر جميع الإطارات في العلاقة (3-23)، (Luiten, 2021) على النحو التالي:

$$MOTA = 1 - \frac{\sum_t (FN_t + FP_t + IDSW_t)}{\sum_t GT_t} \qquad (23\text{-}3)$$

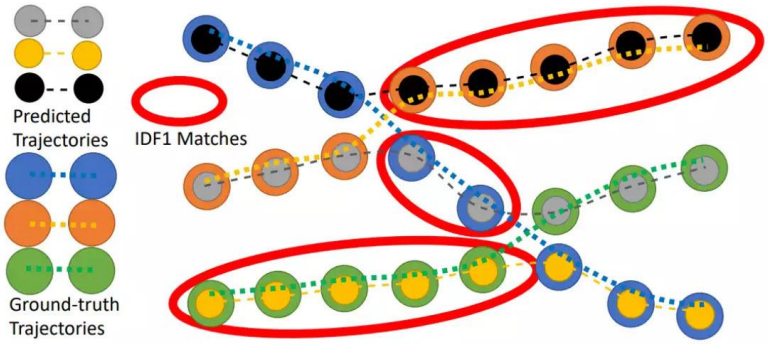
حيث t هو مؤشر الإطار، IDSW هي تبديلات الهوية، وGT هو عدد كائنات الحقيقة الأساسية.

مقياس IDF1: يعالج IDF1 بعض قيود MOTA من خلال التركيز على المدة التي يحدد فيها المتعقب الكائن بشكل صحيح، بدلاً من مجرد حساب الأخطاء. ويستند إلى مفهوم دقة التعريف (IDP) وتذكر التعريف (IDR).

في العلاقة (4-18)، (Luiten, 2021)، يحسب التعيين بين كائنات التنبؤ والكائنات الحقيقية عبر الفيديو بالكامل، بدلاً من إطار بإطار، ويبين الشكل (3-24) توضيحا لمعيار المطابقة بالمسار.

$$IDF1 = \frac{2 * IDTP}{2 * IDTP + IDFP + IDFN} \qquad (24\text{-}3)$$

- IDTP (معرف إيجابي حقيقي): عدد الاكتشافات التي تم تحديدها بشكل صحيح
- IDFP (معرف إيجابي كاذب): تنبؤات المتتبع التي لا تتطابق مع أي حقيقة أساسية
- IDFN (معرف سلبي كاذب): مسارات الحقيقة الأساسية التي لا يتم تتبعها



الشكل (3-48) مطابقة المسارات (Encord2, 2024).

معيار وقت الاستدلال: يحدد وقت الاستدلال مدى سرعة معالجة الخوارزمية لإطار الفيديو والتنبؤ بتسمية الكائن وموقعه. غالبًا ما يتم قياسه بالإطارات في الثانية (FPS). يشير هذا إلى عدد الإطارات التي يمكن للخوارزمية معالجتها وإخراجها كل ثانية. تشير قيمة FPS الأعلى إلى استدلال أسرع.

خلاصة الفصل الثالث:

تم في هذا الفصل عرض بعض المبادئ المهمة في التعلم العميق، التي تخدم موضوع الاطروحة في تمييز الكائنات بعد تداخلها، وتم التركيز على مواضع السمات، الكشف، والملاحقة، حيث تبدأ ملاحقة الكائنات باكتشافها أولاً. وكما لاحظنا تعتبر الشبكات العصبونية التلافيفية (CNN) رائدة في مجالات اكتشاف الكائنات، تجزئة الصور، وتصنيف الصور، أما خوارزمية YOLO وخصوصاً اصداراتها الأخيرة، فتقدم طريقة مناسبة للكشف عالي الجودة والسرعة، بالتالي هي مناسبة جداً للعمل في الزمن الحقيقي بالتالي ستخدم الطريقة المقترحة في هذه الاطروحة. يعتبر التحدي الأساسي في المقترح هو التغيرات المفاجئة في المظهر، والاحتجاب الشديد للجسم. لذلك كان لابد من دراسة خوارزمية ملاحقة تأخذ هذه المعاملات بعين الاعتبار، وهو ما تم الاعتناء به في هذا الفصل بحيث تم التركيز على خوارزمية الملاحقة DeepSORT، كونها ستخدم عملنا وستكون أساساً مرجعياً لمقارنة النتائج، حيث تستخدمها بعض الأبحاث المرجعية وفق طريقة TBD بالاعتماد على كاشف YOLO في عمليات الملاحقة.

الفصل الرابع

الخوارزمية المقترحة لملاحقة وتمييز الكائنات الحرارية المتداخلة

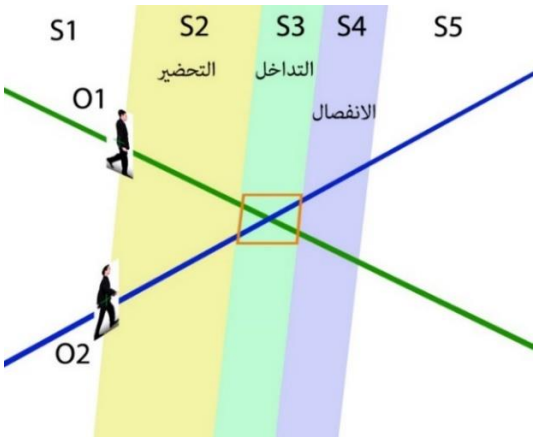
Proposed Algorithm for Overlapped Thermal Objects Tracking and Recognition

مقدمة:

يهتم هذا البحث بتمييز الكائنات بعد تداخلها ويمكن اعتباره آلية جزئية من أي خوارزمية ملاحقة، ونحتاج للعمل في أكثر من مرحلة قبل وأثناء وبعد التداخل. يوصف سيناريو العمل للخوارزمية المطورة على الشكل التالي:

- من المهم جداً لدى ملاحقة الكائن بالحالة العامة الحصول على معلومات هامة تخصه، وتتمثل هذه المعلومات بمساره. وذلك ما ينطبق على مرحلة ما قبل التداخل. لذلك تقوم الخوارزمية في المرحلة S1 بالملاحقة بملاحق مدرب ومرشح كالمن يعتمد على مركز الكائن، كما في الشكل (4-1).
- يتم التنبؤ بمرشح كالمن بتوقيت التداخل وعند عدد معين من الإطارات قبل التداخل تبدأ مرحلة التحضير للتداخل S2 ستكون هنا المعلومات المحصلة هي المسار بالإضافة لشكل الكائن (البصمة).
- المرحلة الثالثة S3، هي مرحلة التداخل حيث ضياع الكائنات المتداخلة، ونشوء كائنات جديدة وهنا لابد من الاحتفاظ بمعاملات التقدير لمسارات الكائنات السابقة وتقديرات تغير الاشكال.
- المرحلة الرابعة S4، يتم فيها نشوء كائنات جديدة وضياع الكائنات الناتجة عن التداخل وتقوم الخوارزمية بمقارنة المعاملات مع معاملات التقدير السابقة.
- في المرحلة S5، تتم العودة للملاحقة العادية كما في S1

يبين الشكل (4-1) تمثلاً مبسطا لمراحل تنفيذ الخوارزمية عند التقاء شخصين وانفصالهما



الشكل (4-1) مراحل التداخل في الطريقة المطورة (الباحث)

تنبع التحديات الرئيسية عادةً من المشكلات الموجودة في الصورة والتي تجعل من الصعب على النماذج إجراء عمليات الكشف بشكل فعال على الصور، بالإضافة لذلك فإن مشاكل عملية الملاحقة وخصوصاً الاحتجاب (Klingler. 2023)، ستزيد الأمر سوءاً. إن بناء خوارزمية مطورة تستطيع التعامل مع الكائنات المتداخلة أثناء حركتها يقتضي فهم أهم المشاكل التي يمكن أن تعترض عملية الملاحقة، بالإضافة للمشاكل الإضافية الناتجة عن الصورة الحرارية على طريقة الحل. نناقش هنا بعض المشكلات الأكثر شيوعًا مع مهمة ملاحقة الكائنات ومن ثم نعرض لطريقة الحل التي اتبعناها.

4-1 مشاكل عملية ملاحقة الكائنات

4-1-1 سرعة الملاحقة

يعد تحسين سرعة الملاحقة Tracking Speed أمراً ضروريًا بشكل خاص لنماذج الملاحقة في الزمن الحقيقي، فبالإضافة للدقة المطلوبة لتنفيذ الملاحقة، من المفترض أن تقوم خوارزميات ملاحقة الكائنات بإجراء عمليات الكشف وتحديد موقع الكائنات ذات الأهمية، وذلك في أقل قدر ممكن من الوقت. يجب استخدام نماذج كشف سريعة واختيارها بعناية وذلك لإدارة الوقت المستغرق لأداء النموذج. يمكن أن تكون التعديلات على الخوارزمية هي العامل المميز بين النموذج الأسرع والنموذج الأبطأ. تؤثر اختيارات التصميم بالإضافة إلى إطار الكشف أيضًا على التوازن بين سرعة ودقة نموذج الكشف عن الكائنات.

4-1-2 تشويه الخلفية

تؤثر خلفيات الصور Image Background المدخلة أو الصور المستخدمة لتدريب النماذج أيضًا على دقة النموذج. فالخلفيات المزدحمة بالكائنات الملاحقة تصعب الأمر كثيراً، فمع وجود خلفية ضبابية أو أحادية اللون، يسهل على نظام الذكاء الاصطناعي اكتشاف وملاحقة الكائنات. أما الخلفيات المزدحمة للغاية أو التي لها نفس لون الكائن فقد تجعل من الصعب ملاحقة الكائنات، ويدخل في ذلك أيضاً نقصان اضاءة المشهد.

4-1-3 مقاييس مكانية متعددة

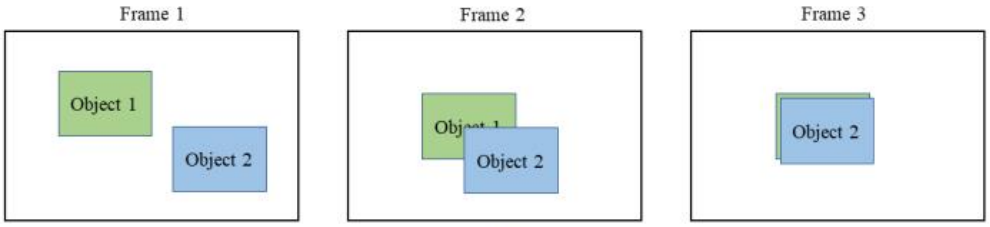
إن وجود أحجام وأبعاد متنوعة للكائنات Multiple Spatial Scales تربك الخوارزميات وتجعلها تعتقد أن الكائنات أكبر أو أصغر من حجمها الفعلي. ويمكن أن تؤثر المفاهيم الخاطئة حول الحجم سلبًا على سرعة الاكتشاف. لمكافحة مشكلة المقاييس المكانية المتنوعة.

4-1-4 مشكلة ربط المسارات

تسمى عملية ربط الأجسام الموجودة بالزمن t بالاكتشافات الجديدة مع الأجسام في اللقطة السابقة t-1 بربط المسارات. ولعل احدى أهم مشكلات ملاحقة الأجسام المتعددة MOT هي صعوبة ربط المسارات، وهي مشكلة أكثر تعقيدًا لأنها تنطوي على تتبع العديد من الأجسام في وقت واحد، ويتطلب ربط المسارات الموجودة بشكل صحيح بالاكتشافات الجديدة المقابلة لها. بالإضافة لصعوبة العمل مع الأهداف التي تتحرك في اتجاهات عشوائية بسرعات متفاوتة. وعلاوة على ذلك، يمكن أن يختلف عدد المسارات الجديدة مقارنة بالمسارات الموجودة بسبب دخول الأجسام إلى المشهد وخروجها منه أو ربما حجبها أو تداخلها مع بعضها البعض. تضيف هذه العوامل تعقيدًا إلى مهمة ربط المسارات. من المهم ملاحظة أن هذا الربط يجب أن يلتزم بالقيد الذي ينص على أنه لا يمكن تعيين كل مسار إلا لكشف واحد والعكس صحيح.

5-1-4 الاحتجاب

في مهام الرؤية بالذكاء الاصطناعي باستخدام التعلم العميق، يحدث الاحتجاب Occlusion عندما تقترب أشياء متعددة من بعضها البعض فتتداخل أو تندمج (Klingler. 2023). يتسبب هذا في حدوث مشكلات لأنظمة ملاحقة الكائنات لأن الكائنات المحجَبة والمحجِبة تُرى على أنها شيء واحد أو أن المحجَبة تتبع للكائن بشكل غير صحيح. يمكن أن يرتبك النظام ويحدد الشيء الذي تم ملاحقته مسبقاً على أنه شيء جديد، وتجدر الإشارة إلى أنه في المشاهد المزدحمة، يكون التعامل مع الاحتجابات صعبًا للغاية (Bashar, 2022). تعطي حساسية الاحتجاب فهماً أكبر للأجزاء الأكثر أهمية في الصورة لتصنيفها بواسطة نظام الملاحقة. تشير حساسية الاحتجاب إلى مقياس حساسية الشبكة الاحتجاب في مناطق بيانات مختلفة. يتم ذلك باستخدام مجموعات فرعية صغيرة من مجموعة البيانات الأصلية. الشكل (4-2) (a) رسم توضيحي لاحتجاب جسمين (أخضر وأزرق). في الإطار 1 الجسمان منفصلان عن بعضهما. في الإطار 2 تداخل جزئي، وفي الإطار 3 تداخل كامل. في الشكل (4-2) (b) المرأة ذات الثوب الأحمر مغطاة تقريبًا بعمود الإنارة والشخص، وهذا مثال على الاحتجاب.



(a)

(b)

الشكل (4-2) الاحتجاب (Occlusion) (Mk Bashar, 2022).

وأخيرا يمكن ذكر أن **تشابه الأهداف** قيد الملاحقة يؤثر كثيراً على عملية الملاحقة.

2-4 ملاحقة الكائنات الحرارية

مع انتشار أجهزة التصوير الحراري، تلقى تقنية ملاحقة الأجسام المزيد من الاهتمام باعتبارها تقنية رؤية ذكية بالغة الأهمية. وهي تستخدم على نطاق واسع في المراقبة بالفيديو (Tiu,2020). لا يزال دور الأشعة تحت الحمراء في مساعدة قيادة السيارات وكثير من تطبيقات المراقبة والقيادة عند مستوى منخفض (Hou, 2022)، وذلك بالمقارنة مع التصوير المرئي بالرغم من أن التصوير الحراري بالأشعة تحت الحمراء جزءًا لا غنى عنه من مستقبل الكاميرات وخصوصاً في مجالات الكشف والقيادة الذاتية.

1-2-4 ميزات العمل في مجال الصورة الحرارية

يتأثر الحساس الحراري بالحرارة المقابلة له، بحيث يمرر تيار متناسب مع الحرارة. يحوَل هذا التيار الى جهد يقاس لمعرفة مقدار درجة الحرارة. بتجميع عدد من الحساسات (مصفوفة MxN) يتم تكوين مصفوفة من الجهود، كل منها ممثل للحرارة المقابلة أمام الحساسات "المشهد الحراري". تحصل هذه الجهود وتحول إلى أرقام في تخزن في ذاكرة الكاميرا الحرارية. من الجدير بالذكر ان بعض الكاميرات تستخدم مبردات للحساسات، وذلك للتخلص من الضجيج المسبب من ارتفاع حرارة الحساس وخصوصا لدى تصوير مشاهد بأمدية بعيدة حيث يكون مستوى الضجيج مؤثر على الاشارة الصغيرة للحساس.

تلتقط الكاميرات الحرارية درجات الحرارة النسبية للمشهد والأهداف المختلفة، ولقد أثبتت هذه الكاميرات أنها تعمل بشكل موثوق، في ظل مختلف الظروف الجوية والإضاءة المختلفة ولمهام مختلفة في مجال الرؤية الحاسوبية (Brenner, 2023) و(Broyles,2022) و(Shin, 2025). تسمح هذه الخاصية في الكاميرات الحرارية لأنظمة الرؤية التي تستخدمها بالاستفادة من الأنماط الحرارية الفريدة للأشياء لتحقيق أداء أكثر قوة وموثوقية.

من أهم ميزات استخدام الصور الملتقطة من الكاميرات الحرارية:

- امكانية فصل الأهداف ذات الطاقات الحرارية العالية، مثل المشاة أو السيارات، عن الخلفيات الباردة، وقد أدى تميزها إلى استخدامها على نطاق واسع، لتحقيق مراقبة على مدار 24 ساعة في مجالات الصناعة والرؤية الروبوتية وأنظمة الدفاع، وذلك ما يمكّن من توفير الدعم لـ صنع القرار (Chen, 2022).
- بسبب القدرة على كشف الأجسام المصدرة للحرارة فإن العمل بالكاميرات الحرارية سيكون مستقلاً نوعاً ما عن ظروف الرؤية الصعبة كانخفاض الضوء وتأثيرات الإضاءة الضجيجية، والرؤية الليلية.
- العمل في الظروف المناخية الصعبة، وذلك بسبب اعتمادها على الطيف الحراري بالتالي تأثير الضباب والثلج منخفض بالمقارنة مع الكاميرات المرئية.

4-2-2 مشاكل الملاحقة ضمن الصور الحرارية

تواجه ملاحقة الأجسام المتعددة (MOT) في التصوير الحراري تحديات جديدة وذلك بالإضافة للتحديات التي تواجه الملاحقة للأجسام المرئية نذكر منها:

- الافتقار إلى السمات المرئية (El Ahmar, 2024)، وهي سمات تبنى على أساسها نماذج الملاحقة، مثل ميزة اللون وحواف وتفاصيل الجسم.
- التشوه والتباين المنخفض (Zha,2024)، (Yang, 2023)، (Liu, 2020).
- الاختلافات الصغيرة داخل الفئة (Zha,2024) حيث الكائنات من الصنف الواحد قريبة إلى التشابه مما يعقد عملية تعريف الكائن بالتالي عملية الملاحقة.
- تعقيد أنماط الحركة (El Ahmar, 2024). بسبب تغير الكثافة (Liu, 2020) وخسارة صلادة شكل الكائن.
- التقاطع الحراري (Liu, 2020) حيث ينتج كائنات متشابهة حرارياً بالرغم من كون ألوانها مختلفة.
- الأهداف الصغيرة (Liu, 2024)، على عكس صور الضوء المرئي، تعكس الصور بالأشعة تحت الحمراء معلومات الإشعاع الحراري للأشياء، مما يسمح باكتشاف الأشياء على مسافات كبيرة في تكنولوجيا القيادة الذاتية الحديثة. لكن تنبعث منها إشعاعات حرارية ضعيفة، وتظهر فرقًا ضئيلًا في درجات الحرارة عن محيطها، هذا يفرض تحديات خاصة. أولاً، غالبًا ما تعاني الأشياء الصغيرة في صور الأشعة تحت الحمراء من التداخل من الخلفيات المعقدة والتباين المنخفض والضجيج، مما يجعل من الصعب تمييزها عن الخلفية. ثانيًا، الحجم الصغير للأهداف بالأشعة تحت الحمراء، غالبًا ما يكون بضعة بكسلات فقط، يجعل تحديد الموقع والتعرف أكثر تحديًا. كذلك تظهر الأهداف الصغيرة في صور الأشعة تحت الحمراء عادةً كنقط ساخنة أو بقع ساخنة، مع معلومات محدودة عن الشكل والملمس، مما يزيد من صعوبة الكشف.

4-2-3 تمييز الكائنات ضمن الصور الحرارية

تنشأ صعوبة تمييز الكائنات عن بعضها في الصورة بشكل مباشر من ازدحام هذه الكائنات وبالذات عند تداخلها، وفي حالات انخفاض القدرة على إعادة تعريف الكائن خلال الوقت بالتالي فشل ملاحقته. وينشأ تعقيد مهمة إعادة تحديد هوية الكائن من الاختلافات في مظهر الكائن (Kniaz, 2021) وقد ينشأ هذا الاختلاف من:

- عوامل عديدة مثل التغييرات في الإضاءة
- المواقع وجهة النظر(Liu, 2023) .

تكون التغييرات بالاضاءة في الصورة الحرارية ضعيفة بالتالي يبقى العاملان الأساسيان هما المواقع واختلاف جهة النظر.

إن المنهجية الأكثر شيوعًا هي الملاحقة بالكشف، حيث يتم تحليل المشكلة إلى مهمتين رئيسيتين:

- الاكتشاف، الذي يحدد موقع الكائنات في كل إطار فيديو
- الارتباط، الذي يطابق الكائنات والمسارات الموجودة أو ينشئ مسارات لكائنات ظهرت حديثًا.

نحصل على المزايا الرئيسية الجيدة للملاحقة بالكشف، عندما نقوم بمقايضات جيدة بين الكفاءة الحسابية والأداء (Ma, 2024). وعند الملاحقة تنحصر المشكلة الأساسية في تهيئة المسار/إنهائه، ويعد حل مشكلة اختفاء المسار وإعادة ظهوره مشكلة صعبة، عندما لا يتم اكتشاف مسار موجود لمدة معينة، إما بسبب الاحتجاب أو ترك المشهد، يتم إنهاؤه عادةً، ثم يتم تهيئته بشكل غير صحيح كمسار جديد إذا ظهر مرة أخرى. وبالتالي، فإن حل مشكلة ظهور المسار وإعادة ظهوره والاحتجاب مشكلتان مترابطتان. في معظم تقنيات الملاحقة المرئية، يتسبب الاحتجاب أيضًا في تبديل الهوية Identification (Id)، وخاصة في المشاهد المزدحمة، والاحتجابات المستمرة. إذاً مشكلة تمييز الكائنات الحرارية وإعادة تعريف هويتها متطلب هام في مجال تداخل الكائنات لدى مراقبتها.

3-4 إطار العمل والحلول

إن تداخل الكائنات خلال اطارات الصورة يمثل مشكلة معقدة ومعظم الطرائق والخوارزميات الحديثة تبحث عن حلول جيدة لهذه المشكلة، وإذا أضيف لمشكلات التداخل مشكلات الصورة الحرارية، سنصطدم بصعوبات أقسى. لذلك تقسم اجرائيات المعالجة في هذا البحث إلى ثلاثة مناحي مختلفة، تنتج عن المشاكل الرئيسية التي تفرضها عملية التداخل للكائنات الحرارية في منطقة التداخل وجوارها، بغض النظر عن بقية مشاكل الملاحقة الأخرى:

- تعزيز سمات الكائنات الحرارية المكتشفة الضعيفة أصلاً، بالإضافة لتشوه الكائن والتباين المنخفض.
- استخدام سمات الحركة لأجزاء الجسم المراقب لتحسين ترابط المسارات.
- تطوير طريقة جديدة لإعادة تعريف الكائن ضمن إطار الصور الحرارية.

يؤمن نموذج الحركة عادة حلا قويا لترابط المسارات ولكن بسبب أن أغلب حالات فشل تمييز الأهداف عن بعضها بعد التداخل يكون في حالات تغير معادلات الحالة للنموذج، بالتالي سيكون الاعتماد بالحل باتجاه طريقة الملاحقة بالكشف، بحيث نؤمن إعادة ارتباط يعتمد بالدرجة الأولى على كشف الكائن وإعادة ربط شكله الحالي بالشكل المتوقع من الحالات السابقة. ومن الجدير بالذكر بأن المحاور الأساسية للخوارزمية المقترحة تدرس بشكل مستقل لسهولة التحليل ولكن في الحقيقة ستعمل هذه المحاور بالتوازي في الوقت ذاته، وكل منحى سيكون مؤثرا بالآخر.

1-3-4 تعزيز سمات الكائنات الحرارية المكتشفة

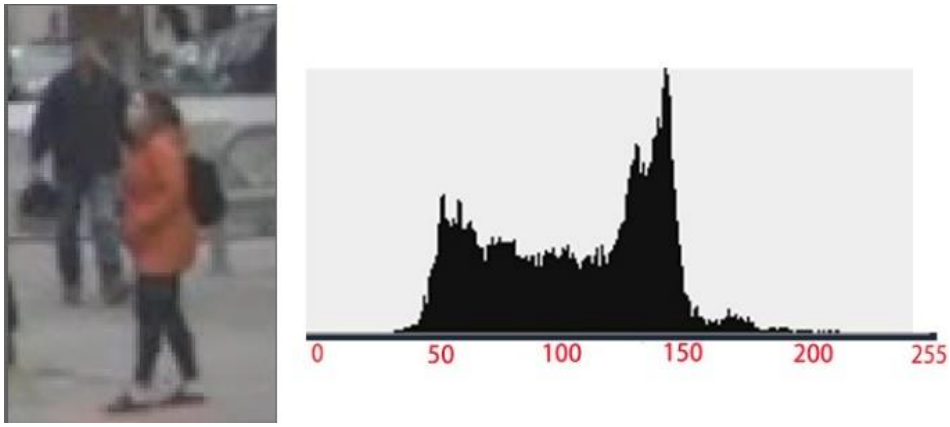
إن الصور الحرارية لها عيوب واضحة، مثل الدقة المنخفضة، وانخفاض نسبة الإشارة إلى الضجيج، ونقص معلومات الملمس واللون، والحواف غير الواضحة وتزيد أوجه القصور هذه من صعوبة تحقيق فهم مستقل للصورة في أنظمة القيادة الذاتية (Ding, 2023).

ساعد الذكاء الاصطناعي الكثير من خوارزميات كشف الكائنات وتصنيفها، وأصبح التعلم العميق عماد الأنظمة ذات الصلة. ولكن التعلم العميق بحاجة للعديد من مجموعات البيانات للتدريب والتحقق، وهو ما اهتم به المطورون حيث تم إنشاء العديد من مجموعات البيانات لمختلف الأصناف في الصور المرئية، ولكن مجموعات البيانات في مجال الصور الحرارية لم يساير المجال المرئي بل تخلف عنه كثيراً بسبب صعوبات بناء مجموعات البيانات وهو ما يمثل عقبة حقيقية أمام تطوير كاشف كائن مخصص للصور الحرارية. لدى البحث عن مجموعات بيانات للصور الحرارية نجد القليل منها مثل FLIR ADAS (Flir,2023)، ويلاحظ بالمقارنة مع مجموعات بيانات التصوير المرئي، بأن عدد الكائنات المصنفة أقل بكثير. وتقدم الصور الحرارية تفاصيل ومعلومات أقل من تلك المرئية، وهي غير واضحة لأسباب مختلفة مثل الظروف الجوية وسوء التركيز والطول الموجي العالي. إن كل ذلك يجعل معالجة الصور الحرارية يمثل تحدياً كبيراً بالمقارنة مع الصورة المرئية، فتغير شكل الهدف وطمس معالمه وانخفاض الدقة ومشاكل الخلفية هي مشاكل عالقة تؤخر التطوير بهذا المجال (Ashrafi, 2022). إذاً بالنسبة لمجموعات عينات الأهداف لتدريب النظام الذكي عليها، فإن جمع الكثير من الصور وترشيحها ووضع علامات عليها مهمة شاقة وباهظة الثمن، كما أن الجودة المنخفضة للصور الحرارية تجعل تثبيت المعلومات الموافقة للصور أمراً صعباً. يلاحظ في الشكل (4-3) صورتين لذات المشهد الصورة الأولى (a) تمثل صورة مرئية أما الصورة (b) فهي الصورة الحرارية. يمكن دراسة كلا الصورتين وتحليلهما، حيث يلاحظ أن هناك العديد من السمات المشتركة. في الصورة المرئية وهي التي تعودت عليها العين وتدرب نماذج الكشف عليها، وتظهر الأشخاص بوضوح أما في الصورة الحرارية فيظهر الجسم بشكل مختلف، لكن بالرغم من الاختلاف نلاحظ أن النقاط الساطعة تمثل نوعاً ما هيكل الجسم بالتالي يمكن تلمّس بعض التشابه في سمات الصورتين.

الشكل (4-3) صورتان لمشهد واحد من كاميرتين مرئية وحرارية (Flir, 2023).

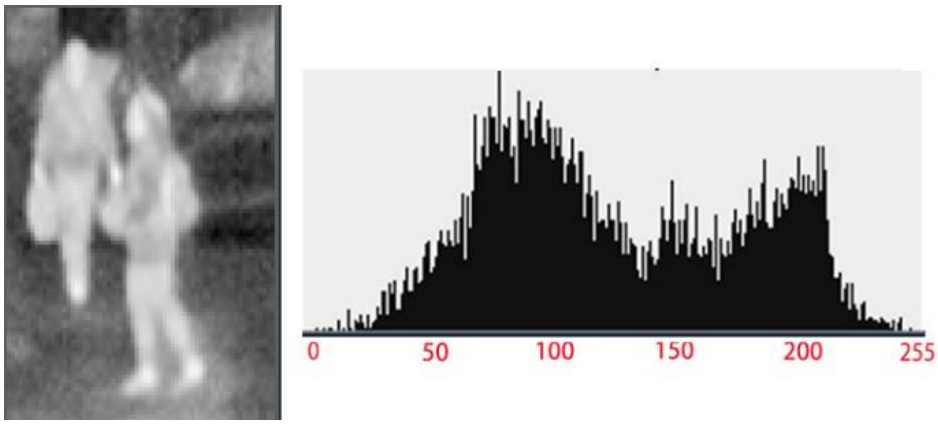
من السمات الضائعة هنا يمكن ملاحظة ضياع الألوان كذلك الحواف الداخلية داخل جسم الكائن كذلك يمكن ملاحظة تأثير الضجيج أو الخلفية المعقدة التي توحي بمناطق حرارية لكنها بالحقيقة ليست الكائنات الهدف. نجري فيما يلي دراسة تحليلية عن صلة الصورة الحرارية بالصورة المرئية وهل يمكن تحسين الصورة الحرارية لتبدو أكثر قبولاً في النماذج المدربة مسبقاً.

يعطينا تحليل الهيستوغرام (Histogram) فكرة عن توزيع المستويات اللونية في الصورة (Gonzalez, 2018) حيث يمثل المستوى 0 في الاحداثي الأفقي اللون الأكثر سواداً يليه مستويات متدرجة الرماديات عددها 255، حيث يمثل المستوى 255 اللون الأكثر بياضاً. وعند كل مستوى من تدرج الرمايات نلاحظ عدد النقاط التي تملك اللون ذاته في الصورة. في الشكل (4-4) يتم اجراء تحليل الهيستوغرام وذلك للمنطقة الحاوية للهدف في تدرج الرماديات للصورة اللونية.



الشكل (4-4) هيستوغرام الصورة المرئية (الباحث)

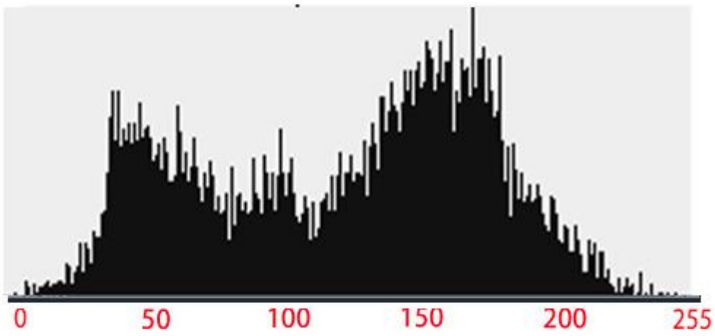
أما في الشكل (4-5) يتم إجراء تحليل الهيستوغرام للمنطقة المحيطة بالهدف في المشهد ذاته لكن للصورة الحرارية.



الشكل (4-5) هيستوغرام الصورة الحرارية (الباحث)

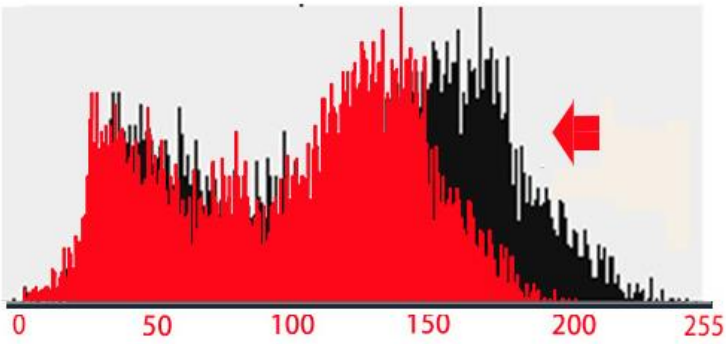
يلاحظ هنا بأن الاختلاف كبير بين توزع مستويات الإضاءة في كلا التحليلين والعائد إلى كون الصور الحرارية لا تأخذ فروقات الإضاءة كما في الصورة المرئية بل تعتمد على فروقات الحرارة.

إن العين يمكنها ملاحظة وجود شخصين في الصورة الحرارية، بالرغم من الاختلاف في توزع سويات الألوان، لذلك نبحث عن أوجه التشابه في معلومات الصورتين في التحاليل المختلفة. في تحليل الهيستوغرام يمكن ملاحظة تشابه بالمعلومات المستخلصة من الصورة المرئية مع مقلوب المستويات المستخلصة من الصورة الحرارية، بحيث إذا قلبنا مخطط الهيستوغرام للصورة الحرارية أفقياً يمكن الحصول على مخطط أكثر تشابها الشكل (4-6).



الشكل (4-6) قلب مخطط الهيستوغرام للصورة الحرارية السابقة (الباحث)

لزيادة نسبة التشابه في مخططنا الناتج مع مخطط الهيستوغرام للصورة المرئية يمكن ضغط المخطط بمنحى المستويات اللونية "تسوية الهيستوغرام Hestogram Equalization" إلى جهة اليسار أي تعتيم كما في الشكل (4-7)



الشكل (4-7) ضغط مخطط الهيستوغرام لجهة اليسار (الباحث)

لتحقيق نسبة التشابه المذكورة وباعتبار القيم ضمن نطاق 8bit يمكن تبديل المستويات اللونية إلى المعكوس في كل بيكسل وفق المعادلة (4-1)، (Parker, 2010):

$$I_{x,y} = 255 - I_{x,y} \qquad (1\text{-}4)$$

إن معالجة انضغاط مستويات الهيستوغرام تشكل معضلة حقيقية بسبب عدم وجود سوى صورة حرارية لدى الكشف، وبالتالي مدى واتجاه قيم الانضغاط غير متوفرة.

في طريقتنا هذه نحاول استخدام قيم الانزياح عن متوسط إضاءة الصورة I_{mean} باعتبار الكثير من الكواشف المدربة تعتمده بطرائق متعددة لتحسين الصورة في مرحلة المعالجة قبل التدريب.

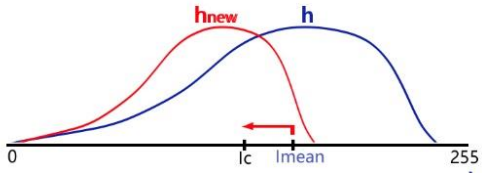
تحسب المعادلة (4-2)، (الباحث) متوسط سطوع أو إضاءة الصورة:

$$I_{mean} = \frac{1}{m.n}\sum_{i=0}^{n}\sum_{j=0}^{m} I_{i,j} \qquad (2\text{-}4)$$

إن نسبة الانتقال اللازم لضغط المستوى تعطى بالعلاقة (4-3)، (الباحث):

$$d = \frac{Imean - I_c}{I_c} \qquad (3\text{-}4)$$

حيث Ic هي مستوى متوسط الإضاءة في الصورة المعدَّلة بالاجراء السابق كما في الشكل(4-8).



الشكل (4-8) مقدار ضغط مخطط الهيستوغرام لجهة اليسار (الباحث)

وتعطي العلاقة (4-4)، (الباحث): معادلة التصحيح هنا لكل بكسل:

$$I_{new} = I_{h(I_{x,y}\,.\,d)} \qquad (4\text{-}4)$$

يجب الأخذ بعين الاعتبار هنا قلب نسبة d حسب اشارتها، حيث الإشارة السالبة تعني تكبير نسبي وليس ضغط، بالتالي الضرب بـ 1/d. لقد تم الاهتمام هنا بالتصحيح حسب الهيستوغرام بقيم الانزياح لمتوسط مستويات الصورة عن المستوى المركزي. قد يفيد تحليل قيم معاملات أخرى مع الهيستوغرام مثل مستوى إضاءة بكسلات الصورة Brightness والتباين Contrast لكن ذلك نظرياً يخضع لظروف التقاط الصورة وسيتم التأكد من درجة الأهمية في التجارب.

أما بالنسبة لاختيار ثوابت ضغط الهيستوغرام ومستوى الإضاءة والتباين فستجرى تجارب متعددة على البيانات القياسية المزدوجة مرئية وحرارية بحيث تكرر التجارب ضمن الظروف المختلفة وفي كل مرة تكرار يعتمد ثابت مختلف للوصول إلى أفضل استجابة لكشف الكائن ضمن كل ظرف.

4-3-2 استخدام سمات الحركة

يتم التعرف على الهدف في كل إطار ببساطة من خلال القيام بالكشف، ولكن ربط الهدف بين الإطارات أمر صعب. هذا هو أحد أكبر التحديات في ملاحقة الكائنات. من الضروري إنشاء اتصال واحد خلال هذه الإطارات، لكن هذا الاتصال يتطلب العديد من أنواع المعلومات للتعرف على الهدف من الأهداف الأخرى بثقة عالية، وهذا ما يستدعي استخدام نماذج الحركة لتقدير موقع الكائن في الإطار التالي ومن أحد أهم الوسائل التي عرضناها سابقاً هي مرشح كالمن الذي يُستخدم باعتباره أحد أكثر الأساليب شيوعاً (MathWorks. 2024)، لتحقيق ملاحق متعدد الكائنات من خلال تقديم أداء قوي في:

- التنبؤ بالموقع المستقبلي للكائن
- تقليل الضجيج الناتجة عن الاكتشافات غير الدقيقة

- تسهيل عملية ربط كائنات متعددة بمساراتها بين الاطارات

تتضمن تقنية ملاحقة الأجسام المتعددة التعامل في وقت واحد مع أشياء متعددة، وتتبع مواقعها، ومراقبة تفاعلاتها. يجب على النظام أيضًا الحفاظ على هوية مسار متسقة لنفس الشيء منذ اكتشافه الأولي، والتعامل مع تغييرات المظهر المفاجئة واحتجابات الكائنات الشديدة. علاوة على ذلك، يحتاج إلى معالجة سيناريوهات معقدة أخرى، مثل دخول الأشياء إلى المشهد والخروج منه.

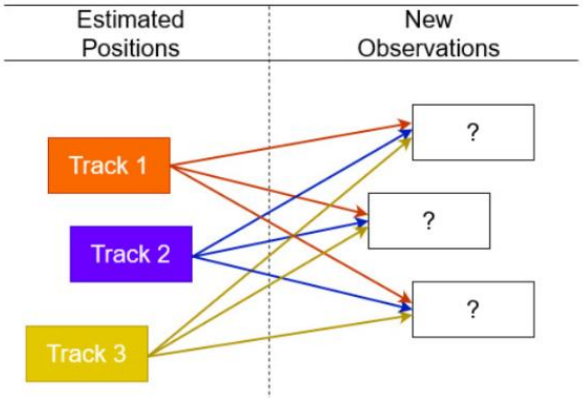
لفهم أفضل لمشكلة ربط المسارات في تتبع متعدد الكائنات، لنركز على الحالات المبسطة التالية في MOT.

لنفترض أنه في الوقت $t - 1$، لدينا ثلاث مسارات موجودة مُسمَّاة بـ Track1 وTrack2 وTrack3. عندئذٍ، في الوقت t، يمكن أن تنشأ مجموعات كشف مختلفة، كما هو موضح بهذه الأمثلة المحتملة:

الحالة الأولى: يكتشف الكاشف ثلاثة أهداف في الوقت الحالي t، وهو نفس عدد المسارات الموجودة في الوقت السابق $t - 1$، كما هو موضح في الشكل (4-9).

الحالة الثانية: اختفاء أحد الأهداف من المشهد، ويحدد الكاشف بدقة وجود هدفين في الوقت الحالي t.

الحالة الثالثة: يظهر هدف جديد في المشهد، ويكتشف الكاشف بشكل صحيح وجود أربعة أهداف في t.



الشكل (4-9) البحث عن الربط (Sadli,2023)

تُظهر الأمثلة المقدمة أعلاه التحدي الكبير المتمثل في ربط المسارات الموجودة بالاكتشافات المقابلة لها في ملاحقة الكائنات المتعددة. من الصعب للغاية تحديد المسار الذي يتوافق مع الاكتشاف الصحيح بدقة. لذلك، هناك حاجة إلى طريقة تحسين فعالة. لإيجاد الحل الأمثل لمشكلة ارتباط المسارات، يمكننا صياغة المشكلة على النحو التالي:

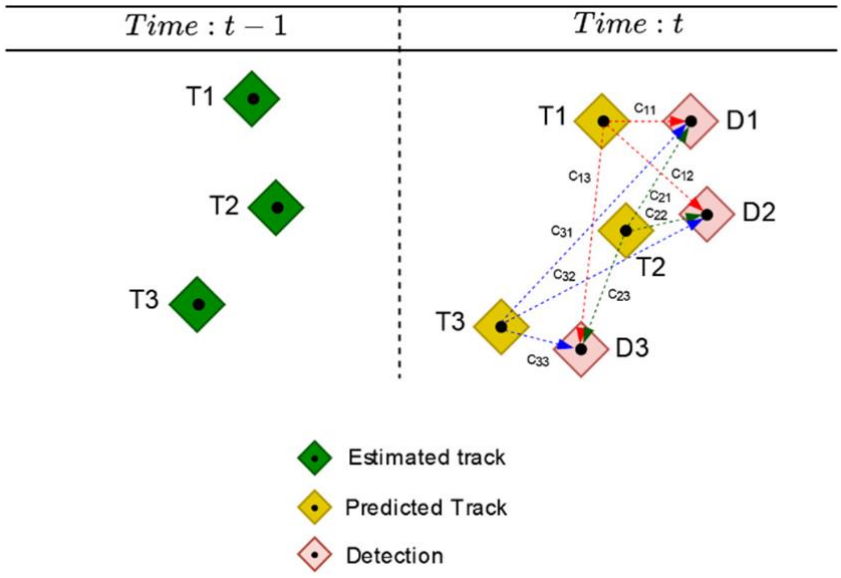
مصفوفة التكلفة $C = [C_{ij}]$، حيث تمثل C_{ij} تكلفة تعيين المسار i للمكتشف j، نوجد المصفوفة $X = [x_{ij}]$ بحيث يكون ناتج ما يلي أصغري معطى بالعلاقة (4-5)، (Sadli,2023):

$$\sum_{i=1}^{n} \sum_{j=1}^{m} C_{ij}\, x_{ij} \qquad (5\text{-}4)$$

المتغير x_{ij} ثنائي، مما يضمن ارتباط كل كشف بمسار واحد فقط وارتباط كل مسار بكشف واحد فقط. لذلك، يجب أن يكون x_{ij} إما 0 أو 1، أي $x_{ij} \in \{0,1\}$.

تُعد مصفوفة التكلفة مكونًا حيويًا في تحسين مشكلة ارتباط المسار في الملاحق متعدد الكائنات. إنها مصفوفة ثنائية الأبعاد حيث يمثل كل عنصر تكلفة تعيين مسار لاكتشاف جديد. بعبارة أخرى، يمثل كل صف من مصفوفة التكلفة مسارًا، بينما يمثل كل عمود اكتشافًا جديدًا. يمكن حساب تكلفة تعيين مسارات موجودة لاكتشافات جديدة في التتبع متعدد الكائنات باستخدام مقاييس مختلفة، مثل المسافة، والتقاطع على الاتحاد (IoU)، وتشابه المظهر. يعد اختيار المقياس المناسب لإنشاء مصفوفة التكلفة أمرًا بالغ الأهمية لتحقيق أداء التتبع الأمثل. يعتمد اختيار المقياس على خصائص الكائنات التي سيتم تعقبها والمتطلبات المحددة لنظام التتبع.

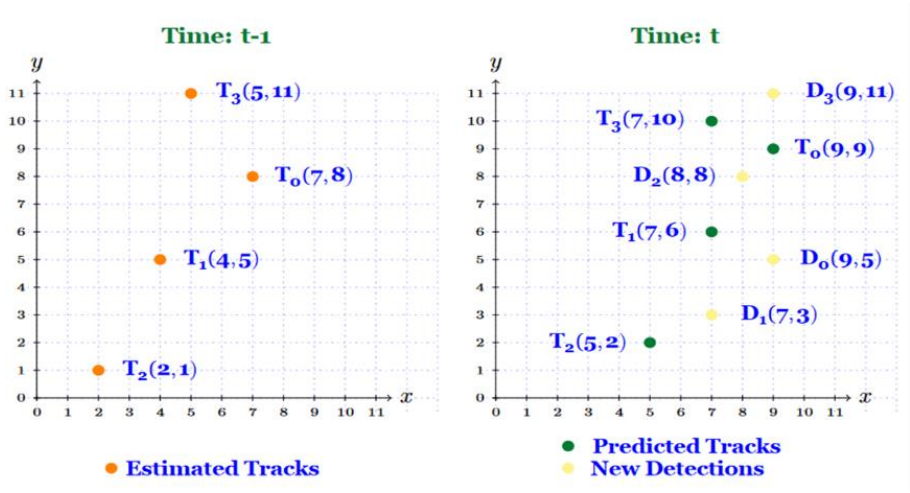
يمكن في كل إطار جديد وباستخدام التقدير الحالي بكالمن توقع مواقع الكائنات في الإطار الجديد التالي، وعند اكتشاف الإطار التالي يتم تقدير احداثيات الكائنات الحالية. في الشكل (4-10)، نورد مثالاً عملياً حيث توجد ثلاث مسارات مقدرة في الوقت $t-1$، بالإضافة إلى ثلاث اكتشافات جديدة وثلاثة مسارات متوقعة في الزمن t.



الشكل (4-10) تقدير الحالة الجديدة للكائن أثناء عملية الملاحقة (Sadli, 2023)

يتم إنشاء مصفوفة تكلفة 3×3 يشار إليها بـ C، والتي تمثل تكاليف كل ارتباط محتمل. يشير C_{12} إلى تكلفة تعيين المسار الأول للكشف الثاني، ويمثل C_{33} تكلفة تعيين المسار الثالث للكشف الثالث. ومن المهم ملاحظة أن مصفوفة التكلفة يتم حسابها بناءً على مواضع المسارات المتوقعة والاكتشافات الجديدة.

بالنظر إلى السيناريو في الشكل (4-11). في الوقت $t-1$، توجد أربعة مسارات مقدرة ممثلة بدوائر برتقالية. وفي الوقت t، لدينا أربع مسارات متوقعة ممثلة بدوائر خضراء وأربعة اكتشافات جديدة ممثلة بدوائر صفراء.



الشكل (4-11) مثال لمعطيات إطارين (Sadli, 2023)

وعندما ننشئ مصفوفة التكلفة لهذا السيناريو، كما هو موضح في الشكل أعلاه، في الوقت t، ستتواجد أربعة مسارات متوقعة T 0 وT 1 وT 2 وT 3 وأربعة اكتشافات D 0 وD 1 وD 2 وD 3.

يري الشكل (4-12) ناتج مصفوفة التكلفة:

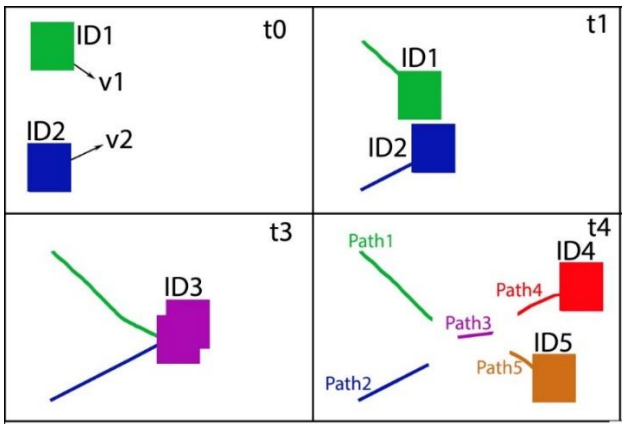
$$\mathbf{C} = \begin{array}{c} \\ T_0 \\ T_1 \\ T_2 \\ T_3 \end{array} \begin{array}{c} \begin{array}{cccc} D_0 & D_1 & D_2 & D_3 \end{array} \\ \begin{bmatrix} 4 & 6.32 & 1.41 & 2 \\ 2.24 & 3 & 2.24 & 5.39 \\ 5 & 2.24 & 6.71 & 9.85 \\ 5.39 & 7 & 2.24 & 2.24 \end{bmatrix} \end{array}$$

الشكل (4-12) مصفوفة التكلفة عند الربط (Sadli, 2023)

بالنسبة لمصفوفة التكلفة الأصغر، تقدم طريقة القوة المسيطرة Brute-force نهجًا مباشرًا لاستكشاف جميع الحلول المحتملة وتحديد الحل الأمثل. بشكل عام، تحتوي مصفوفة التكلفة $n \times n$ على $n!$ حلول ممكنة، على سبيل المثال، تحتوي مصفوفة التكلفة 3×3 على 6 حلول محتملة، وتحتوي مصفوفة التكلفة 4×4 على 24 حلًا، وهكذا. ومع ذلك، مع نمو مصفوفة التكلفة وزيادة مساحة الحل بشكل عاملي، تصبح طريقة القوة المسيطرة مكلفة حسابيًا بسرعة وهي بالتأكيد غير مناسبة للتطبيقات في الزمن الحقيقي. إحدى الخوارزميات المستخدمة على نطاق واسع لهذا الغرض هي الخوارزمية المجرية Hungarian algorithm. تعتمد هذه الخوارزمية على مبدأ التحسين التكراري ويمكنها حل المشكلة في زمن معالجة $O(n^3)$ "مكعب معالجات كافة النقاط n" وهذا يجعلها أسرع بكثير من طريقة القوة المسيطرة للقيم الأكبر من n. في الأعوام الماضية بدأت الأبحاث المتعلقة بـ MOT تعتمد على استخدام نماذج الحركة بطرائق مختلفة. يصف البحث (Pramod,2018) عملية التقدير بالمهمة جدًا في نظام المراقبة، وفيها تتم معرفة مواقع إطارات الهدف بواسطة مرشح كالمان عن طريق التعرف على الألوان،

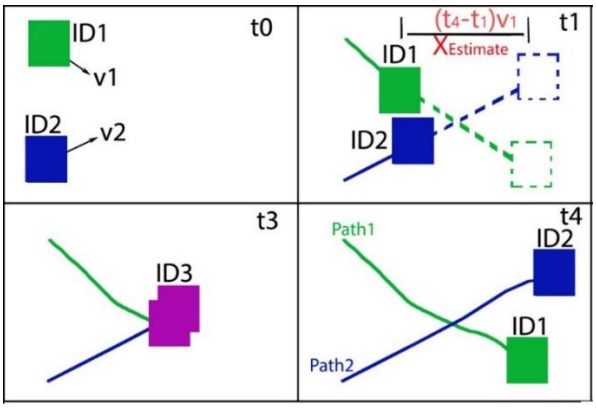
وتقدر المواضع الجديدة في الإطار التالي من خلال ادخال المواضع المكتشفة حاليا والمواضع المقدرة سابقاً. في طريقة CenterTrack التي اقترحها (Zhou,2020)المستندة إلى CenterNet هي إطار عمل ملاحقة قائم على النقاط، حيث يتم تمثيل كل كائن بنقطة واحدة في مركز المربع المحدد له ثم تؤدي ملاحقة هذه النقطة المركزية عبر الزمن إلى تبسيط اكتشاف حالة الملاحقة وارتباط الكائن المستهدف عبر الزمن.

يبين الشكل (4-13) رسما مفصلاً لتصرف نظام ملاحقة عند مراقبة متحركين متداخلين، ويلاحظ ضياع المسارات ونشأة مسارات جديدة.



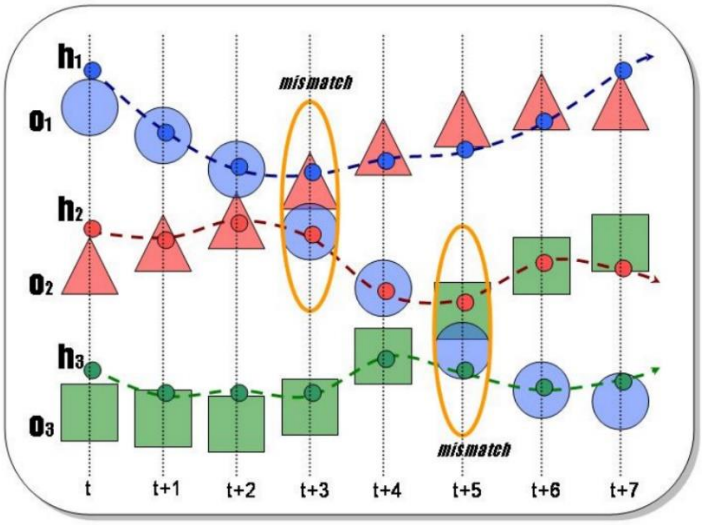
الشكل (4-13) تصرف نظام ملاحقة عند مراقبة متحركين متداخلين (الباحث)

يبين الشكل (4-14) رسما مفصلاً لاستخدام المرشح في حل معضلة التداخل الناتج عن تقاطع متحركان.



الشكل (4-14) التداخل الناتج عن تقاطع متحركين مع نموذج حركة (الباحث)

بالرغم من استخدام المرشح لحل الكثير من مشاكل الملاحقة لكن تواجه معظم خوارزميات الملاحقة المتعددة MOT العديد من المشاكل ولعل أهمها هو تغير نموذج الحركة من قبل الكائن، مثل تغيير الكائن لاتجاهه فجأة أو تغيير السرعة أو التسارع وهكذا. يبين الشكل (4-15) إحدى المشاكل المتعلقة بتغيير الاتجاه وفشل الملاحق في الأزمنة t+3 وt+5 ففي الزمن t+3 تم تبديل الكائنان المثلثي والدائري بسبب تنبؤ المرشح الخاطئ، أما في الزمن t+5 فهناك فشل آخر وتبيل الكائنات الدائري والمربع للسبب السابق ذاته

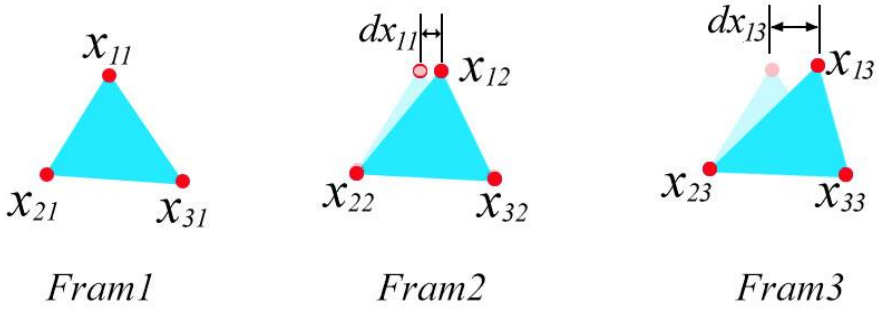


الشكل (4-15) فشل المرشح بسبب تغيير اتجاه الحركة (Guo,2022).

لابد من الذكر أن من أسباب فشل المرشح في عملية الملاحقة هو تغير نقطة المركز تبعاً لتغير شكل الكائن، كمثال حالتنا حيث هناك عدم ثباتية بشكل الكائن الحراري.

طريقة الحل المقترحة

للمساهمة في تقليل الأخطاء يجب الانتباه لشكل الكائن بحيث تتم مراقبة تحولات هذا الشكل بالإضافة لمراقبة موقعه، وهذا ما يسمح لخوارزمية الملاحقة بالتمييز بعد التداخل باستخدام بصمة الشكل بالإضافة لنموذج الحركة، لذلك نقترح مراقبة عدة نقاط في الكائن المكتشف. يوضح الشكل (4-16) تمثيلا لحركة كائن متغير الشكل مع الزمن (غير صلد) حيث تتم مراقبة أكثر من نقطة.



الشكل (4-16) مراقبة أكثر من نقطة لكائن متغير الشكل (الباحث)

بفرض أن الكائن بين الإطار الأول والثاني قد تحرك بمقدار Δx، وفي المثال المعطى يتغير موقع x_{12} عن المتوقع بمقدار dx_{11} لذلك يمكن كتابة معادلات فاصلة الموقع (4-6)، (الباحث):

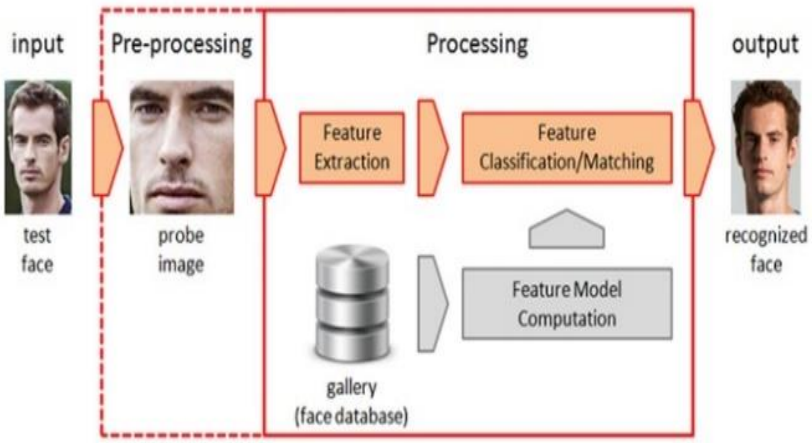
$$\begin{aligned} x_{12} &= x_{11} + \Delta x + dx_{11} \\ x_{22} &= x_{21} + \Delta x + dx_{21} \\ x_{32} &= x_{31} + \Delta x + dx_{31} \end{aligned} \qquad (6\text{-}4)$$

وهكذا بالنسبة للإطار التالي. إن تقدير القيمة x_{12}, y_{12} واحداثيات النقاط الأخرى في الطور الثاني ثم التالي سيؤسس لتقدير مجموعة نقاط تمثل بصمة الكائن تدخل بصورة احتمالية عالية لإعادة تعريف وتمييز الكائن بعد انتهاء التداخل، بحيث نكون هنا قد أخذنا سمات الحركة لهذه النقاط بعين الاعتبار، ومن كون هذه النقاط تتحرك بشكل مترابط بالنسبة للكائن الواحد فسيكون تمييز الكائن أسهل بعد التداخل

4-3-3 إعادة تعريف الكائن ضمن الصور الحرارية

يعتبر التعرف على الوجه أهم ما يمكن الاعتماد عليه في تمييز الشخص. ويبين الشكلان (4-17) و (4-18) طرائق ومراحل التعرف على الوجه Face Recognition عن طريق مقارنة خصائص الوجه المكتشف Face Detection مع الخصائص المشتركة Feature Classification/Matching للعديد من الصور العائدة لذات الشخص في مجموعة البيانات Face dataset والتي تم تدريب النظام مسبقاً عليها (Olszewska, 2016).

الشكل (4-17) التعرف على الوجه. (Olszewska, 2016)



الشكل (4-18) مراحل التعرف على الوجه. (Olszewska, 2016)

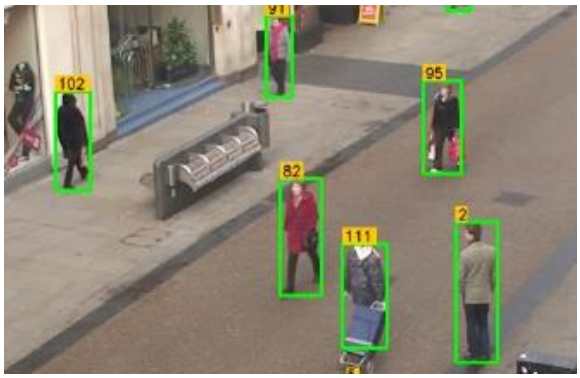
تشمل الخصائص ملامح الوجه المكتشفة، على سبيل المثال في الخوارزميات البسيطة: أعلى الرأس، زاوية الحاجب الأيمن الزاوية، الزاوية اليمنى للحاجب الأيسر مركز العين اليمنى، الزاوية اليمنى للعين اليسرى، وهكذا ...

من أهم العوامل التي تؤدي لعدم نجاح التعرف على الشخص هو الحجم الصغير لمساحة الوجه المكتشفة، كأن تكون مسافة التقاط الصور كبيرة وهي الحالة العامة لدى ملاحقة عدة أشخاص. لذلك تلجأ الخوارزميات هنا إلى الاكتفاء بكشف الجسم ووضع رقم تعريفي له.

تعطى الهوية التعريفية للكائنات ضمن الصورة الحاوية على كائنات متعددة كرقم تعريفي ID:

id=1, id=2, id=3, ….

غالباً ما يعتمد بترتيبه على طريقة المعالجة وزمن الظهور للهدف في المشهد كما في الشكل (4-19)



الشكل (4-19) الرقم التعريفي للكائن (Serpush, 2020)

تكتفي خوارزميات الملاحقة هنا بالرقم التعريفي لكل جسم مكتشف ضمن كل صنف بحيث تحاول المحافظة على هذا الرقم لذات الجسم مهما مرت مراحل الملاحقة بصعوبات كالتداخل والانقطاعات وخروج الأجسام من المشهد وعودتها ثانية، مما يدخل الملاحقة في مشاكل جمة تمثل تحديات تتسابق البحوث المختلفة على حلها، فمنها من يعتمد على الملامح ومنها ما تأخذ في الاعتبار بدلاً من ذلك تفاصيل الملابس وألوانها وسمات شكل الجسم والسمات الأخرى المتعلقة بالمظهر، ومنهم من ذهب إلى خصائص حركته في المشي(Simhi,2020).

في الصورة الحرارية نلاحظ أنها عبارة عن تمثيل لاختلافات درجات الحرارة المنبعثة من الأجسام المختلفة المحتواة في الصورة، حيث يكون الجسم المصدر للحرارة ممثلاً بمناطق اضاءة أكثر من الجسم الأقل اصداراً للحرارة، وهو ما يمكن من التمييز بين الأهداف الحية المصدرة للحرارة والخلفية، وخصوصاً في الظلام.

يعرض الشكل (4-20) صورة حرارية لمجموعة من الأشخاص، ويلاحظ بأن النقاط الممثلة للأشخاص تكون أشد اضاءة من باقي أجزاء الصورة. وبالنسبة لعملية الكشف تكون بتمييز هذه النقاط المكونة للهدف عن نقاط الخلفية وإيجاد احداثيات مركز هذا الجسم.

الشكل (4-20) صورة حرارية لمجموعة أشخاص (الباحث)

لو نظرنا بالعين المجردة إلى الأشخاص الموجودين بالصورة نلاحظ أننا يمكننا كشف وتمييز هذه الكائنات عن باقي الأهداف، ولكن ملامح الأشخاص وهيئتهم مختلفة عما هو موجود في الصور المرئية RGB بالتالي لا يمكن تمييزهم عن بعضهم وتعريف هويتهم.

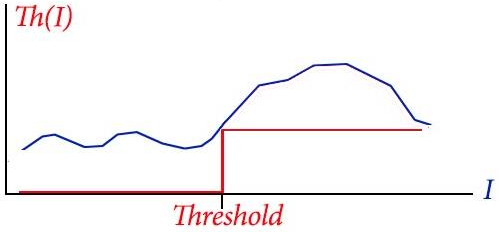
كما مر معنا في الدراسات المرجعية السابقة هناك تركيز على مشاكل انخفاض جودة الكشف والأهم من ذلك الصعوبة الكبيرة في تصنيف الكائنات ضمن الصنف الواحد أي "التعرف على الهوية" لذلك تقدمنا بطرح طريقة تساعد في إيجاد هوية افتراضية للشخص المكتشف وهي بحد ذاتها تساعد مع أي خوارزمية للملاحقة في دوام التعريف لذات الكائن في الفيديو الحراري.

الطريقة المقترحة:

في البداية يتم تحسين الصورة الحرارية وتحضيرها للمعالجة ثم يُستخدم كاشف كائنات ذو وثوقية مرتفعة YOLO، وعند تطبيقه على الصورة يلاحظ بأن النموذج قد كشف الأشخاص الموجودين وأحاطهم بمستطيل إحاطة كما في الشكل (4-21). للتغلب على مشكلة طمس المعالم في الصورة الحرارية يتم البحث عن الملامح الآنية للكائن المكتشف، لذلك تقوم الخوارزمية بالبحث داخل قناع كل هدف على سمات تستطيع من خلالها تمييز الأشخاص عن بعضهم. تم هنا الاعتماد على العلاقة (4-7) المستخدمة في التجزئة "تابع العتبة" (Gonzalez, 2018) وهو التابع الذي يحول الصورة إلى لونين أبيض وأسود فقط كما في الشكل (4-22) الذي يعبر عن تابع العتبة باستخدام الهيستوغرام.

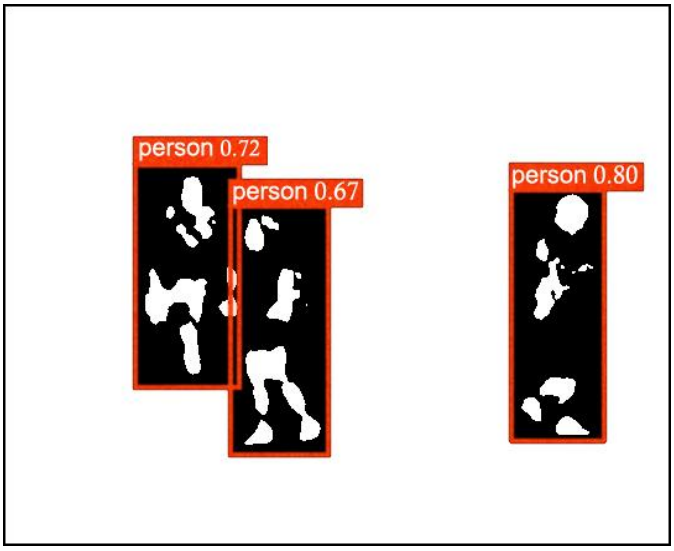
$$Th(I) = \begin{cases} 0 & if\ I \leq Threshold \\ 1 & if\ I > Threshold \end{cases} \qquad (7\text{-}4)$$

الشكل (4-21) كشف الكائنات البشرية بـ YOLO-V8 (الباحث)



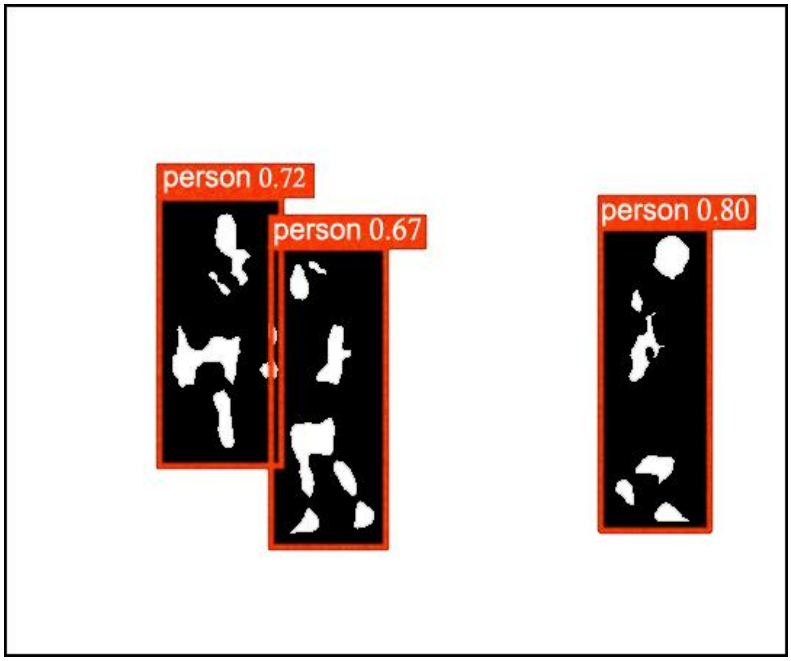
الشكل (4-22) تابع العتبة Threshold، المنحني باللون الأزرق يمثل قيمة الاضاءة للنقطة (الباحث)

تم اختيار عتبة Threshold متكيفة مع Imean متوسط إضاءة الصورة في المنطقة المحصورة ضمن قناع الكائن الهدف "مستطيل الكشف" ويؤخذ بعين الاعتبار الحصول على عدد من الأهداف الجزئية ضمن المجال [5,10] كما في الشكل (4-23).



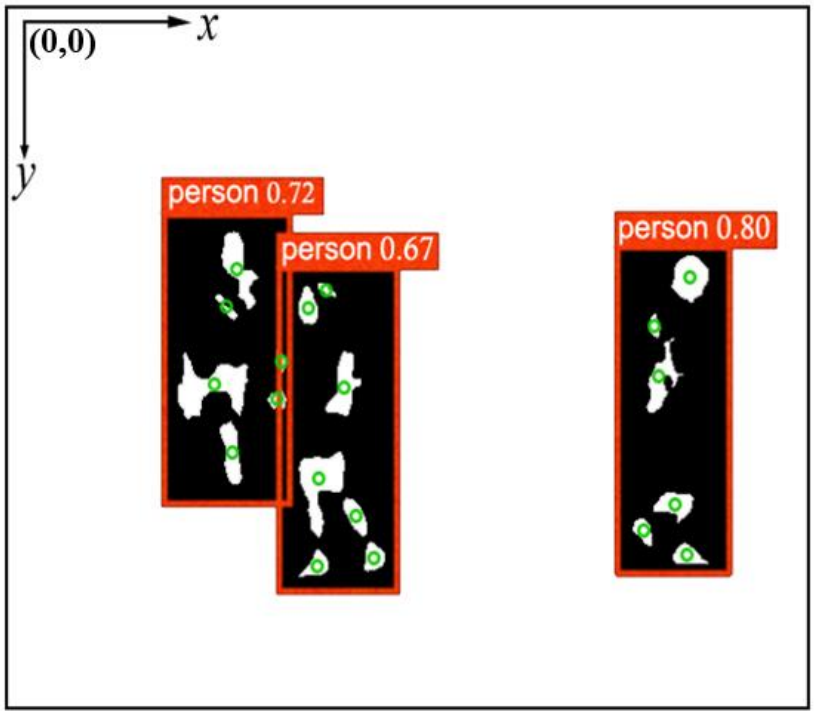
الشكل (4-23) تجزئة الكائنات ضمن الأقنعة بتابع العتبة (الباحث)

الآن يتم حذف الأجزاء الصغيرة بالعمليات المورفولوجية. (Gonzalez, 2018) كما في الشكل (4-24).

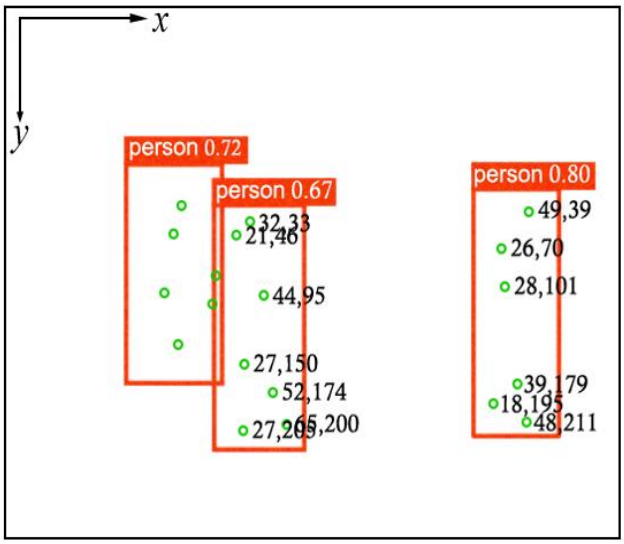


الشكل (4-24) عمليات مورفولوجية لإزالة الأجزاء الصغيرة (الباحث)

ان الناتج الآن هو ثلاثة أهداف بتفاصيل مختلفة، يمكن تعريف هذه التفاصيل ضمن مصفوفة "مصفوفة الأجزاء الداخلية" لكل هدف كما في الشكل (4-25)، تحوي كل مصفوفة بشكل أساسي على مواقع هذه الأجزاء وهي تمثل الاحداثيات الديكارتية بالنسبة للزاوية العليا اليسرى لمستطيل الاحاطة"، يبين الشكل (4-26) الاحداثيات النسبية لأجزاء الهدفين الأيمن والأوسط/ مع العلم أنه تم حذف الاحداثيات على الكائن الأيسر منعاً لتراكب الاحداثيات.



الشكل (4-25) تحديد مراكز الأجزاء (الباحث)

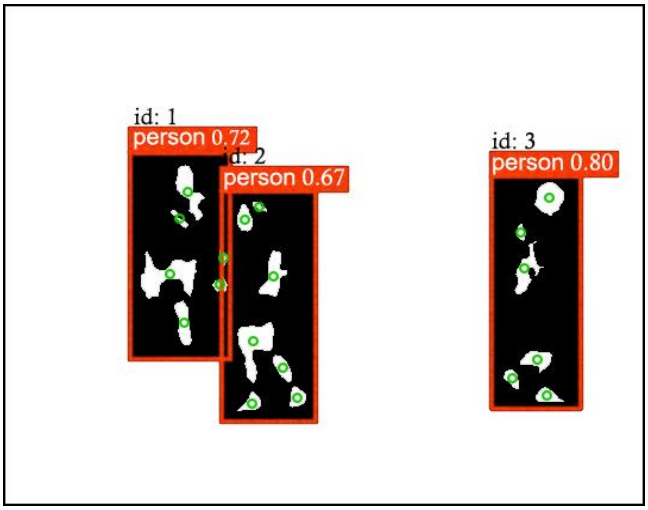


الشكل (4-26) مواقع (الاحداثيات النسبية) لأجزاء الهدفين الأيمن والأوسط (الباحث)

الآن بإمكاننا تعريف هوية كل هدف بحيث التعريف سيعبر عن رقم مصفوفة الأجزاء الداخلية Person n:

$$Person2 = \begin{bmatrix} (32,33) \\ (21,40) \\ (44,95) \\ (27,150) \\ (52,174) \\ (65,200) \\ (27,205) \end{bmatrix} \qquad Person3 = \begin{bmatrix} (49,39) \\ (26,70) \\ (28,101) \\ (39,179) \\ (18,195) \\ (48,211) \end{bmatrix}$$

يلاحظ أن أبعاد الصفوف غير متساوية، الأمر الذي يدخل كسمة للهدف، ويبين الرسم في الشكل (4-27) تعريف هوية كل هدف ومراكز الأجزاء الداخلية التي تم توصيفها في المصفوفات السابقة، ويلاحظ أن أبعاد الصف يحدد حسب ناتج عمليات التجزئة والمورفولوجي.

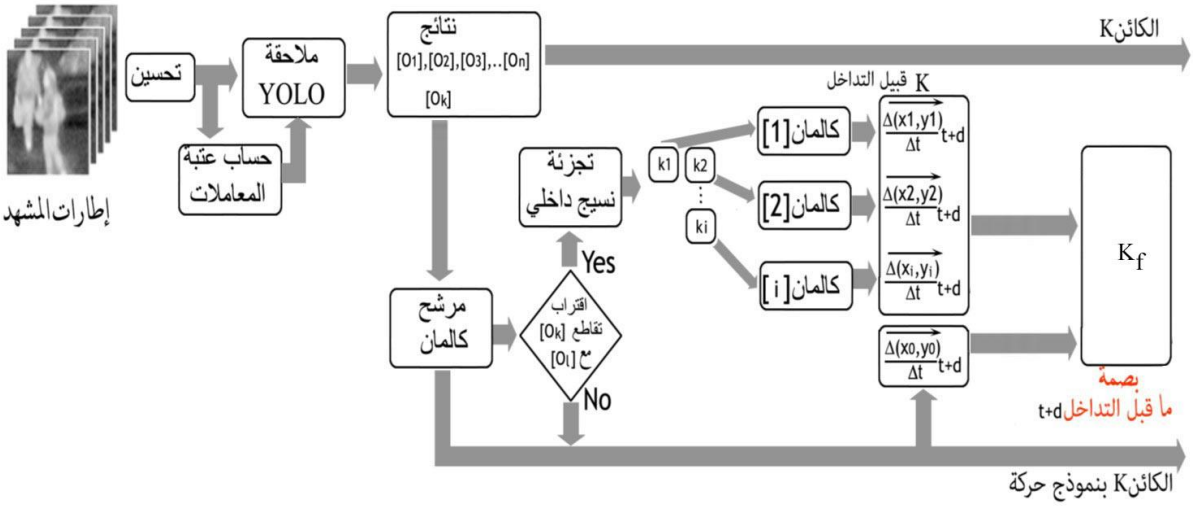


الشكل (4-27) تعريف هوية كل هدف (الباحث)

يلاحظ من قيم عناصر مصفوفة بصمة كل كائن، بأن هذه القيم بمجملها تشكل عنصراً فريداً (بصمة) سيساهم هذا العنصر باستعادة المسار الذي كان قبل التداخل.

4-3-4 المنهجية الأساسية للخوارزمية المقترحة

تمثل الخطوات المدروسة السابقة أساساً للطريقة المقترحة، ويمكن ترتيب هذه الخطوات ضمن مخطط مسلكي كما يبين الشكل (4-28) وذلك لتعامل الخوارزمية قبل التداخل من تحسين وموافقة الصورة الحرارية ومن ثم كشف الكائنات(O) ومن ثم تجزئة هذه الصورة ومتابعة أجزائها K بكالمن وتشكيل البصمة للكائن بمعاملات المرشح.



الشكل (4-28) مخطط مسلكي لتعامل الخوارزمية قبل التداخل (الباحث)

1-4-3-4 تحسين سمات الصورة الحرارية

في البداية يتم قلب التدرج الرمادي للصورة وتخفيض الضجيج فيها بمرشح تمرير منخفض قياس 3×3 ثم يتم حساب الإضاءة المتوسطة I_{mean} ومن ثم الهيستوغرام وتحسب نسبة الانتقال d اللازم لضغط الهيستوغرام، وبعدها يتم الحصول على الصورة الجديدة I_{new} من معادلة التصحيح "حاصل جداء الصورة القديمة بمعامل الضغط المستنتج من الهيستوغرام المضغوط" كما مر معنا في الفقرة (4-3-1).

2-4-3-4 تعديل معاملات الكاشف

يتم هنا استخدام الكاشف YOLO V8 لميزاته المذكورة في الفقرة (3-3-2)، ونحاول تكييف هذا الكاشف بتعديل معاملاته لخدمة كشف أفضل. يتم في الإطارات الأولى للعمل، تحديد متوسط الإضاءة للخلفية حسب ما يلي:

في العلاقة (4-8) يطرح ناتج الكشف في الشكل (4-29-b) من الصورة الأصلية في الشكل (4-29-a) والناتج هو الصورة في الشكل (4-29-c).

$$I_c = I_a - I_b \qquad (8\text{-}4)$$

في العلاقة (4-9)، (الباحث) يتم حساب متوسط اضاءة الخلفية $I_{Back\ mean}$.

$$I_{Back\ mean} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} I_{c\ i,j}}{N.M} \ : Ic_{i,j} \neq 255 \qquad (9\text{-}4)$$

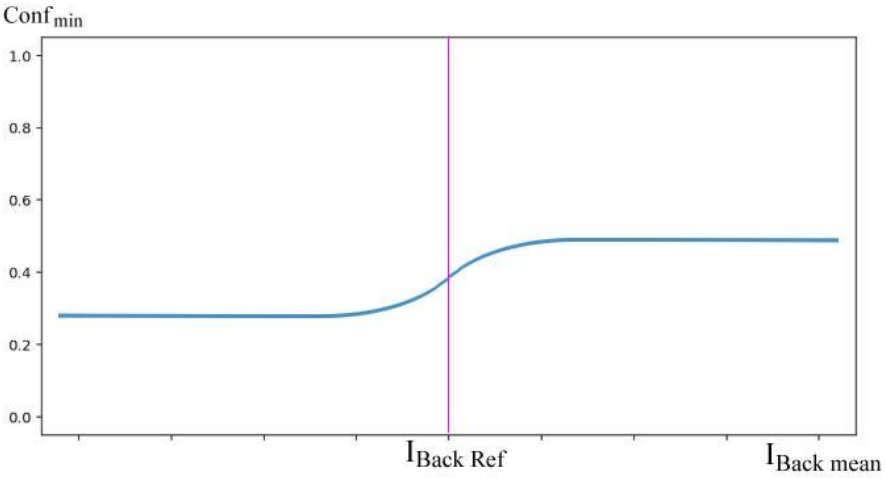
الشكل (4-29) الصورة c ناتج الكاشف ابتداء من الصورة a في الطريقة المطورة (الباحث)

إن تحديد متوسط اضاءة الخلفية سيساهم بتحديد وقت المراقبة هل هو ليلا أو نهارا، حيث أثناء تحليلنا للصور الحرارية وجدنا أنا هناك فوارق مهمة في الصورة النهارية عن الليلية حيث تكون الخلفية في الأولى متأثرة بحرارة الشمس وتعكس طيفاً حرارياً يزيد في تعقيد المشهد بالإضافة أنه سيساهم في تغير متوسط الإضاءة بالتالي التأثير على جودة الكاشف إذا تم التعديل على معاملاته، لذلك يمكن التعديل على معاملات الكاشف بشكل نسبي حسب قيمة $I_{Back\ mean}$. كذلك يتم اعتماد قلب الصورة فقط في الصور الليلية وذلك لتناسب المشهد الليلي المقلوب مع المنظر النهاري أكثر.

عادة ما يتم اختيار معامل ثقة الكشف للكاشف Confidence بقيمة 0.5. في الخوارزمية المقترحة يتم تعديل هذا المقدار بحسب $I_{Back\ mean}$. لقد اعتمد في ذلك على مجموعة من التحاليل للصور الحرارية ومقارنتها مع مجموعة من الصور المرئية لذات الكائنات في مجموعة البيانات المختبرة، وبينت نتائج الكشف للمجموعتين بأن أداء الكاشف الحراري في الليل يستطيع التغلب على الكاشف الحراري نهارا، وخصوصاً للأهداف الأصغر وتعليل ذلك بأن البصمة الحرارية ليلاً أكثر وضوحاً وتمايزاً بسبب انخفاض حرارة الخلفيات. لذلك تم ربط معامل ثقة الكاشف بمستوى اضاءة الخلفية من خلال العلاقة (4-9)، (الباحث) التي تم استنتاجها من تابع التفعيل كما سنرى لاحقاً وتم تحديد ثوابتها تجريبيا:

$$\text{Conf} = 0.3 + \frac{0.2}{1+e^{\frac{I_{Back\ mean} - I_{Back\ Ref}}{25}}} \qquad (10\text{-}4)$$

يلاحظ في الشكل (4-30) رسماً بيانياً للتابع المعتمد

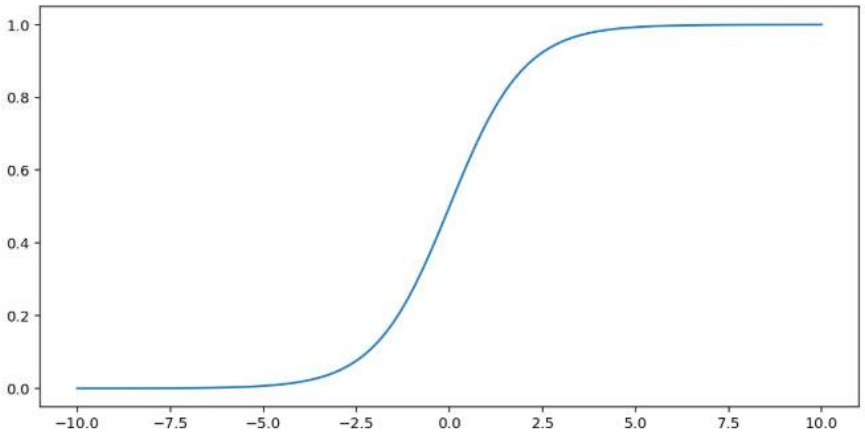


الشكل (4-30) تابع معامل الثقة في الطريقة المطورة (الباحث)

تم الوصول إلى العلاقة السابقة بالتعديل على تابع التفعيل ReLU الشكل (4-31) المستخدم في الشبكات العصبونية ويعطى بالعلاقة (4-10)، (Machine learning, 2024):

$$F(x) = \frac{1}{1+e^{-x}} \qquad (11\text{-}4)$$

بعد ذلك ستتم الملاحقة بنموذج ملاحقة قوي لكافة الأهداف، وللتغلب على مشكلة التدريب (وقت وعتاد وامكانيات تقنية)، تم استخدام الملاحق DeepSORT.



الشكل (4-31) تابع التفعيل ReLU (Machine learning, 2024)

4-3-4-3 تقدير الحالات التالية بمرشح كالمن

يتم التقدير باستخدام مرشح كالمن لاستقراء الحالة التالية (التنبؤ) بحركة الكائنات المختلفة ومعرفة توقيت التداخل "يترجم كعدد إطارات" بين أي كائنين مثلاً O_k, O_l، فإذا كان عدد الإطارات أقل من حد مرجعي معين يتم حساب زمن التلاقي والانفصال بنموذج الحركة لكل إطار.

4-4-3-4 تجزئة الكائن ومراقبة النسيج الداخلي

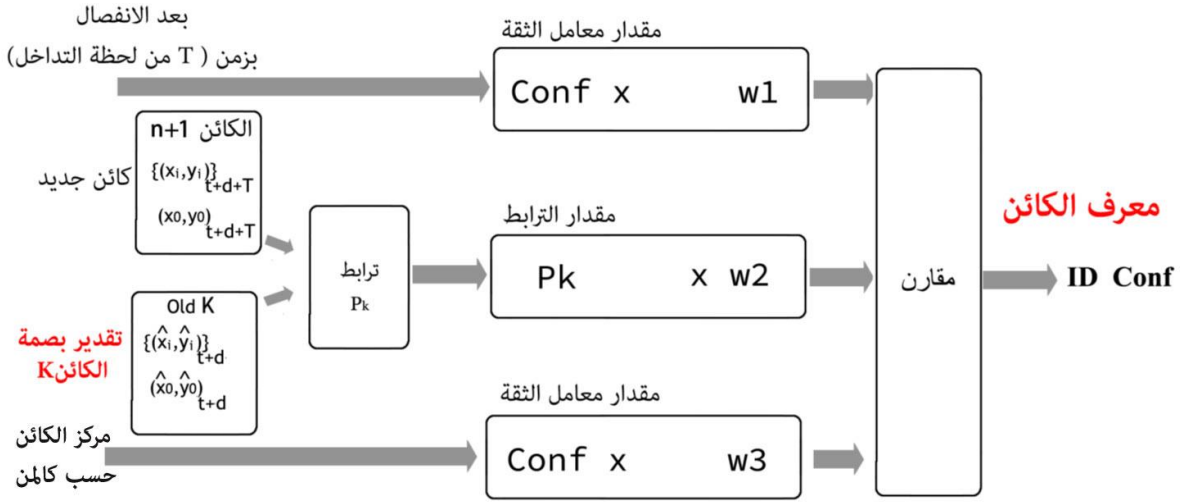
أثناء انتظار التلاقي يتم تجزئة الهدف مثلاً O_k المرشح للتداخل بعتبة تكيفية للحصول على أجزاء كنسيج داخلي متميز، لاحظ الشكل (4-28) الذي يبين الأجزاء $K_1, K_2, \dots, K_i$ التابعة للكائن O_k. وتتم ملاحقة مراكز هذه الأجزاء وتقدير وتسجيل مساراتها، بمرشح كالمن لغاية حصول التداخل.

4-4-3-5 تقدير بصمة الكائن بعد التداخل

إن ناتج تقدير مراكز الأجزاء الداخلية للكائن ستمثل تقديرا لبصمة الكائن بعد زمن T هو زمن التداخل باستخدام النموذج الديناميكي للحركة، وقد استخدم هنا تأخير زمني مقداره d طيلة فترة استخدام نموذج الحركة للتقدير وهو مقدر بحيث يكون أكبر من زمن التداخل المتوقع.

4-4-3-6 استنتاج الكائن الهدف ما بعد التداخل

تنتقل الخوارزمية الآن إلى المرحلة الرابعة S4 أي ما بعد التداخل، يتم فيها نشوء كائنات جديدة وضياع الكائنات الناتجة عن التداخل والشكل (4-32) يعطي توضيحاً لتصرف الخوارزمية في مرحلة ما بعد التداخل. بعد عملية الكشف والملاحقة لعدة إطارات تقوم الخوارزمية بمقارنة بصمة كل كائن جديد في الزمن t+d+T مع معاملات التقدير السابقة للعنصر المرشح للتداخل المحسوبة في الزمن t+d.

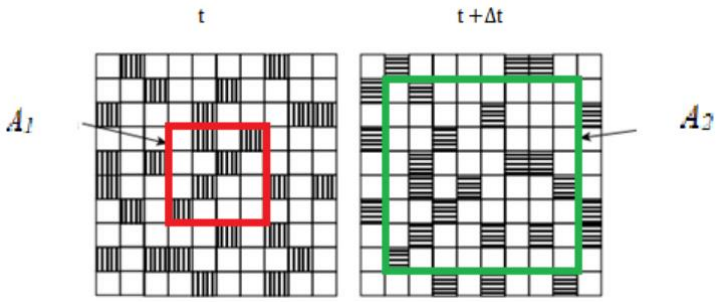


الشكل (4-32) توضيح لتصرف الخوارزمية في مرحلة ما بعد التداخل (الباحث)

من أجل عملية المقارنة في الصور تستخدم الكثير من الخوارزميات تابع الترابط المتصالب Cross correlation بين معلومات الهدفين (Gao, 1988) و (Wieneke,2017) و (Meng, 2021) حيث من المحتمل وجود إزاحة بين الموقع الحقيقي للهدف بعد التداخل وموقع الهدف المتوقع قبل التداخل. لذلك يتم اجراء إزاحة t وتحسب قيمة الترابط (مجموع متوسطات حاصل ضرب القيمتين من الصورة الأولى ومن الصورة الثانية) ومن ثم تزاح بقيمة أكبر وهكذا إلى نهاية الحدود. وفي النهاية نأخذ القيمة العليا كخرج لتابع الترابط المتصالب (Elliott, 2022) لاحظ الشكل (4-33)، وتعطى قيمة الترابط بالعلاقة (4-11)، (Elliott, 2022):

$$C\ \ (r,s) = \sum_{i=0}^{D-1} \sum_{j=0}^{D-1} A_1(i,j) A_2(i+r, j+s) \qquad (11\text{-}4)$$

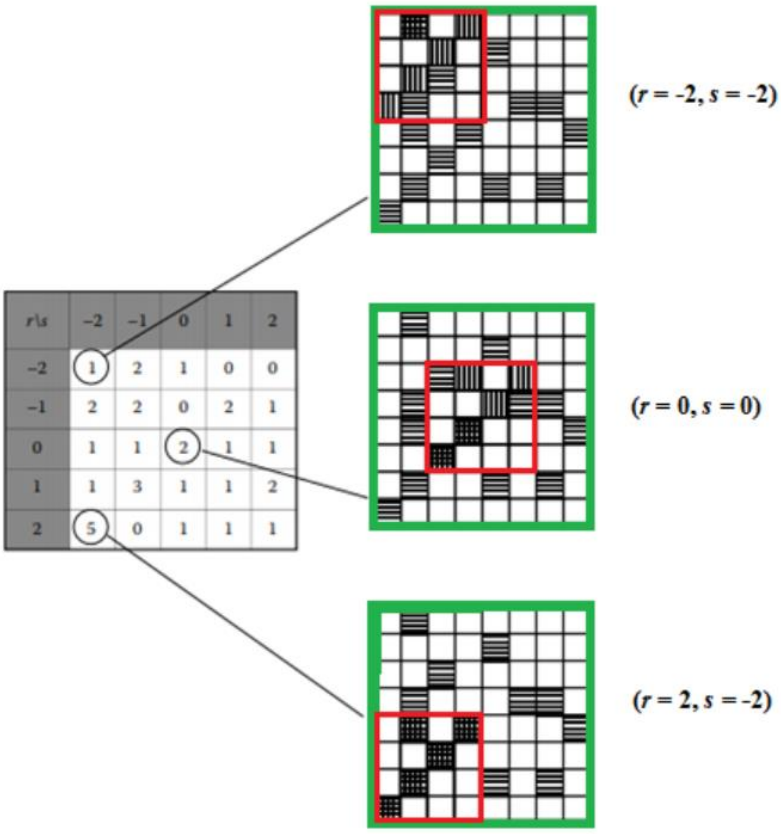
حيث r,s احداثيات مركز المربع الأحمر "الازاحة" الذي نبحث عن مدى تطابق محتوياته مع محتوى المربع الأخضر. يبين الشكل (4-33) صورتين A1 وA2 ويراد الوصول إلى مطابقة A1 مع ما يقابلها من A2.



الشكل (4-33) مطابقة صورتين (Elliott, 2022)

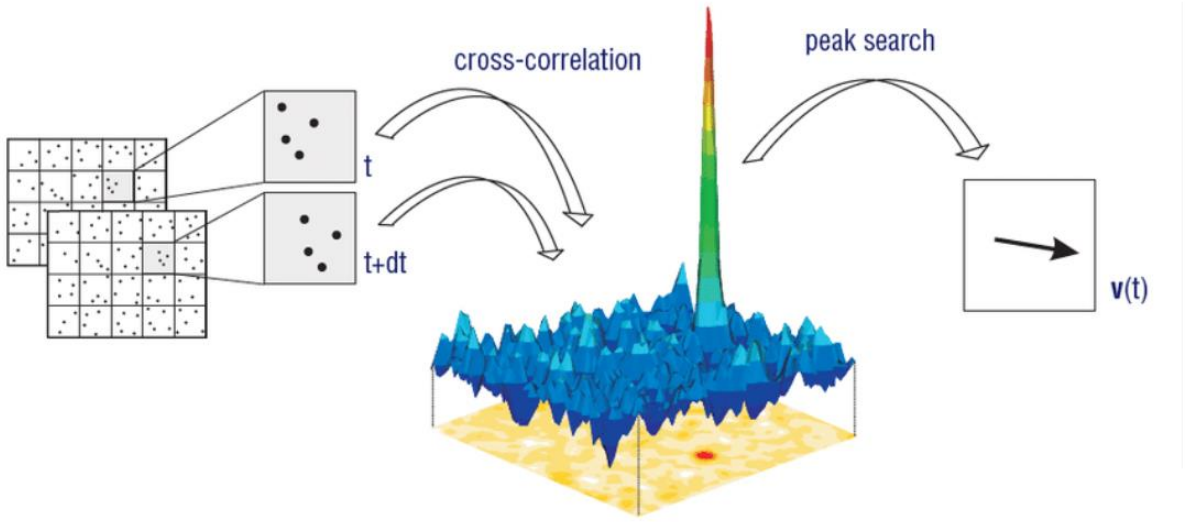
بتطبيق تابع الترابط المتصالب يتم الوصول للمطلوب وتكون الإزاحات كما في الشكل (4-34) هي:

(r = 2, s = -2) حيث عن هذه القيم تكون عندما قيمة الترابط أعظم ما يمكن.



الشكل (4-34) اختيار معاملات أعلى ترابط (Elliott, 2022)

يرينا الشكل (4-35) قيم تابع الترابط في مختلف نقاط الصورة عند البحث عن هدف ما في صورة مأخوذة بعد زمن معين dt



الشكل (4-35) تمثيل ثلاثي الابعاد لقيم تابع الترابط في مختلف نقاط الصورة (Wieneke,2017)

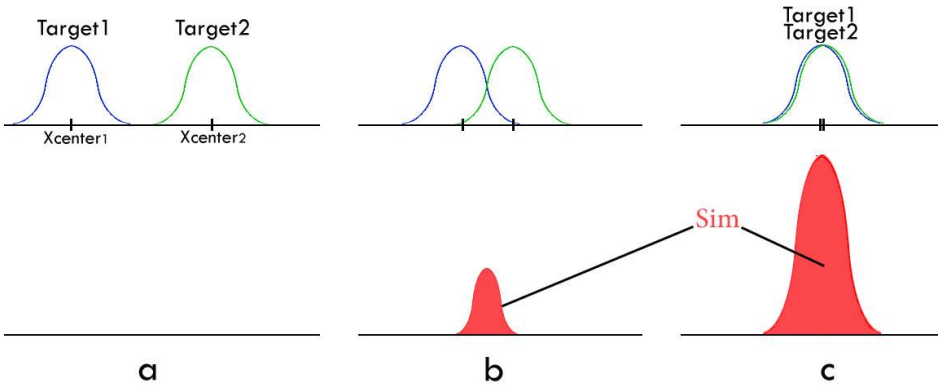
ان استخدام تابع الترابط المتصالب لكشف ترابط الكائنات الهدف قبل وبعد التداخل قد يأخذ تكلفة زمنية كبيرة وخصوصاً في الصور المتعددة الكائنات بسبب الحسابات الضخمة لكافة نقاط الهدف أو أجزاؤه، لذلك تم اتباع طريقة أخرى مستندة على تابع الترابط المتصالب لمراكز الأجزاء وجوارها.

إن اعتماد المراكز فقط سيؤدي لفقدان الترابط وذلك بسبب حتمية إزاحة الهدف ولو بعدة نقاط. لذلك يُستخدم التوزيع الطبيعي لاحتمالية وجود الهدف قبل وبعد التداخل، وذلك كما ذكرنا في الفقرة 3-5-3 بالتالي من الضروري استخدام التوزيع الغاوصي الذي سبق واستخدم في مرشح كالمن، حيث يمكن بتكبير مقدار الانحراف، الوصول لنتائج مهمة وموثوقة أكثر. يبين الشكل (4-36) الطريقة المستخدمة حيث من مصفوفة بصمة الكائن يُشكل مصفوفة موسعة تشمل مركز الكائن وجوارها بعرض 4σ وستكون محصلة مضاريب القيم من الكائن الأول بالقيم المقابلة من الكائن الثاني هي قيمة التشابه Sim كما في العلاقة (4-12)، (الباحث)، وتجرى العملية على المحورين الاحداثيين باستخدام المتحولين i على x و j على y.

$$Sim = \sum_j \sum_i Target1_i . Target2_i \qquad (13\text{-}4)$$

حيث Target1 هي احتمالية وجود مركز الهدف عند الاحداثي (i, j).

في الشكل (4-36-a) نرى مركزي الهدفين Target1 وTarget2 بعيدين بحيث ناتج عملية الترابط معدوم وفي الشكل (4-36-b) نرى الهدفين قريبين بحيث يتداخل تابعي الكثافة الاحتمالية لهما بالتالي تنتج قيمة للترابط هي Max c والتي تكون قيمة كبيرة كما في الشكل (4-36-c) بسبب شبه الانطباق الحاصل. لكشف الكائن المرتبط بالكائن المدروس. يتم جمع قيم التشابه لجميع الأجزاء لكل كائن بعد التداخل، وسيكون الكائن المشابه هو صاحب أعلى قيمة تشابه لمجموع الأجزاء.



الشكل (4-36) احتمالية وجود الهدفين Target1 وTarget2 (الباحث)

من أجل وضع الجواب السابق كاحتمال يمكن هنا تعريف متحول جديد يعبر عن احتمال كون هذا الكائن هو الكائن المطلوب P أو لا، وهو يعطى بالعلاقة (4-13)، (الباحث):

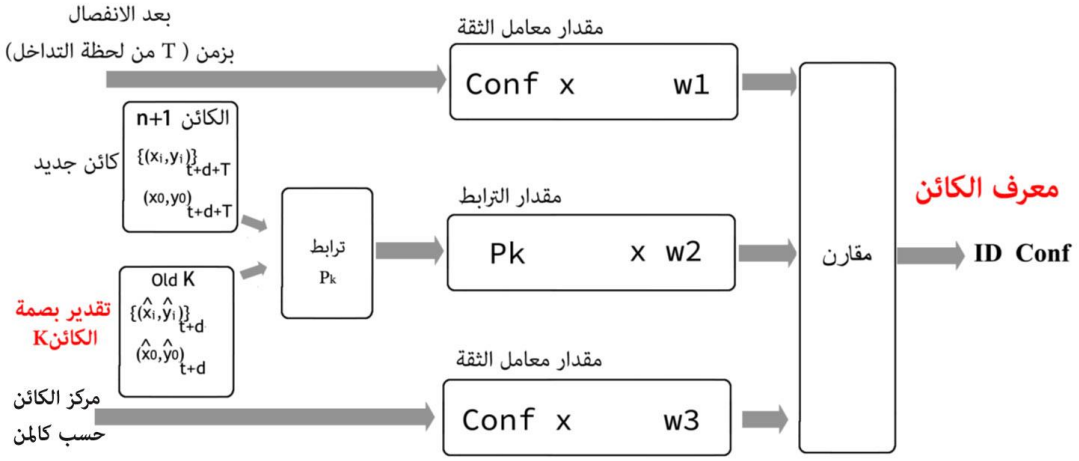
$$P_k = \frac{Max(\sum Sim)}{n} \qquad (13\text{-}4)$$

حيث n هي عدد نقاط المستنتجة من تجزئة الكائن، وk هو رقم معرف الكائن.

4-3-4-7 ربط المسار بين مرحلتي ما قبل التداخل وما بعده

يبين الشكل (4-37) عمل الخوارزمية لاعتماد النتائج بعد التداخل. لربط المسار واختيار الكائن الصحيح بعد التداخل، وهو الكائن المراد ربطه مع الكائن K قبل التداخل، وهنا نقارن بين ثلاث نتائج:

- الناتج الأول وهو الحاصل من خوارزمية الملاحقة الأساسية.
- الناتج الثاني وهو ناتج البصمة الحرارية المتوقعة بعملية التجزئة في طريقتنا.
- الناتج الثالث وهو الحاصل الاساسي من خوارزمية الملاحقة الأساسية مترافق مع احتمال مرشح كالمن لمركز احداثياته.



الشكل (4-37) عمل الخوارزمية لاعتماد النتائج بعد التداخل (الباحث)

تؤخذ النتائج بالعلاقة بشكل مصفوفي كما يلي:

$$Result = [Result_i \quad Result_{i+1} \quad Result_{i+2} \quad . \quad Result_{i+n}]$$

حيث: i هو أصغر رقم لمعرف موجود قبل التداخل.

n هو عدد الأهداف الأعظمي قبل التداخل أو بعد التداخل وتصبح عناصر المصفوفة السابقة وفق العلاقة (14-4)، الباحث:

$$Result = [w1 \quad w2 \quad w3].\begin{bmatrix} conf_i & conf_{i+1} & conf_{i+2} & \dots\dots & conf_{i+n} \\ P_i & P_{i+1} & P_{i+2} & \dots\dots & P_{i+2} \\ Cconf_i & Cconf_{i+1} & Cconf_{i+2} & \dots & Cconf_{i+n} \end{bmatrix} \quad (14\text{-}4)$$

يعامل كل ناتج بعوامل تثقيل w1, w2, w3 للحصول على النتيجة $Result_k$، وقد تم الوصول لتحديد كل معامل بتجارب الاختبار.

يمثل Conf معامل الثقة الناتج سواء من طريقة الملاحقة أو من الملاحقة بنموذج حركة. أما P_k فهو معامل احتمالي يمثل الترابط بالطريقة المقترحة. ناتج عملية المقارنة سيكون كما يلي:

- 0: أي لا يوجد ارتباط مع أي كائن حالي بالتالي ويحدد ذلك عندما $Result_k \leqslant 0.4$ وهو معرف الكائن الذي سيرتبط بالكائن K بعد التداخل. وذلك سيعني أن الكائن k اختفى.
- k: أي الكائن المدروس بعد التداخل هو ذاته الكائن k قبل التداخل بالتالي يأخذ ذات المعرف k. ويتم ذلك بشرط $Result_k > 0.4$
- New: رقم جديد عندما $Result_{all} \leqslant 0.4$ لا يوجد ترابط بين الكائن الجديد وكل الكائنات السابقة.

خلاصة الفصل الرابع:

في هذا الفصل تم عرض للمشاكل التي تعترض عملية الملاحقة وخصوصاً في الصورة الحرارية، وتم تحديد المشاكل الأكثر تأثراً في موضوع التداخل من ضعف سمات الشكل، تغير شكل الكائن، وأخطاء اعتماد نماذج الحركة. بعد ذلك تم التطرق لطرائق الحل. فتم اعتماد تحسين الصورة الحرارية والتعديل بمعاملات الكشف قبل تطبيق الكاشف YOLO مما يحسن من سمات الكشف، ومن ثم يتم تجزئة الكائن المكتشف لتشكيل البصمة الحرارية المؤلفة من مراكز ثقل الأهداف التي نسبت لمركز ثقل الكائن، بالاضافة لسرعات هذه الأجزاء. أما بعد التداخل فيتم اختبار مدى مطابقة كل كائن مع الكائنات التي كانت موجودة قبل التداخل، وعندها يتم اعادة تعريف هذه الكائنات وانشاء الجديد وحذف المختفي.

الفصل الخامس

التنفيذ والنتائج العملية

Implementation and Experimental Results

1-5 أدوات البحث

1-1-5 العتاد المستخدم

من أجل تجارب التحليل وتطوير الخوارزمية تم استخدام حاسوب من النوع Asus بالمواصفات التالية:

- وحدة معالجة مركزية Intel core i7, 2.6 GHz
- ذاكرة نفاذ عشوائي 8 GB

من أجل اختبار الخوارزمية المحسنة استخدمت منصة ColabResearch للتغلب على سرعة المعالجة بالحواسب، وهي تؤمن معالجة ضمن المسرع T4GPU وهو كذلك يستعمل Google Drive لتخزين النموذج والبرمجيات بالتالي الاستقلال عن مشاكل استخدام الحواسب الشخصية.

لتحصيل المشاهد الحرارية وتكوين مجموعات الاختبار الخاصة تم استخدام كاميرا حرارية غير مبردة بالمواصفات التالية:

- الابعاد بالبيكسل 384 x 288 P
- النظام NTSC
- طول الموجة 8 – 14 µm
- البعد المحرقي 75mm
- زاوية الرؤيا 14 درجة

2-1-5 البرمجيات

نظام التشغيل Windows 10.

تم استخدام برمجيات Matlab النسخة R2023b من أجل عمليات التحليل والمحاكاة وبعض التجارب.

لاختبار الخوارزمية المقترحة، تم استخدام لغة Python النسخة 3.12.

تم استخدام برمجيات منصة Colab، أو "Colaboratory"، من أجل اجراء تجارب جزئية على الخوارزمية واختبار الأداء العام. ميزة هذه البرمجيات أنها تتيح كتابة وتنفيذ Python في المتصفح، مع عدم الحاجة إلى أي تعديل، وكذلك تمكن من الوصول إلى وحدات معالجة الرسوميات مجانًا.

3-1-5 المكتبات

مكتبة google.colab/drive: التي تمكن من الوصول الى ملفاتنا المخزنة لدى google من أي مكان في العالم، دون الحاجة لتثبيت أي برمجية.https://colab.research.google.com /

مكتبة os: تؤمن الوظائف المعتمدة على نظام التشغيل مثل قراءة وكتابة الملفات والتعامل مع مساراتها ضمن الحاسوب.https://docs.python.org/3/library/os.html

مكتبة cv2 : لتنفيذ المهمات الخاصة بمعالجة الصور وتحليل مقاطع الفيديو، وتعلم الآلة، وأوامرها وتوابعها شبيهة بأوامر وتوابع الماتلاب MATLAB. https://pypi.org/project/opencv-python/

مكتبة matplotlib: هي مكتبة شاملة تضم حزمة رسوميات ثنائية الأبعاد تستخدم في Python لتطوير التطبيقات والبرمجة النصية التفاعلية وتوليد صور بجودة النشر عبر واجهات المستخدم وأنظمة التشغيل (Hunter, 2007). (https://pypi.org/project/matplotlib)

مكتبة PIL: تضيف مكتبة الصور Python Imaging Library قدرات معالجة الصور إلى مترجم Python

وتوفر هذه المكتبة دعماً واسع النطاق لتنسيقات الملفات، وتمثيلًا داخليًا فعالاً، وقدرات معالجة صور قوية إلى حد ما ويعتبر تصميم مكتبة الصور الأساسية للوصول السريع إلى البيانات المخزنة في عدد قليل من تنسيقات البيكسل الأساسية. يجب أن توفر أساسًا متينًا لأداة معالجة الصور العامة. (https://pypi.org/project/pillow).

مكتبة numpy: من المكتبات المهمة التي تعتمد عليها بشكل قوي العديد من مكتبات البيانات وتعلم الآلة، كذلك تتيح هذه المكتبة قدرة التعامل مع المصفوفات بطريقة أفضل من ال Lists الموجودة تلقائيًا كشكل من أشكال تراكيب البيانات في البايثون وتُعتبر أكثر كفاءةً ومتانة.(https://numpy.org/)

مكتبة skimage: عبارة عن مجموعة من الخوارزميات لمعالجة الصور. وهي متاحة مجانًا وبدون أي قيود. تتميز ببرمجيات عالية الجودة. (https://scikit-image.org/)

مكتبة Ikomia: هي أداة تطوير مفتوحة المصدر مصممة لتبسيط تطوير ونشر حلول الرؤية الحاسوبية. فهي تمكن من دمج الخوارزميات الحديثة من مصادر متعددة، مثل OpenCV وDetectron2 وOpenMMLab وYOLO، في التطبيقات.(https://ikomia-dev.github.io)

2-5 مجموعات البيانات

بالنسبة لمجموعات البيانات المستخدمة في التجارب، فقد تم استخدام بيانات Flir_ADAS_V2، وهذه البيانات مكونة من 9233 صورة مرئية و9711 صورة حرارية مقسمة للتدريب والتحقق والاختبار. ونظراً لاستخدامنا نموذج مدربا مسبقا فقد استعملنا مجموعة الاختبار المكونة من 3749 صورة مزدوجة مرئية وحرارية.

تشكل هذه الصور لقطات فيديوية تم أخذها عبر زوج من الكاميرات الحرارية والمرئية باستخدام Teledyne FLIR Tau 2 مقاس 13 مم f/1.0 مع مجال رؤية أفقي 45 درجة (HFOV) ومجال رؤية رأسي 37 درجة (VFOV). كانت الكاميرا الحرارية تعمل في الوضع الخطي T. تم التقاط الصور المرئية باستخدام كاميرا Teledyne FLIR BlackFly S BFS-U3-51S5C (IMX250) مع عدسة HFOV Edmund Optics بزاوية 52.8 درجة.

نظراً لفقر مجموعات البيانات المعيارية السابقة بالحالات المعالجة فقد استخدمنا مجموعة لقطات حرارية تم تحصيلها خلال فترة الدراسة، وتم انتقاء هذه اللقطات بحيث تخدم اختبارات النتائج في الطريقة المقترحة، ولقد تم تحصيل اللقطات بكاميرا حرارية بالمواصفات التي ذكرت سابقا.

في المجموعة Adataset تمت التجارب وفق الشروط:

- المشهد ليلي
- المسافة 35 متر
- عدد الإطارات 1006
- عدد اللقطات 3
- زمن الإطار 40m.s
- الميل الافقي لمحور الكاميرا 5- درجات.

في المجموعة Bdataset تمت التجارب وفق الشروط:

- المشهد نهاري
- المسافة 200 متر
- عدد الإطارات 2000
- زمن الإطار 40ms
- عدد اللقطات 10
- الميل الافقي لمحور الكاميرا 5- درجات.

3-5 مقاييس الأداء

تعتبر مقاييس الأداء من أهم ما يمكّن الخوارزمية من تحقيق أداء ذو كفاءة عالية، ولابد أن يكون تصميم الخوارزمية جيداً ومناسباً للغرض المطلوب، لكي تكون عملية تحليل الخوارزمية ناجحة، وعملية تقييم أداء الخوارزمية دقيقة، وللحصول على ذلك تم اتباع عدة طرائق:

- منذ البداية تم التركيز على المشكلة، حيث الكثير من الخوارزميات كانت قد ركزت على طرائق الملاحقة بشكل عام كما بالدراسات المرجعية، حيث تم التركيز بشكل دقيق على مشكلة تمييز الكائنات بالصورة الحرارية، وحددت المتطلبات والقيود المفروضة على الحل كجزء من مشكلة الملاحقة.
- تم ترتيب الخطوات بالترتيب المنطقي وتمت التجارب بشكل مجزأ لكل خطوة.
- تم اختيار الخوارزميات قيد الاستخدام بشكل مدروس لتخدم الناتج المنشود، حيث تم اختيار خوارزمية كشف حديثة، بالاضافة لخوارزمية ملاحقة ممتازة، حيث من خواصها اعتماد الكشف بطريقة عملها.
- تمت الاختبارات في كل مرحلة وتم تعديل على الخوارزمية المطورة بحيث يتم الحصول على كفاءة أفضل.

يتم قياس أداء الخوارزمية عادة بناءً على عدة معايير، مثل وقت التنفيذ، حجم الذاكرة، دقة الحل، معدل الخطأ، وغيرها من المعايير. يمكن تحديد معدل الخطأ بمعرفة عدد الحلول الخاطئة التي يتم إنتاجها بواسطة الخوارزمية.

في حالات ملاحقة الكائنات تستهلك كل الخوارزميات حجماً كبيراً من الذاكرة بسبب التكلفة البرمجية الكبيرة، لذلك سيهمل معيار حجم الذاكرة، وسينحصر الاستخدام بالمعايير الثلاثة الأخرى.

1-3-5 معيار دقة الحل

نناقش أداء الخوارزمية الجزئي خلال مرحلة الكشف باستخدام معيار معامل ثقة الكشف للكاشف Confidence ومتوسط الدقة (AP) باعتبار المراد من هذه المرحلة الكشف بأعلى جودة.

سيتم ضمن إجراءات الاختبار تنفيذ خوارزمية الكشف على مجموعتي البيانات Flir_ADAS_V2 المرئية والحرارية ومقارنة النتائج، ثم يعاد تنفيذ خوارزمية الكشف المحسنة على ذات البيانات الحرارية، وذلك لمعرفة مدى التحسن.

سيتم اجراء مقارنة بين أداء كشف مركز الكائن باستخدام الملاحقة لمركز الكائن ضمن بيانات منتقاة من المجموعات المتوفرة لحركة كائن يسير باتجاه مستقيم وبسرعة ثابتة بواسطة الكاشف YOLO V8 ثم باستخدام الطريقة المحسنة ومعيار المقارنة سيكون باستخدام المنحنيات البيانية لحركة مركز الكائن في الطريقتين.

نظرا لكون المقاييس المدروسة في الفقرة (4-4-2) بمجملها مقاييس شاملة في الكشف والملاحقة، وكون طريقتنا تعالج مشاكل جزئية من مشاكل الملاحقة، بالإضافة لذلك عالجت طريقتنا المشاكل بالصور الحرارية. بالتالي كان

صعبا الحصول على دراسات قريبة من بحثنا لاعتماد نتائج الدراسات كسبيل للمقارنة، لذلك سنستخدم طريقة تحليلية لبيان مقدرة الطريقة المقترحة في مجمل مساراتها مع مقدار التحسين في كل مرحلة ومقارنة ذلك مع بعض الخوارزميات الاساسية الجيدة الأداء، وذلك بذات مجموعات البيانات سواء المجموعات المعيارية أو مجموعات اللقطات الفيديوية التي تم التقاطها خصيصاً للعمل.

كما تم الذكر في السابق في الفقرة (4-4-2) تعتبر الخوارزميات Kalman وDeepSORT وYOLO V8 من بين أفضل عشر خوارزميات مستخدمة للملاحقة (Encord. 2025) وفي الطريقة المطورة تم استخدام مزيج من هذه الخوارزميات، وقد استخدم كالمن من أجل مراقبة وملاحقة نقاط المراكز، بالتالي سيكون التركيز بالاختبار على مجموعة السمات الإضافية التي اضافتها الطريقة لكل منهما، وذلك باستخدام بعض المعايير والمقاييس السابقة.

5-3-2 معيار معدل الخطأ

هو المعيار الأهم في اختبار طريقتنا حيث تهدف الخوارزمية المقترحة للتقليل من الأخطاء المرتكبة خلال عملية التداخل/ وهي حالة خاصة من الملاحقة، وهنا المقارنة ستتم بشكل احصائي لعدد مرات الخطأ في الطريقة والمقارنة مع طريقة الملاحقة بخوارزمية DeepSORT وYOLO V8.

إن اللجوء للعملية الإحصائية وعدم اعتماد معايير قياسية هنا ناتج عن كون المشكلة المعالجة هي مشكلة خاصة من مشاكل الملاحقة، ولا يراد لأي مشكلة أخرى أن تكون مؤثرة. لذلك تم اختيار مجموعة اختبار جزئية من مجموعات البيانات العامة، بحيث تحوي العديد من حوادث التداخل الواضحة.

5-3-3 معيار زمن التنفيذ

لقد تم عنونة الأطروحة بـ " تطوير خوارزمية تمييز الكائنات المتداخلة في الزمن الحقيقي باعتماد السمات الطيفية الحرارية والحركية"، إذاً المجال الزمني الذي نعمل عليه هو الموافق للزمن الحقيقي، بالتالي يمكن تحديد معيار زمني لتنفيذ الخوارزمية ضمن حد زمني.

إن معدلات الإطارات الأعلى تعد بتتبع أفضل للحركة السريعة، ولكن أنظمة الرؤية المتقدمة في الوقت الحقيقي نادرًا ما تتجاوز النطاق القياسي 10-60 هرتز، حيث الحساب المطلوب سيكون كبيرًا جدًا. لذلك يتم تخفيف معدل الإطارات من خلال انخفاض التكلفة الحسابية لكل إطار في أجهزة الملاحقة التي تستفيد من التنبؤ (Handa,2021). ان حدود زمن استجابة العين للأحداث لا تتعدى الـ 100 ms ومعظم كاميرات تحصيل الصور الحالية تعمل على تردد تحصيل الإطارات أكبر من 10 Frame/s.

في الطريقة المقترحة يتم ملاحقة هدف بالحالة الطبيعية لذلك يمكن اعتبار أن المعيار الزمني للمعالجة ضمن الخوارزمية يجب أن لا يتعدى الـ 100 ms الموافق لتردد تحصيل العينات 10 fps.

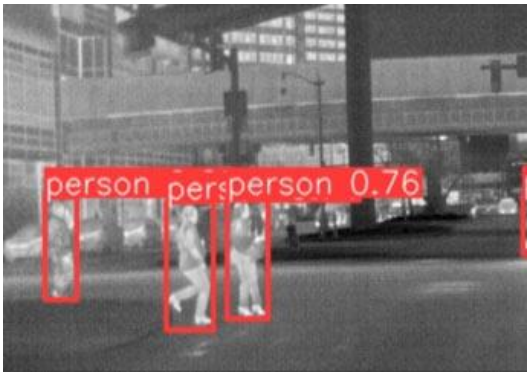
4-5 إجراءات التنفيذ العملي

1-4-5 موافقة الصورة الحرارية

في البداية تم تطبيق الكاشف YOLO-V8 على صورة تم انتقائها من ضمن صور الاختبار المعيارية، وهي صورة مرئية من مجموعة العينات ضعيفة الإضاءة كما يبين الشكل (5-1)، وتم تنفيذ النموذج المدرب YOLO-V8 عليها، يلاحظ أن النموذج قد استطاع كشف ثلاثة أشخاص بمعامل ثقة كبير وصل بأحد الأشخاص إلى 88%.

الشكل (5-1) نتائج YOLO-V8 على صورة مرئية

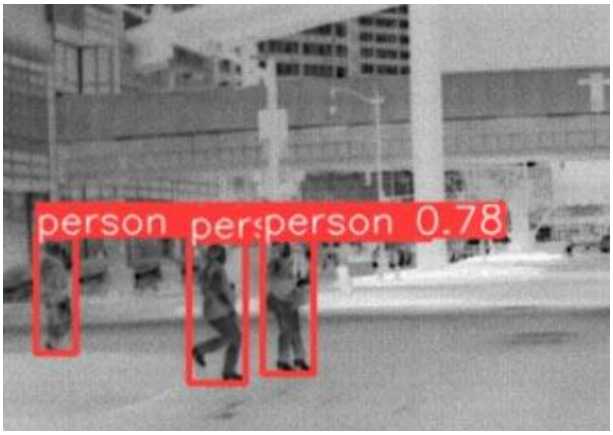
يمثل الشكل (5-2)، صورة حرارية من مجموعة البيانات الحرارية القرينة للمجموعة المرئية، بحيث تكون مقابلة للصورة في الشكل السابق (5-1)، وتم تنفيذ النموذج المدرب عليها، يلاحظ أن النموذج قد استطاع كشف ثلاثة أشخاص بمعامل ثقة أخفض ووصل للشخص السابق ذاته إلى 76%.



الشكل (5-2) نتائج YOLO-V8 على الصورة المماثلة الحرارية (الباحث)

يعزى انخفاض جودة الكشف لانخفاض سمات الهدف بالصورة الحرارية عن تلك الموجودة بالمرئية.

الآن يتم تطبيق التحسين على الصورة المتمثل في قلب الصورة لموافقتها مع الصورة المرئية. تعاد التجربة كما في الشكل (5-3) الذي يرينا ناتج النموذج المدرب على الصورة الحرارية وفق الطريقة المقترحة ويلاحظ التحسين في ثقة الكشف بمقدار 2%.



الشكل (5-3) نتائج YOLO-V8 وفق الطريقة المحسنة (الباحث)

عند إعادة التجربة بشكل متكرر على صور مجموعة البيانات الأخرى تمت ملاحظة ما يلي:

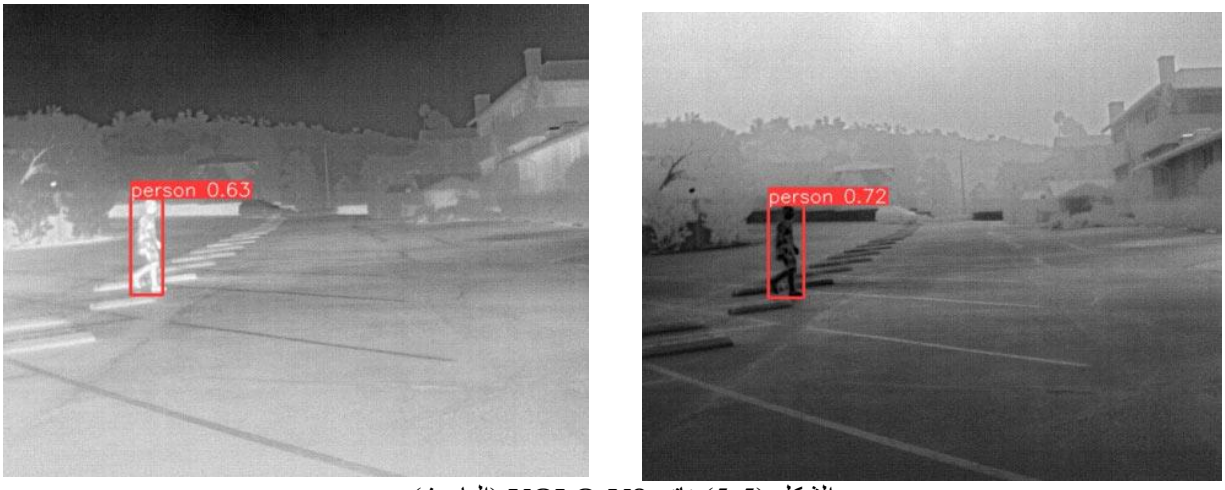
- فشل الكشف للكائنات الصغيرة
- انخفاض في قيمة الثقة لدى زيادة عدد الكائنات أو تعقيد الخلفية

سبب المشكلة الأولى هو تشوه الكائن عند صغر حجمه وهي مشكلة مشتركة بين المرئي والحراري ولكن أثرها أكبر بالحراري، أما عن طريق تخفيف هذا الأثر فستعالج بتخفيض معامل الثقة المقبول لدى الكاشف بالتالي الحصول على كشف أفضل للكائنات الصغيرة، كما تم الذكر في الفقرة (4-3-1). يعود سبب المشكلة الثانية إلى تعقيد الحسابات عند المشهد المعقد، وسيكون الحل هو تبسيط عدد المعالجات. في طريقتنا تم تطبيق الكاشف فقط في الأماكن المتوقعة لوجود الكائنات. بالطبع سيساهم نموذج حركة الكائن بذلك خلال خوارزمية الملاحقة، ولكن أضفنا لذلك اجراء إضافي وهو البحث عن الكائنات بالمناطق الأشد نصوعاً، باعتبارها الأماكن الأكثر احتمالاً لوجود الكائن بسبب تصديره للحرارة.

يرينا الشكل (5-4) صورة في ظل انارة ضعيفة جدا (صورة ليلية) ويُلاحظ عدم تمكن النموذج من الكشف. وهي احدى سيئات الصورة المرئية في ظروف الطقس السيئة والظلام.

الشكل (5-4) صورة ليلية لشخص (Flir_ADAS_V2)

الصورة اليسرى في الشكل (5-5) تمثل ناتج الكشف ضمن الصورة الحرارية لنفس المشهد بالنموذج YOULO-V8 وفي الصورة اليمنى الناتج المحسن وفق طريقة تعديل إضاءة الصورة.

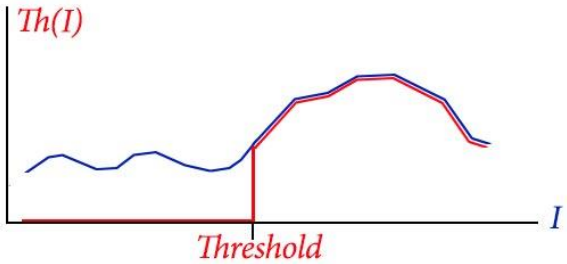


الشكل (5-5) ناتج YOLO-V8 (الباحث)

وكما ذُكر سابقاً، لتحسين أفضل لظروف الكشف يمكن تجزئة الصورة المطلوب كشف الكائنات البشرية فيها بحيث يتم الكشف فقط بالأماكن المحتملة وهي الأماكن المحددة في مناطق الشدة العالية. ولتنفيذ ذلك يتم تجزئة الصورة على مرحلتين: الأولى باستخدام تابع العتبة أي يسمى بعتبة تدريج الرمادي المستخدمة في التجزئة "تابع العتبة" كما رأينا في العلاقة (4-7) لكن بتعديل بسيط، حيث نعتمد العلاقة (5-1).

إن I هي شدة الإضاءة أو النصوع وThreshold عتبة التابع والتي ستستنتج قيمة مجالها المسموح لدى اجراء الاختبار على مجموعة البيانات، وسنعطي هنا الشدة $I_{i,j}$ القيمة 0 للبكسلات التي شدتها أقل من العتبة والقيم ذاتها للقيم فوق العتبة كما في الشكل (5-6) الذي يرينا مخطط الهيستوغرام وناتج التجزئة، مع التأكيد بأن Threshold هي التي أصبحت بعد إجراءات طريقتنا السابقة ممثلة لعتبة التجزئة المتكيفة ضمن المجال المسموح.

$$Th(I) = \begin{cases} 0 & if\ I < Threshold \\ I & if\ I > Threshold \end{cases} \qquad (1\text{-}5)$$



الشكل (5-6) تابع العتبة المستخدم (الباحث)

يري الشكل (5-7) صورة حرارية لمشهد في الشارع وفيه عدد من السيارات ومجموعة من الأشخاص

الشكل (5-7) مشهد حراري (Flir_ADAS_V2)

في الشكل (5-8) باستخدام تابع العتبة المتكيف يتم تجزئة المشهد.

الشكل (5-8) تجزئة المشهد (الباحث)

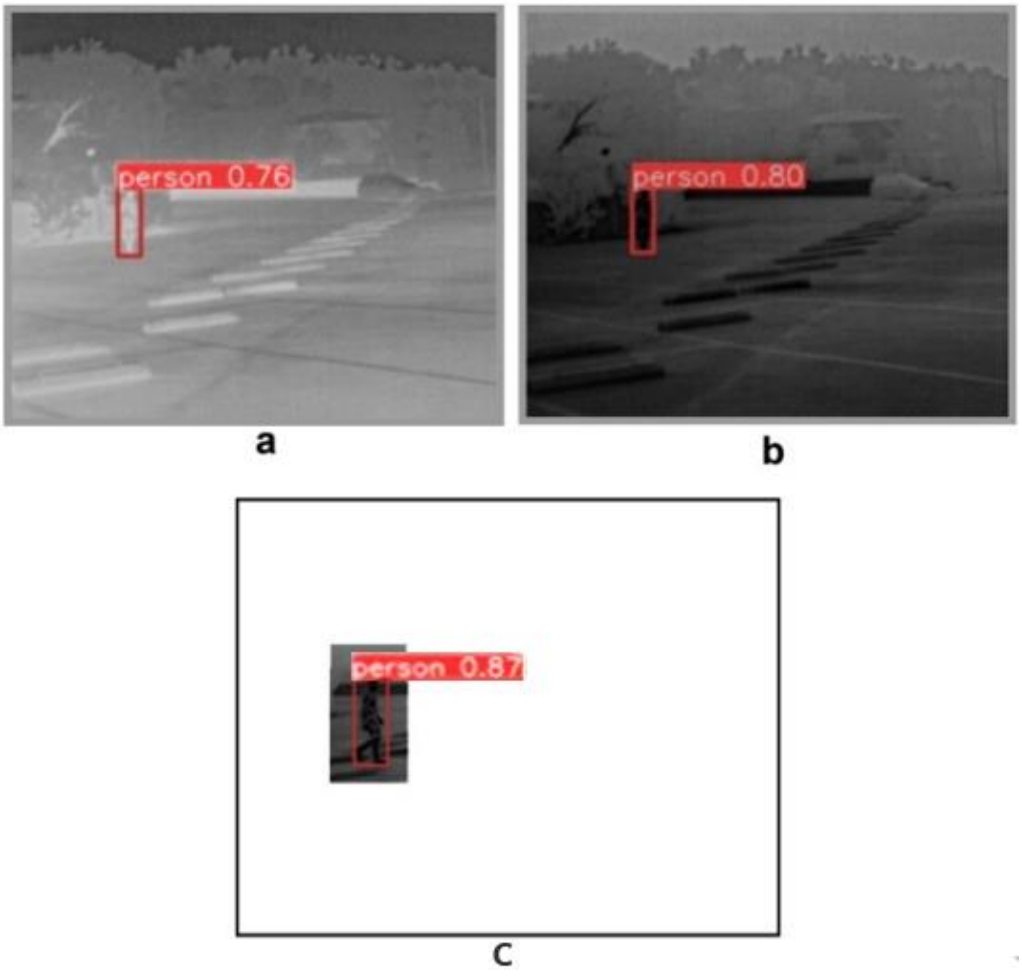
بعد أخذ الصورة المجزأة يمكن البحث بالصورة الأصلية ضمن منطقة شدة كبيرة (مستطيل أبعاده مساوية لـ K من أبعاد الهدف المجزأ)، بالعمليات المورفولوجية يمكن تحديد الهدف المحتمل بانه هو الذي تعداد نقاطه أكبر من عتبة معينة Nu كما في الشكل (5-9).

الشكل (5-9) تحديد مستطيل مناطق الشدة المرتفعة (الباحث)

وفي الشكل (5-10) يتم تطبيق تابع العتبة المذكور.

الشكل (5-10) البحث ضمن مستطيل مناطق الشدة المرتفعة (الباحث)

بعد ذلك سيتم تطبيق النموذج المدرب مسبقا، مع الأخذ بعين الاعتبار تعديل معاملاته السابقة، للحصول على النتائج المرجوة. لمعرفة تأثير مختلف التدابير المتخذة لتحسين الكشف تم اجراء تجارب عديدة على مجموعات البيانات المذكورة ابتداءً من اللقطات التي يظهر فيها هدف واحد إلى متعددة الأهداف.



الشكل (5-11) ناتج YOLO-V8 بعد اجراء التجزئة بتابع العتبة (الباحث)

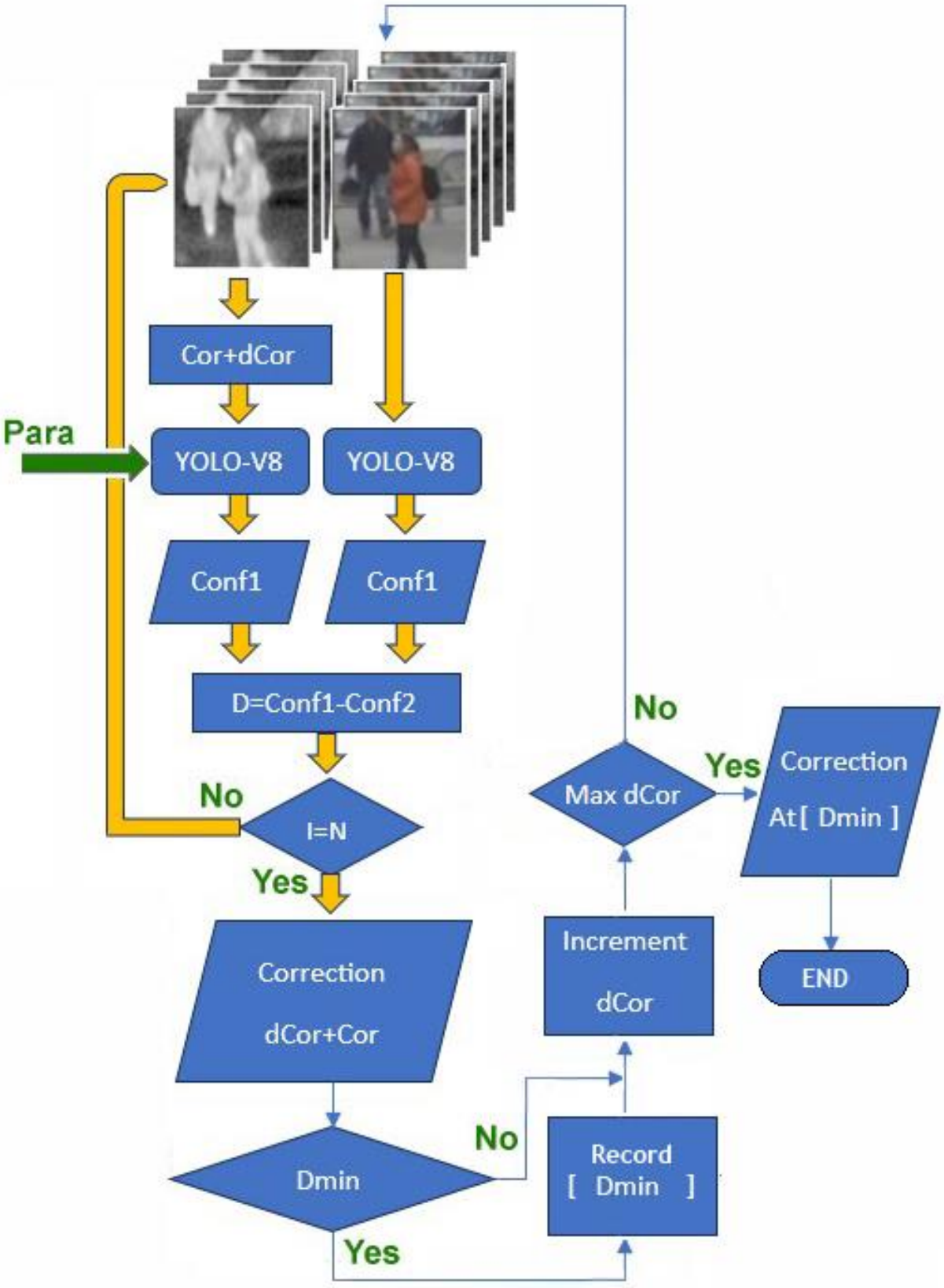
يظهر الشكل (5-11) ناتج الكاشف في حالات مختلفة، عند (a) تطبيق الكاشف المدرب فقط، أما في (b) عكس الصورة وتحسينها، وأخيرا في (c) إضافة تجزئة للصورة بحسب الحرارة، وتطبيق اجراءات الكشف المحسنة كاملة. بالمقارنة مع الشكل (5-5) يلاحظ التحسن الكبير على جودة الكشف ولكن الحالة ليست عامة حيث يمكن ألا تتحسن الحالة أبدا وذلك حسب الصور، ولكن الفائدة الأكبر من حالة التجزئة هي سرعة المعالجة. والتعويض عن التأخيرات بسبب إجراءات التحسين.

وقد تم التأكد من ذلك على صور متعددة من مجموعة البيانات حيث تم أخذ 300 صورة من إطار فيديوي ليلي يظهر فيه شخص يقطع الشارع. وتمت عليه عدة اختبارات.

للحصول على أفضل تعديل مناسب للخوارزمية المطورة للكشف ضمن الصورة، تم اختبار الطريقة على مجموعتين من الصور المرئية والحرارية الخاصة بالأشخاص. الشكل (5-12) يبين مخطط تدفقي لإجراءات خوارزمية التعديل على معاملات التحسين المختلفة، وعند الانتهاء من معالجة الصور يخزن الانزياح في دقة الكشف الناتج D، ثم يتم التكرار عدة مرات حتى الوصول إلى أفضل معامل تصحيح، بحيث في كل مرة يتم تغير قيم التعديلات dCor على معاملات التصحيح Cor.

إن معامل التصحيح هنا عبارة عن مصفوفة من المعاملات تحوي قيمة صغيرة من نسبة عتبة الكشف K ونسبة الانتقال اللوني اللازم لضغط المستوى d والاضاءة والتباين، بالإضافة لعتبة تابع التجزئة وقد اختيرت بمقدار $\mp 20\%$ من متوسط الإضاءة بحيث يتم الضبط المتوسط الأمثل لمجمل الصور بحيث نحصل على مصفوفة أفضل النتائج، وهي المقابلة لفرق معامل ثقة أصغري Dmin عند أكبر عدد من المعاملات المتغيرة حيث تختار من مصفوفة أقضل الفروقات المسجلة [Dmin].

يمثل Para في الشكل نسبة عتبة التجزئة في IOU الخاصة بـ YOLO-V8 حيث يمكن التضحية بنسبة معقولة تحت 50% باعتبار أن الكائن الملاحق هو كائن حراري. وقد تم كذلك التجريب وإعادة الإجراءات السابقة بالشكل (5-12) وفي كل مرة يتم تغير النسبة حتى تم الوصول إلى 30%.



الشكل (5-12) المخطط المسلكي لخوارزمية تعديل التصحيحات (الباحث)

بالنسبة للإضاءة والتباين وعتبة التجزئة فتظهر التجارب أنه في الحالة العامة لا تأثير يذكر ولكن يزداد التأثير في حالة تداخل الهدف مع خلفية حرارية. في الشكل (5-13) يلاحظ وجود خلفية متداخلة مع الهدف تم تعديل معاملات الاضاءة والتباين حسب ما يلي:

Alpha=1.3, beta=30

وقد أظهرت النتائج تحسناً بمعامل الثقة من 0.83 إلى 0.84، أي 0.8%

الشكل (5-13) التحسين بتعديل الإضاءة والتباين (الباحث)

أما مجال عتبة التجزئة فكان:

Threshold= [50,139]

يبين الجدول (5-1) ثوابت التحسين على المعاملات المختلفة للصورة والنموذج Para+CorK، وقد ثبتت هذه الثوابت بعد التجارب المختلفة، ويبين الجدول (5-2) نتائج التحسين بتطبيق طريقة الكشف خاصتنا Proposed بالمقارنة مع نتائج YOLO-V8 لمجموعات مختلفة من بيانات التحقق السابقة عند درجات اضاءة مختلفة للصور المرئية RGB والصور الحرارية Thermal:

الجدول (5-1) ثوابت التحسين (الباحث)

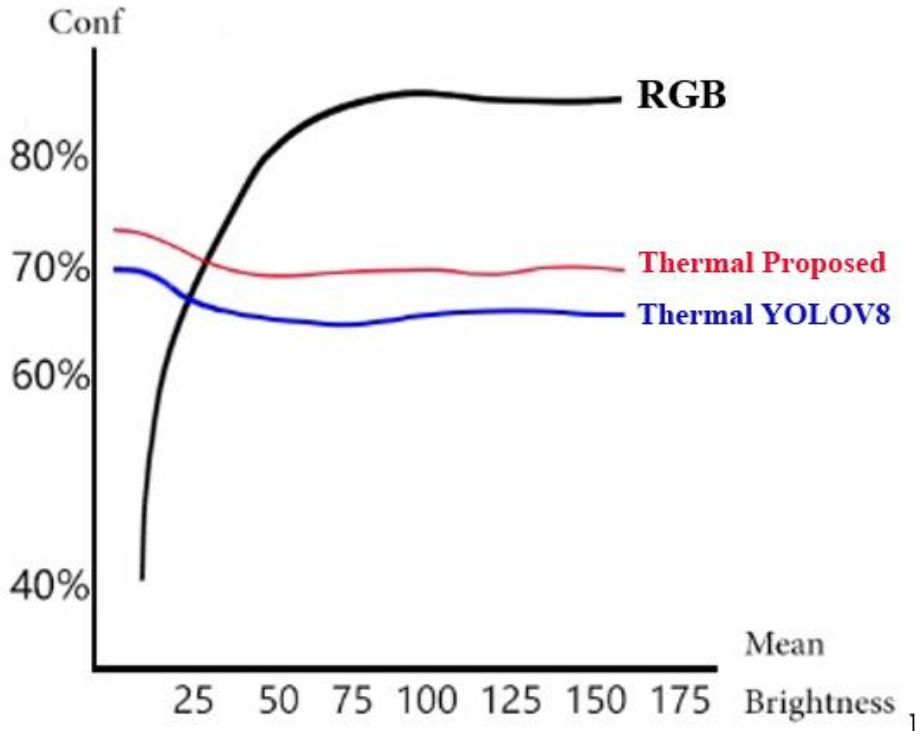
d	Nu	K	Brightness beta	Contrast alpha	IOU
d+0.1	100	3	30	1.3	0.35

الجدول (5-2) نتائج اختبار الطريقة المطورة في مجال الكشف (الباحث)

Condition	RGB	Thermal	Thermal
Mean Brightness	YOLOV8	YOLOV8	Proposed
25	-	67.6	73.3
50	77.7	66.4	69.9
75	85.2	64.7	71.3
100	88.1	65.7	71.8
150	86.2	66.3	72

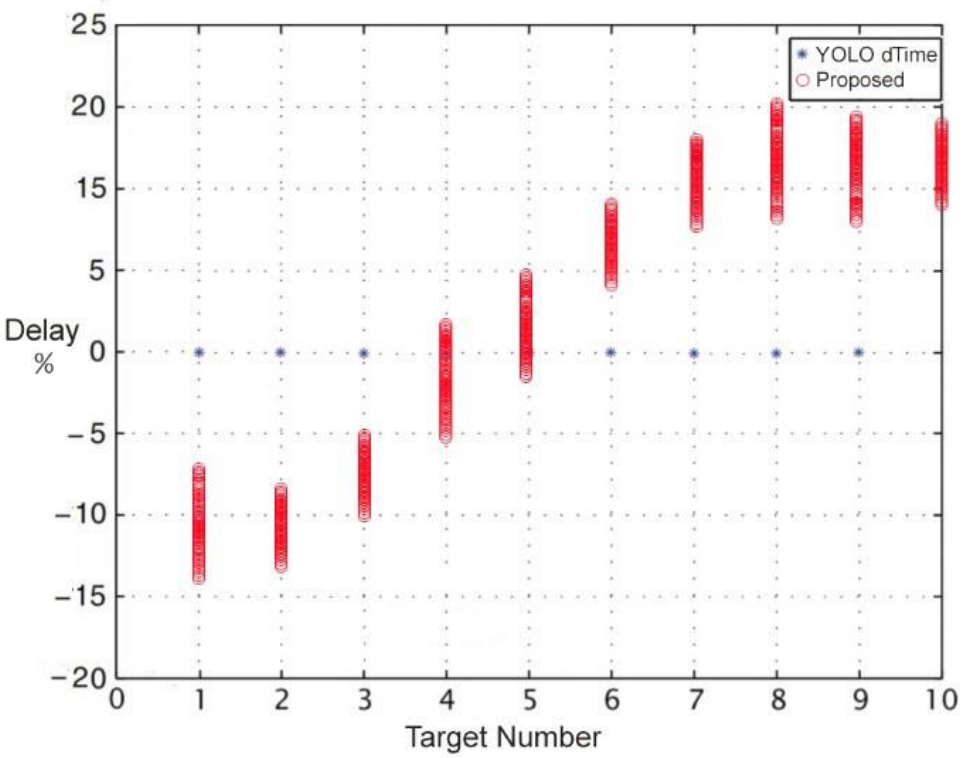
يلاحظ بأن التحسين على أداء الكاشف بين Proposed وYOLOV8 في الحالات المختلفة هو ضمن المجال [3,10] % ويلاحظ التحسن الكبير في ظروف الطقس السيئة (الظلام) بسبب عدم فعالية الكاشف المرئي.

من الملاحظ أثناء التجارب ارتباط وثوقية الكشف في النموذج بدرجة الاضاءة. حيث عند ظروف الإضاءة السيئة ستتقدم جودة الكشف في الكاميرا الحرارية عن نظيرتها المرئية. لتحديد درجات الإضاءة في الصورة تم استخدام متوسط الإضاءة واختبار مجموعة البيانات السابقة بالإضافة لمجموعة تدريب أخرى خاصة بالصور المرئية من مجموعات التدريب MOT17 ويبين المنحني التالي في الشكل (5-14) نتائج اختبارات الكشف في عدة مجموعات بإضاءة مختلفة لكل منها



الشكل (5-14) علاقة معامل الثقة بإضاءة الصورة (الباحث)

يبين الشكل نتائج اختبارات الكاشف على صور البيانات في مجموعات Mot17, Flir_ADAS_V2 حيث المنحني الأسود يمثل الناتج للصور المرئية لنموذج YOLO-V8 والمنحني الأزرق يمثل الناتج للصور الحرارية أما المنحني الأحمر فيمثل الناتج المحسن، ويلاحظ فعالية استخدام الكاميرا الحرارية وخصوصاً لدى انخفاض الاضاءة للمشهد. أما الشكل (5-15) فيبين نسبة أزمنة التأخير Delay للطريقة المقترحة على أزمنة تنفيذ UOLO-V8 المرئية والتي تطرأ على النموذج حسب الخوارزمية ويلاحظ بأن الخوارزمية ستختصر زمن التنفيذ في حالات أعداد الأهداف المكتشفة Target Number ضمن المشهد صغيرة ويظهر ذلك بالتأخير السالب، وعند الأعداد الكبيرة للأهداف ستكون هناك نسبة تأخير بسيطة من مرتبة % [7,20]



الشكل (5-15) التأخيرات الزمنية (الباحث)

5-4-2 الملاحقة بالخوارزمية المطورة الأساسية

في المرحلة الأولى يتم تنفيذ خوارزمية YOLO-DeepSORT بالإضافة لخوارزمية الملاحقة المطورة وبالاعتماد على المعاملات التي توصلنا لها في الفقرة (5-4-1) على الصورة الحرارية على بيانات ADAS-FLIR وتقارن مع نتائج الدراسة المرجعية (Ibrahim,2023) كونها تعتمد على YOLO-DeepSORT.

في المرحلة الثانية يتم تنفيذ خوارزمية الملاحقة المطورة وباعتماد ذات المعاملات على البيانات Our التي صممت لهذا الغرض. يبين الجدول (5-3) نتائج الاختبارات بمعايير الدقة MOTA ومعيار مؤشر التعريف IDF1

الجدول (5-3) اختبارات أولية لخوارزمية الملاحقة المطورة (الباحث)

Manar	IDF1	MOTA	IDF1	MOTA
Ibrahim,2023	54.6	63.8	-	
Data Set	ADAS-FLIR		Our	
YOLO-DeepSORT	54.1	63.6	50.3	59.2
Proposed	53.9	63.7	51.4	58.9

يلاحظ من النتائج في الجداول السابقة بأن نتائج استخدام خوارزمية الدراسة المرجعية السابقة قريبة نوعا ما مع نتائج خوارزمية ADAS-FLIR وكذلك قريبة من نتائج الخوارزمية المقترحة بالرغم من التحسين الواضح في مجال الكشف كما بيّنا في النتائج بالفقرة (5-4-1).

بالنظر إلى النتائج السابقة نلاحظ عند تطبيق خوارزمية المقترحة على مجموعات الاختبار، يتبين أن النتيجة النهائية للخوارزمية في أفضل الحالات ستكون مطابقة لخوارزمية الملاحقة أو الملاحقة مع نموذج حركة، ولدى مراجعة أساسيات الخوارزمية باستخدام التحليل برمجياً تم التوصل إلى مشكلات تتعلق بـ: مركز ثقل الكائن، عدم صلادة "تغير شكل" الكائن الحراري، ومشكلة نموذج الحركة. وقد تم حل هذه المشاكل، كما سنوضح بالفقرة التالية.

5-4 تحسين الخوارزمية المطورة

5-4-1 مشكلة مركز ثقل الكائن وحلها

لقد أدى اعتماد الطريقة المطورة على مركز الكائن إلى عدم ثباتية احداثيات بعض الأجزاء خلال الحركة بسبب ارتباط هذه الاحداثيات بالمستطيل المحيط وهذا المستطيل تتغير أبعاده حسب موقع أطراف الشخص "أطراف مفصلية" بالتالي انخفاض الدقة في تحديد مركز الهدف بالنسبة للزاوية اليسرى العلوية للمستطيل، مما يؤثر على احداثيات مراكز الأجزاء الداخلية. لحل هذه المشكلة تمت المعالجة بتحسين حساب مركز الثقل. تم اعتماد مركز قناع الهدف بديلا عن مركز مستطيل الإحاطة أو زاويته العلوية اليسرى. واستخدمت عزوم نقاط الهدف في حسابه (Wirth, 2004)، (Gahramanova, 2019) كما في العلاقة (5-2):

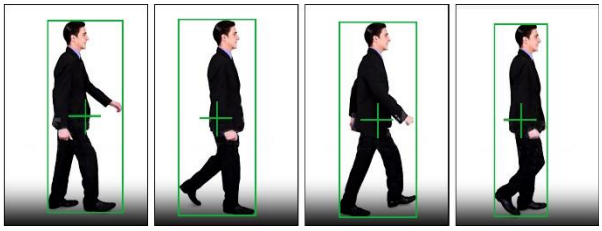
$$m_{pq} = \sum_{-\infty}^{+\infty} \sum_{-\infty}^{+\infty} i^p \; j^q f(i,j) \qquad (2\text{-}5)$$

حيث $(x,y), (i,j)$ هي إحداثيات البيكسيل بالمنطقة (إحداثيات البيكسل في الصور الرقمية). ويكون العزم بالنسبة للمركز كما بالعلاقة (5-3):

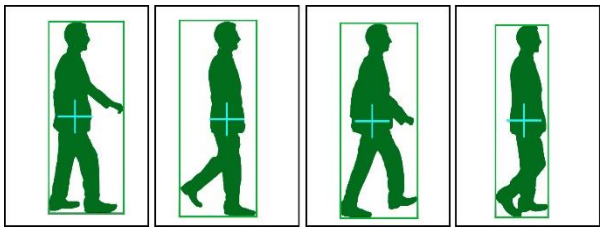
$$\mu_{pq} = \sum_{i=-\infty}^{i=+\infty} \sum_{i=-\infty}^{i=+\infty} (i - x_c)^p (j - y_c)^q \, f(i,j) \quad (3\text{-}5)$$

حيث x_c , y_c هي إحداثيات ثقل الكائن (مركز الكائن)

يبين الشكل (5-16) مركز الكائن بالاعتماد على الطرائق التقليدية (ناتج الكاشف)، بينما الشكل (5-17) يبين المركز بالاعتماد على التحسين المذكور.



الشكل (5-16) مركز الهدف بالاعتماد على ناتج خوارزمية الكشف



الشكل (5-17) مركز الهدف بالاعتماد على الطريقة المحسنة

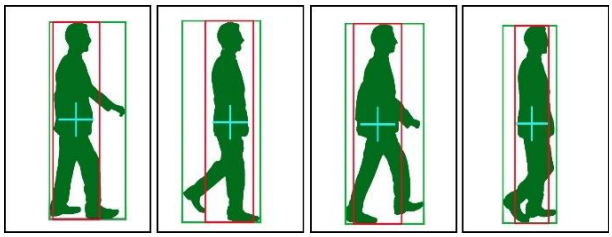
5-4-2 مشكلة عدم ثبات شكل الكائن الحراري أثناء الحركة وطريقة الحل

تأتي مشكلة صلادة "عدم ثبات شكل" الكائن من مصدرين:

- حركة الأطراف للكائنات، وهي مشكلة محلولة بالنسبة للكاشف، الذي درب على بيانات كثيرة لمختلف الحالات بما فيها وضعية الأطراف، ولكن في حالتنا ستؤثر هذه الحالات على مركز الكائن كما ذكرنا بالفقرة السابقة حيث حركة الأيدي ستزيح مركز الكائن.
- تغير شكل الكائن الحراري عند الحركة، وهي مشكلة ثانويةً هنا لأن تأثيرها صغير باستخدام تجزئة الكائن

للتقليل من التأثير للأطراف المفصلية لجأنا لحلين:

تم اعتماد الأجزاء الداخلية التي احداثياتها الأفقية قريبة من المركز ضمن مجال 60% من عرض قناع الكائن ومتمركزة مع الاحداثي الأفقي لمركز الهدف المحسوب كما هو مبين بالشكل (5-18)



الشكل (5-18) اهمال الأجزاء البعيدة عن المركز أفقياً

بالإضافة لذلك وبملاحظة أن النقاط العلوية للهدف أكثر ثباتاً من النقاط السفلية "الأطراف المتمفصلة" تم تثقيل الإزاحة الأفقية لمركز الكائن Wx_0 مع بشكل متناسب مع العزوم العلوية.

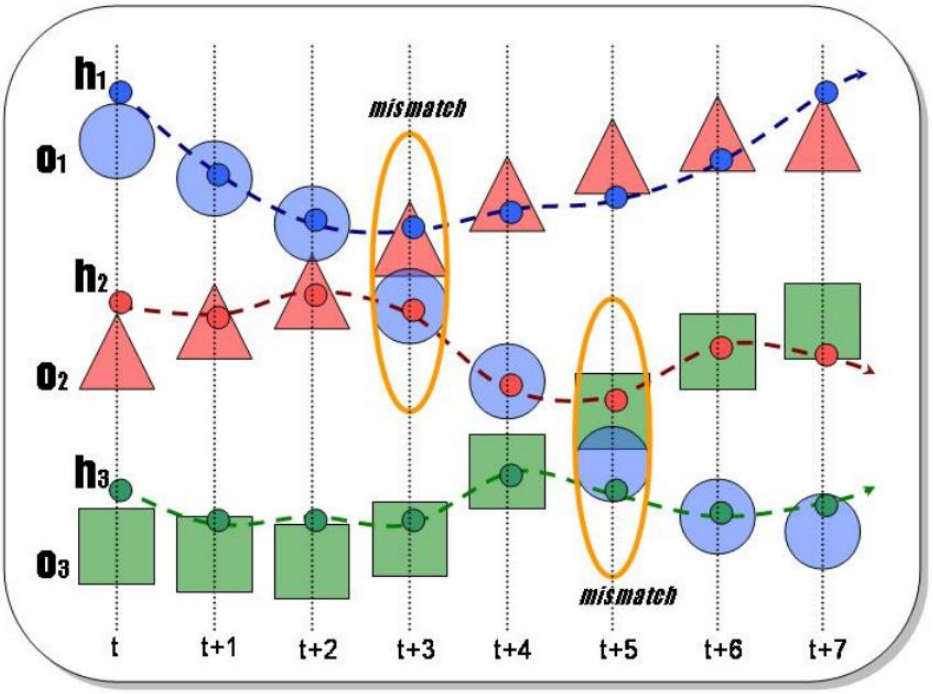
تبين العلاقة (4-5) (الباحث) التثقيل عند السطر y يكون:

$$Wx_0 = \frac{y_{max}-y}{y_{max}} \qquad (4\text{-}5)$$

3-4-5 مشكلات نموذج الحركة والاجراءات

كما يلاحظ من الفقرة (4-3-2) فإن نموذج الحركة سيُخدع عندما يغير المتحرك الاتجاه خلال فترة التداخل بالتالي سيتم تبادل الرقم التعريفي للكائنات. وقد تم استخدام نموذج حركة بسرعة ثابتة، ولم يُستخدم متنبئ للتسارع بسبب التعقيد الحسابي ومشكلة الزمن، حيث تغيير الاتجاه هنا هي مشكلة تسارع أي تغيير السرعة.

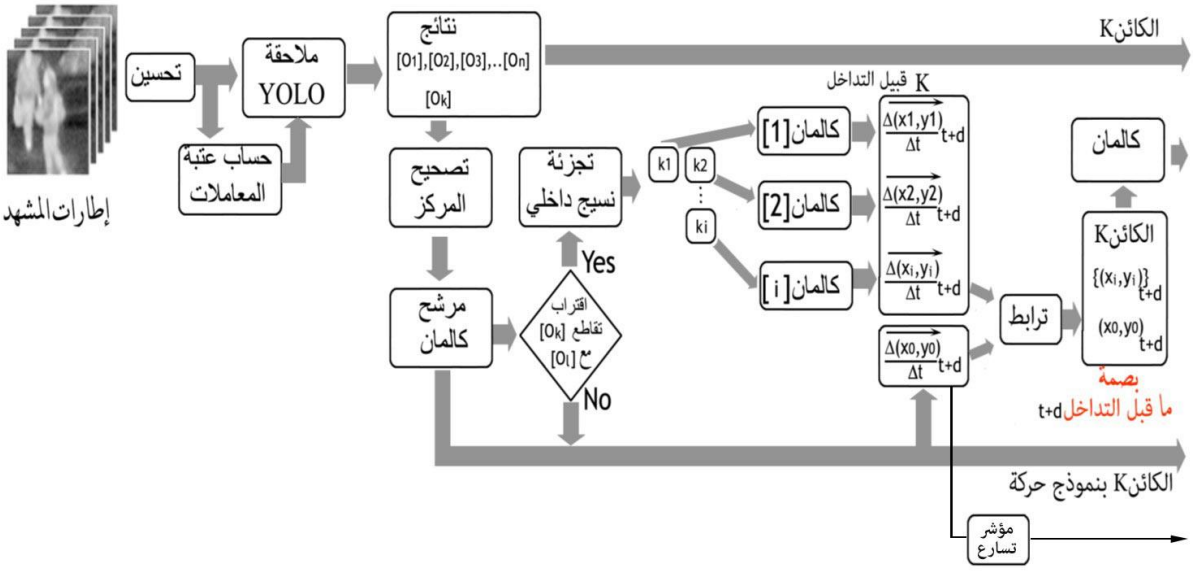
لحل هذه المشكلة سيلحق معامل التسارع بالخوارزمية وبالتحديد بمرشح كالمن، وذلك فقط لدى مراقبة تغير السرعة، وذلك يحد نوعاً ما من الخطأ ولكن على حساب الزمن، ويذكر أن استخدام هذا الحل سيكون مقتصداً ومشروطا. من أجل حل أكثر شمولية سنربط معامل التثقيل w2 المرتبط بشكل الكائن مع هذه الحالات، أي سيكون w2 متغيراً بحيث تكون قيمته عادية في الحالات الاعتيادية لحركة الأجسام ولكن يأخذ w2 قيمة أكبر لدى حالات المناورة أي تغير السرعة نقصاناً بأحد المحورين وتوقع انعدام السرعة خلال التداخل. بالتالي سيكون اختبار الشكل هنا له أهمية راجحة على اختبار نموذج الحركة.



الشكل (5-19) فشل المرشح بسبب تغيير اتجاه الحركة (Guo,2022).

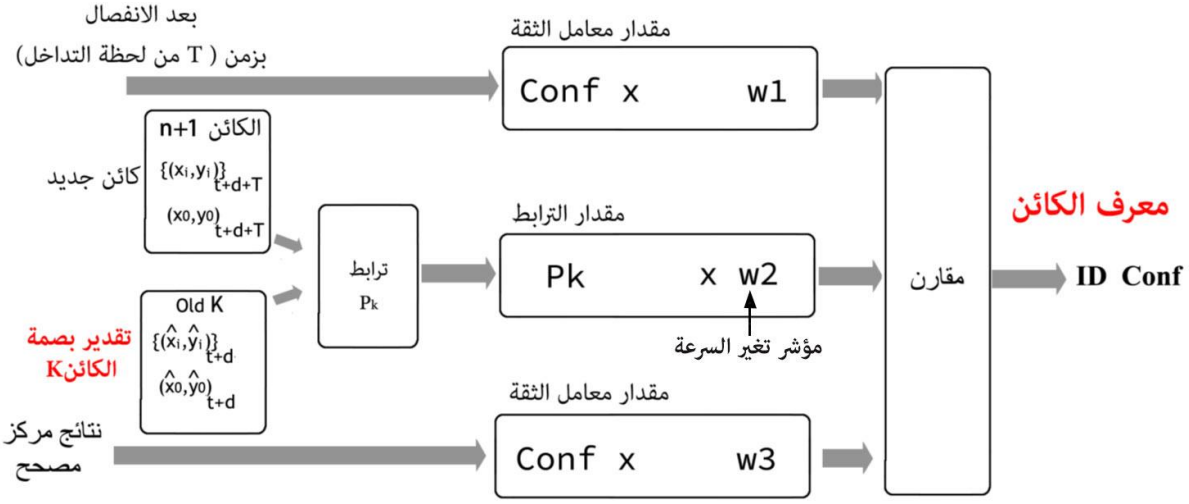
4-4-5 إجراءات التحسين على الخوارزمية

يبين الشكل (5-20) مخططا تدفقياً لتصرف الخوارزمية قبل التداخل وقد تم ادخال التحسينات السابقة عليها



الشكل (5-20) تصرف الخوارزمية المحسنة قبل التداخل (الباحث)

يلاحظ في الشكل (5-20) تصحيح المراكز للكائنات بعد نتائج الملاحقة مما سيعطي نتائج أكثر ضبطاً لمراكز الأجزاء بالتالي تقدير أفضل للمرشح كالمن، كذلك نلاحظ استنتاج مؤشر التسارع (تغيير السرعة) للكائن من أجل التثقيل في المرحلة التالية وذلك كما هو موضح في الشكل (5-21)



الشكل (5-21) تصرف الخوارزمية المحسنة بعد التداخل (الباحث)

ويلاحظ في هذه المرحلة أن المقارنة الثالثة ستتم مع تقديرات المراكز المصححة للكائنات، مما يزيد من الكفاءة.

5-4-5 اختبارات الخوارزمية المحسنة

مثلما تم في الفقرة (5-4-2) في المرحلة الأولى يتم تنفيذ خوارزمية الملاحقة المطورة باعتماد المعاملات التي توصلنا لها في الفقرة (5-4-1) على الصور الحرارية من بيانات ADAS-FlIR وتقارن مع نتائج الدراسة المرجعية (Ibrahim,2023) ومثيلتها طريقة خوارزمية YOLOv5-DeepSORT في الدراسة المرجعية (Sanyam, 2022) وكذلك تقارن مع الخوارزمية ذاتها بكاشف حديث YOLOv8-DeepSORT . في المرحلة الثانية يتم تنفيذ الخوارزمية السابقة بالإضافة لخوارزمية الملاحقة المطورة بالاعتماد على ذات المعاملات، على الصور الحرارية في بيانات مخصصة للتداخل والتي صممت للطريقة المطورة.

يبين الجدول (5-4) نتائج الاختبارات بمعايير الدقة MOTA ومعيار مؤشر التعريف IDF1

الجدول (5-4) اختبارات خوارزمية الملاحقة المطورة (الباحث)

Manar	IDF1	MOTA	IDF1	MOTA
	ADAS- FLIR Data Set		Our Data Set	
(Sanyam, 2022)	26.6	41	25.1	39.5
(Ibrahim,2023)	54.6	**63.8**	-	-
YOLOv8-DeepSORT	54.1	63.6	50.3	59.2
Proposed	**57.1**	63.4	**54.4**	**59.4**
Improvement	**+4.6%**	-0.6%	**+8.1%**	**+0.4%**

يلاحظ من الجداول بأن هناك نتائج جيدة للخوارزمية المقترحة في المعامل IDF1 تعادل بالمجمل 8.1% بالنسبة لأعلى نتائج في الدراسات المرجعية مقابل تقدم بسيط في معيار الدقة العامة MOTA. ونعلل ذلك بأن الخوارزمية المقترحة أدت أداءً جيدا نسبياً بالملاحقة ولاسيما الأداء وباتجاه الحفاظ على هوية الكائنات بعد التداخل. لكن هناك معاملات أخرى لم تلحظها وهو شيء طبيعي كون الخوارزمية قد خُصصت لتمييز الكائنات بعد التداخل.

إن الهدف في الخوارزمية المقترحة يلامس بشكل أكبر مؤشر تبديل الهوية. ومؤشر تبديل الهوية (IDSW) متضمن في MOTA الذي يستخدم مفهوم تتبع الهوية لمقاييس اكتشاف الكائنات. تلجأ بعض الأبحاث لاستخدام المعيار MOTA كما ورد بالعلاقة والبعض الآخر يلجأ لتثقيل معاملات المعيار حسب التطبيق والبحث بالثوابت c1, c2, c3، ويتم حساب مقياس MOTA وفق العلاقة (5-4)، (Liang, 2020) عبر جميع الإطارات:

$$MOTA = 1 - \frac{\sum_t (c1FN_t + c2FP_t + c3IDSW_t)}{\sum_t GT_t} \qquad (4\text{-}5)$$

البعض يعتبر c1=1 وآخرون في بعض الأعمال يعتبرون $c3 = log_{10}$ (Liang, 2020).

ونظراً لندرة وجود دراسات وتجارب معيارية لخوارزميات ملاحقة حديثة على مشاهد حرارية مخصصة لمشاهد التداخل فقط. سوف يتم تطبيق هذه الخوارزمية DeepSORT باستخدام YOLO V8 على مشاهد منتقاة من البيانات الحرارية المعيارية والبيانات التي تم تصميمها خصيصا للاختبار وستقارن النتائج بالتركيز على مؤشر تبديل الهوية مع نتائج تطبيق الخوارزمية المقترحة على ذات البيانات.

تجربة على سيناريو حركة معقد

تمت التجربة بالكاميرا الحرارية لمتطوعين وفق الشروط التالية:

- المشهد ليلي، والمسافة 35 متر
- عدد الإطارات 106، زمن الإطار 40 m.s
- الميل الشاقولي لمحور الكاميرا 5 درجات

يتم اجراء تداخل على مرحلتين الأولى بسرعة ثابتة للشخصين، والثانية بسرعة متغيرة أثناء التداخل، كذلك يُجرى في ذات السيناريو اختباراً لتغيير شكل الهدف، ولتسهيل البيان يتم شرح المراحل الأربعة كل على حدة. **في المرحلة الأولى** الشكل (5-22) تداخل شخصان، الشخص برقم التعريف ID:2 وهو شخص متوقف (مؤشر أزرق). والشخص برقم التعريف ID:3 وهو شخص يمشي يساراً بسرعة ثابتة (مؤشر أخضر) ويتقاطع مع ID:2.

الشكل (5-22) اختبار الخوارزمية في المرحلة الأولى على مشهد معقد (الباحث)

في المرحلة الثانية الشكل (5-23) تغير شكل الكائن الشخص الذي سيأخذ رقم التعريف ID:1 (المؤشر الزهري) الذي كان يلتقط شيئا من الأرض ووقف فجأة.

الشكل (5-23) اختبار الخوارزمية في المرحلة الثانية على مشهد معقد (الباحث)

في المرحلة الثالثة الشكل (5-24) تغير شكل الكائن الشخص الذي سيأخذ رقم التعريف ID:3 (المؤشر الزهري) حيث يقوم بالدوران حول نفسه.

في المرحلة الرابعة الشكل (5-24) يلاحظ تداخل شخصان، الشخص برقم التعريف ID:3 (المؤشر الأخضر) وهو شخص يمشي باتجاه الشخص المؤشر إليه باللون الزهري برقم التعريف ID:1 وهو شخص شبه متوقف، وقد قرر أخيراً المسير يساراً لحظة تداخله مع الشخص الثاني الذي تباطئ مما أدى لأن يسبق الشخص الآخر.

الشكل (5-24) اختبار الخوارزمية في المرحلة الثالثة على مشهد معقد (الباحث)

إن هذا الجزء من السيناريو يعد من أصعب اللحظات للملاحق الشكل (5-25) حيث تم تغيير السرعة للمتحركين في لحظة التلاقي كذلك الحال بالنسبة للتسارع، وهو ما يؤدي إلى فشل نموذج الحركة في خوارزميات الملاحقة.

الشكل (5-25) اختبار الخوارزمية في المرحلة الرابعة على مشهد معقد (الباحث)

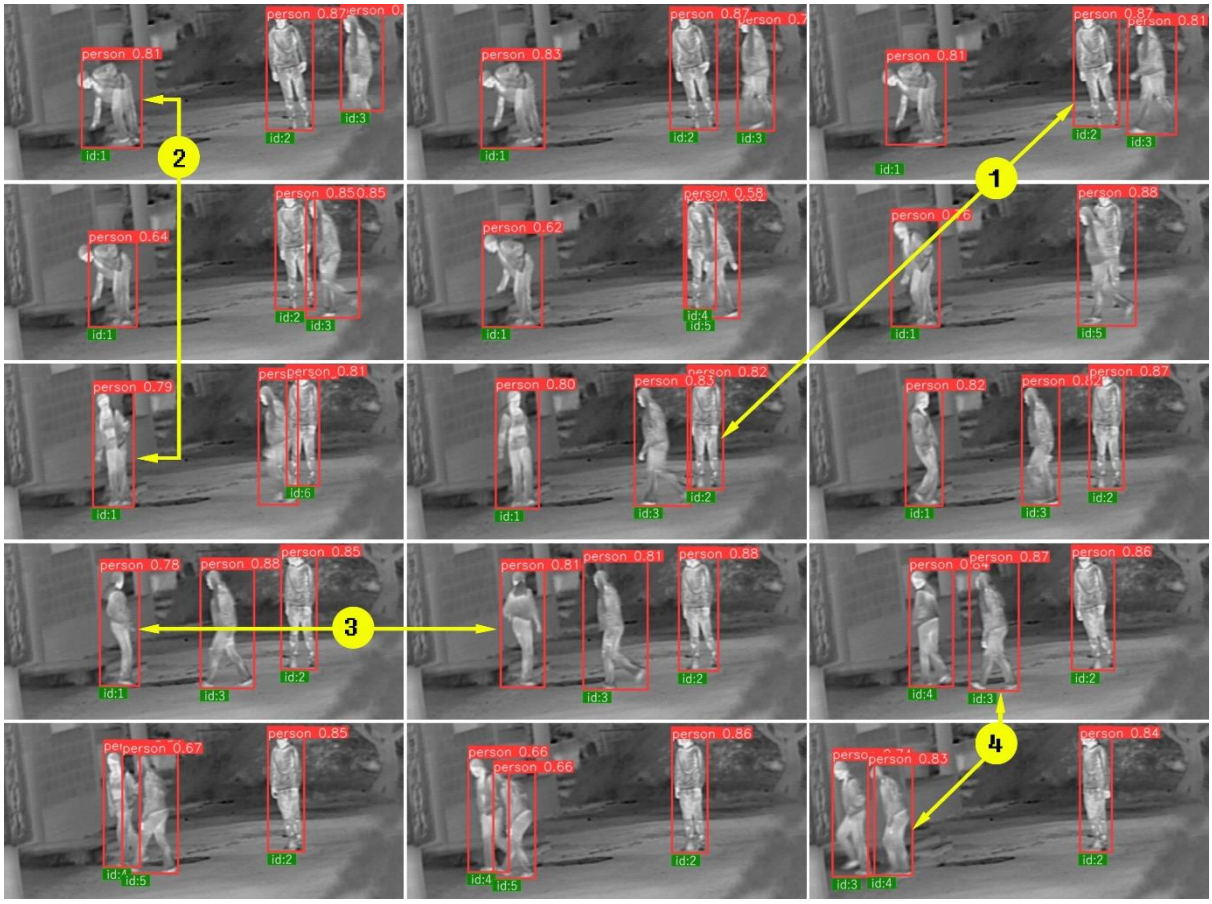
يبين الشكل (5-26) نتائج الملاحقة باعتماد خوارزميةDeepSORT باستخدام YOLO ويلاحظ ما يلي:

في الحالة (1) حالة التداخل: تمكن الملاحق DeepSORT من استدراك ضياع الهدف وتبديل معرفه وعاد باستخدام نموذجه الحركي لتعريف الأهداف المتداخلة بشكل صحيح.

كذلك **في الحالة (2)** حالة الملاحقة مع تغير تدريجي بالشكل، تمكن الملاحق من ملاحقة سمات الشكل المتغيرة للهدف.

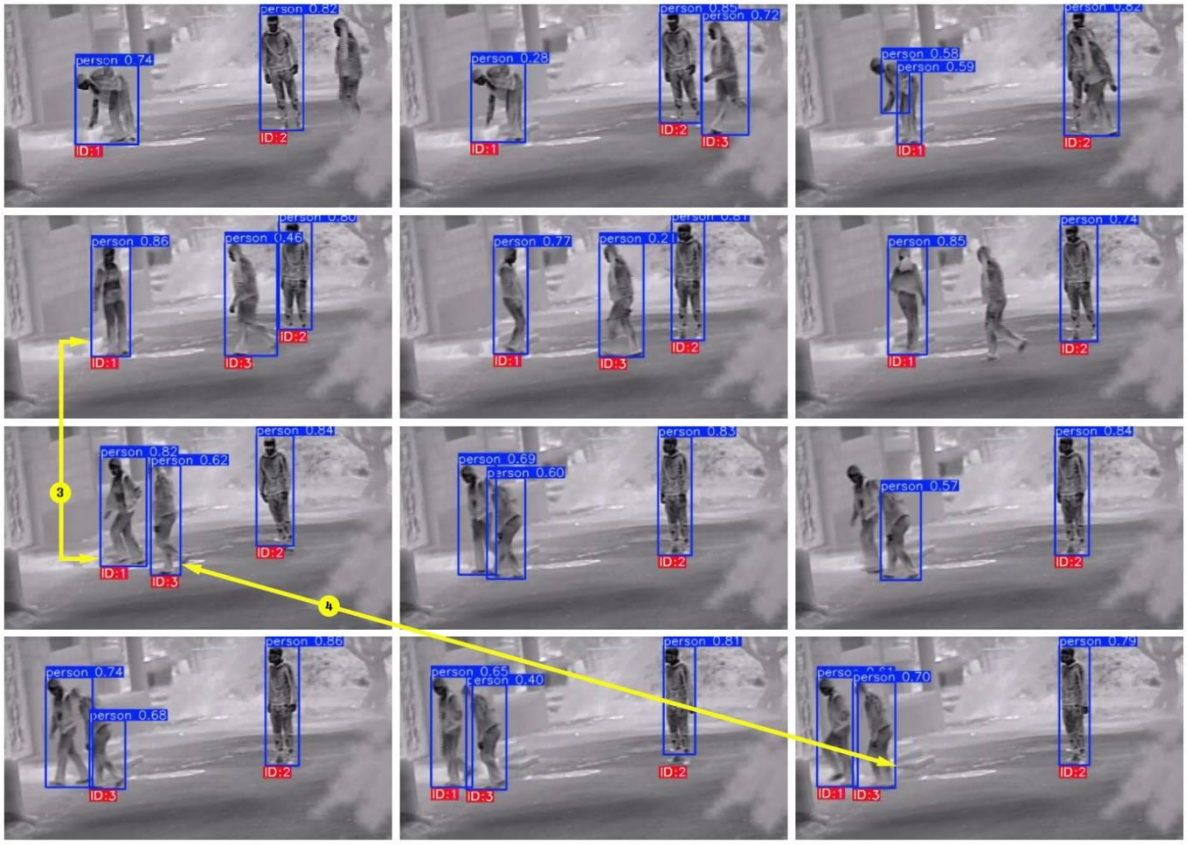
في الحالة (3) حالة الملاحقة مع دوران لجسم الكائن، أدت لتغير سريع بالشكل وضياع الكائن السابق ID:1 من حلبة الملاحقة، ليقوم بعدها الملاحق بتعريف جديد له ID:4.

في الحالة (4) لدى بدء التحام الكائنين يقدر الملاحق بأن الأهداف السابقة ضاعت ويرمز الهدفين برموز جديدة. وبسبب تبادل السرعتين تقريبا فشل نموذج الحركة بعد الانفصال من تحديد المعرف الصحيح حيث تم تبادل المعرفات.



الشكل (5-26) نتائج الملاحقة باعتماد خوارزميةDeepSORT باستخدام YOLO. (الباحث)

يبين الشكل (5-27) نتائج الملاحقة باعتماد الخوارزمية المطورة لنفس اللقطات الاختبارية والحالات المدروسة، ويلاحظ أنه لا يوجد أخطاء هنا في كافة العمليات. ففي الحالة (1) والحالة (2) تتم الأمور بنجاح كما في الخوارزمية السابقة. أما في حالتي الفشل السابقتين (3) و (4)، يلاحظ هنا أن خوارزميتنا قد نجحت في تخطي إعادة تمييز الأهداف المتداخلة. وهي حالتي الدوران وتغيير الشكل والاحتجاب.



الشكل (5-27) نتائج الملاحقة باعتماد الخوارزمية المطورة. (الباحث)

يتم اختيار خوارزمية أخرى أثبتت فعاليتها في إعادة تعريف الكائن في ملاحقة الكائنات المتعددة على مشاهد مرئية وهي DetTrack وهي المستخدمة ضمن الدراسة المرجعية في البحث (Gao, 2023)، حيث وصلت الدقة MOTA إلى 73% ومدى الحفاظ على الهوية IDF1 إلى 70.9%على المشاهد المرئية لمجموعات MOT20. ستستخدم هنا على مشاهد حرارية من مجموعتنا الخاصة والتي انتخبت خصيصا لاختبارات التمييز بعد التداخل، وذلك بالإضافة لخوارزمية DetTrack تم اختبار خوارزمية DeepSORT باستخدام كاشف YOLO والمستخدمة بالدراسة المرجعية (Sanyam, 2022) كذلك الدراسة (Ibrahim,2023) على ذات المشاهد المنتقاة من البيانات الحرارية الخاصة بنا وستقارن النتائج بالتركيز على مؤشر الحفاظ على الهوية والدقة ومعدل الإطارات. إن الجداول التالية تبين نتائج الاختبارات على مجموعة من البيانات التي قسمت لثمان مشاهد كل مشهد يتم التركيز على عدد مختلف من الأشخاص بين الشخصين وتسعة أشخاص

الجدول (5-5) نتائج اختبار مؤشر استعادة الهوية IDF1

Proposed IDF1	DetTrack IDF1	YOLOv8DeepSort IDF1	YOLOv5DeepSort IDF1	Person No.
59	51.9	51.6	45.2	2
57.1	52.3	51.4	45.4	3
57.9	53.3	50.7	42.9	4
56.8	49.9	47.1	43.3	5
56.5	48.1	47.9	40.4	6
54.9	52	50.1	38.1	7
50.1	**50.3**	48.9	26.9	8
40	**50.9**	47.5	26.5	9

الجدول (5-6) نتائج اختبار مؤشر الدقة لملاحقة الكائنات المتعددة MOTA

Proposed MOTA	DetTrack MOTA	YoloDeepSort MOTA	YOLOv5DeepSort MOTA	Person No.
64.6	64	63.3	25.1	2
64	63.5	63.1	26.3	3
64.5	63.2	63.4	24.9	4
64	63.5	63.1	24.3	5
63.6	63.3	63.4	24.4	6
61.8	**63.7**	63.1	24.1	7
61	**62.5**	62.4	24	8
59	**62.3**	62.1	23.7	9

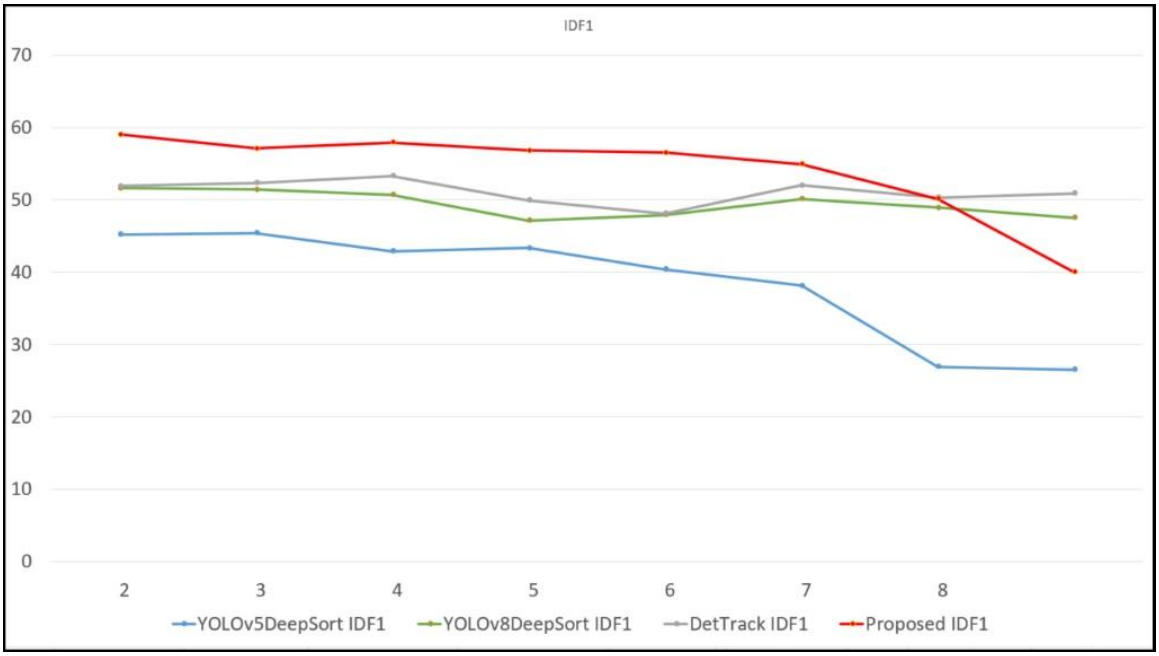
الجدول (5-7) نتائج اختبار دورة المعالجة T

Proposed T	DetTrack T	YoloDeepSort T	YOLOv5DeepSort T	Perso No.
32.9	35.1	38.5	33.3	2
35.1	36.3	37.5	33.2	3
35.2	36.4	38.1	35	4
35.9	36.1	37.9	37.5	5
40.1	**37.6**	38.1	39.6	6
45.5	**37.3**	38.3	39	7
47.4	**38.6**	39	39.6	8
55.4	**39.5**	40.7	39.8	9

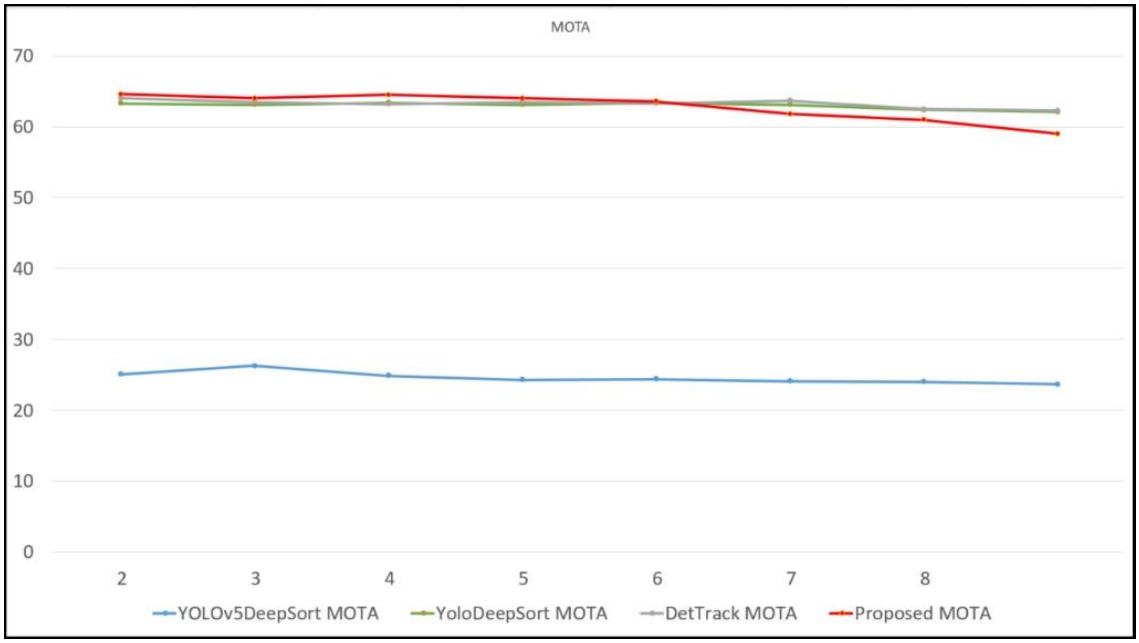
الجدول (5-8) نتائج اختبار معدل الإطارات FPS

Proposed FPS	DetTrack FPS	YoloDeepSort FPS	YOLOv5DeepSort FPS	Person No.
30.40	28.49	25.97	30.0	2
28.49	27.55	26.67	30.1	3
28.41	27.47	26.25	28.5	4
27.86	27.70	26.39	26.6	5
24.94	**26.60**	26.25	25.2	6
21.98	**26.81**	26.11	25.6	7
21.10	**25.91**	25.64	25.2	8
18.05	**25.32**	24.57	25.1	9

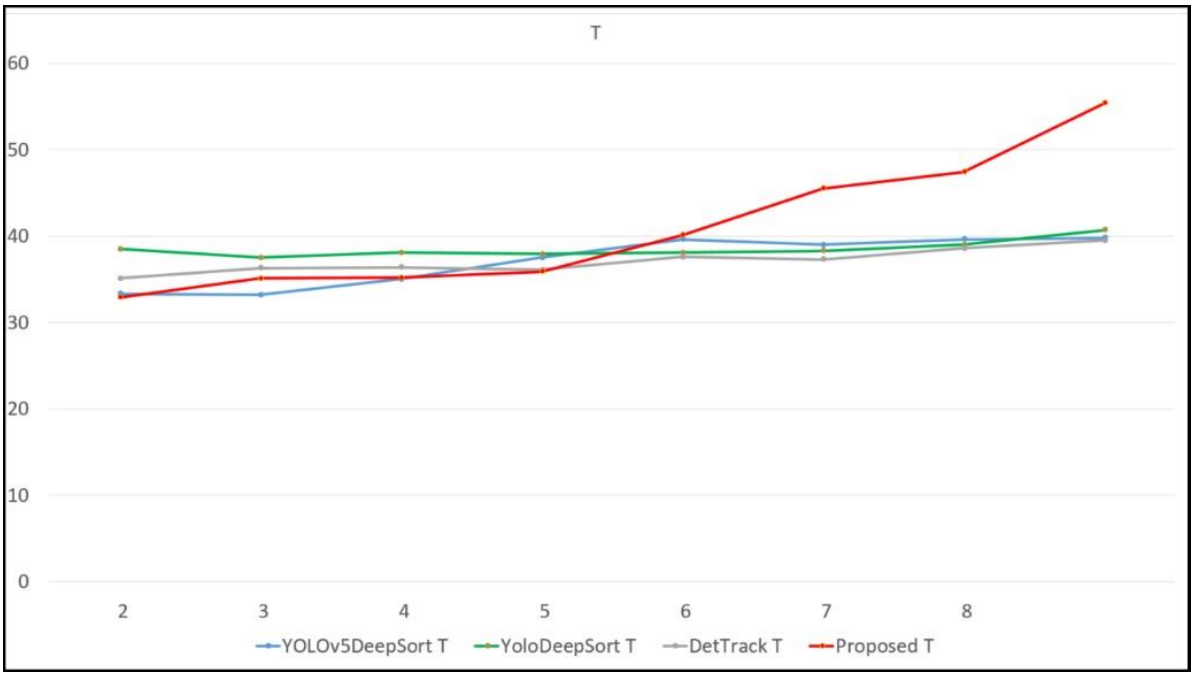
تبين الأشكال التالية نتائج الاختبارات المختلفة على مخططات بيانية، المخطط في الشكل (5-28) يمثل نتائج اختبار مؤشر استعادة الهوية IDF1 وهو المعامل الأهم لبيان قوة الطريقة المقترحة، وفيه يلاحظ أن طريقتنا تظل محافظة على تفوقها وخصوصاً في أعداد الكائنات الصغيرة، أما مخطط الدقة MOTA في الشكل (5-29) فيلاحظ تفوق صغير نسبي في الأعداد الصغيرة ولكن ضمن الأعداد الكبيرة يكون هنالك بعض التقصير، لكن ذلك لا ينتقص من أهمية الطريقة المطورة التي خُصصت بالأصل للتمييز بعد التداخل وليس لحالات الملاحقة العامة. أما في المخططان في الشكل (5-3) والشكل (5-31) يلاحظ أن الطريقة المقترحة تظل فعالة ضمن الزمن الحقيقي حتى في حالة الأعداد الكبيرة للكائنات ولا زالت أزمنة المعالجة بعيدة عن الحد الزمني المقترح.



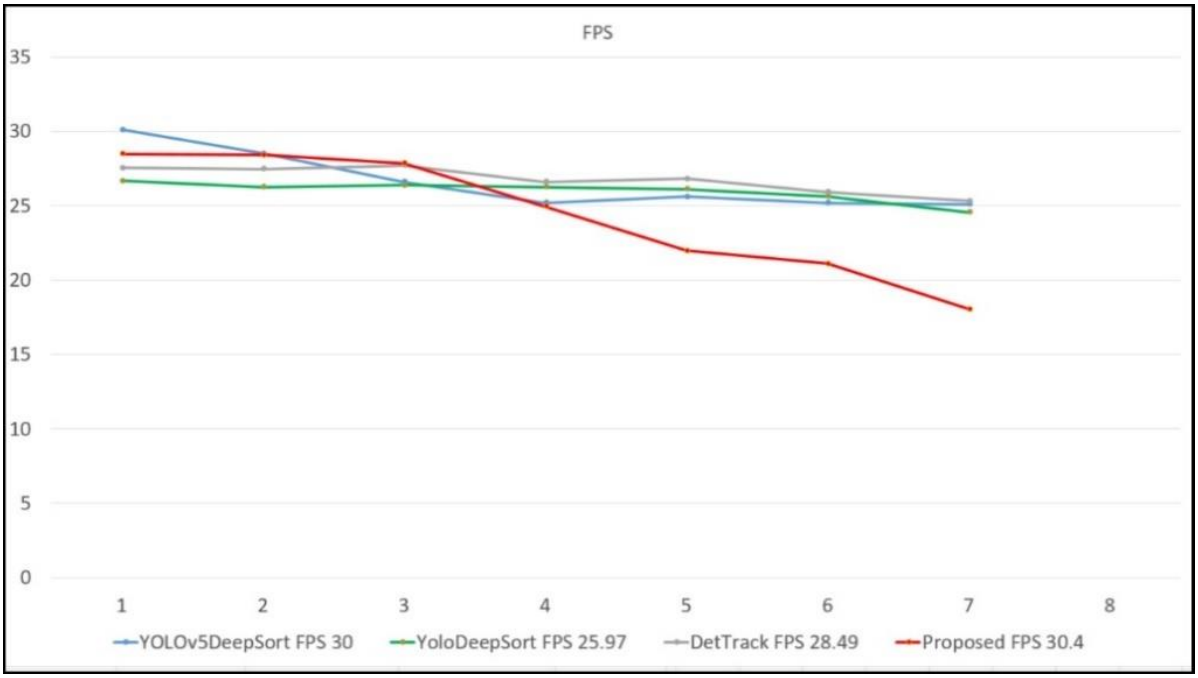
الشكل (5-28) نتائج اختبار مؤشر استعادة الهوية IDF1 (الباحث)



الشكل (5-29) نتائج اختبار الدقة MOTA (الباحث)



الشكل (5-30) نتائج اختبار زمن التنفيذ "ms" (الباحث)



الشكل (5-31) نتائج اختبار السرعة "عدد الإطارات في الثانية FPS" (الباحث)

5-5 الاستنتاجات

بدراسة نتائج تطبيق إجراءات الخوارزمية المقترحة لتمييز الكائنات الحرارية بعد التداخل يلاحظ أن هذه الإجراءات أدت لمجموعة من الاستنتاجات:

1. التحسين الجيد على أداء الكشف في الحالات المختلفة هو ضمن المجال % [3,10] وخصوصاً في ظروف الطقس السيء.
2. اختصار زمن التنفيذ للكشف في حالات أعداد الأهداف المكتشفة Target Number ضمن المشهد ذو الأعداد الصغيرة للكائنات الهدف، وعند الأعداد الكبيرة للأهداف ستكون هناك نسبة تأخير بسيطة من لاتتجاوز 20%.
3. تحسن نوعي بعدد الحالات الناجحة بالمتابعة للأهداف الحرارية بعد التداخل أكير 3%.
4. تحسن نوعي بزمن معالجة الإطار للحالات الناجحة في حالة الأعداد المنخفضة وصلت لـ 3%.
5. بقاء حالات النجاح بالمتابعة بعد التداخل بالزمن الحقيقي حتى لأعداد كبيرة مثلا 9 أشخاص.
6. تجاوزت إجراءات الخوارزمية المقترحة نوعا ما مشكلتَي الكائن المفصلي والكائن الحراري متغيري الشكل.

6-5 التوصيات

من نتائج هذا البحث يمكن تقديم التوصيات التالية:

1. الإجراءات بمجملها مستقلة عن أي خوارزمية ملاحقة بالتالي يمكن استخدامها مع أي خوارزمية متطورة، مما يزيد من نجاعتها.
2. عند وجود أعداد كبيرة من الكائنات سيؤدي لانخفاض سرعة المعالجة. لذلك من الممكن تنفيذ اجرائيات الخوارزمية بشكل انتقائي "تكيفي" أي عدم اتباع اجراءات الخوارزمبة بالكامل طيلة الملاحقة، وهو ما يمكن أن يؤدي لزيادة السرعة.
3. يكون النجاح بالتمييز الناتج مرهوناً بالعدد الصغير للأشخاص وزمن التداخل الصغير ولكن من الممكن زيادة النسبة هذه بزيادة زمن الحجب المعطى لنموذج الحركة. أو بزيادة تمييزية الكاميرا. ولكن ذلك بحاجة لاختبار مدى تأثيره على العمل بالزمن الحقيقي.
4. ارتبط نجاح النتائج مع عدم تجاوز الكائن لحدود السرعة الزاوية المحددة (2Deg/s). لزيادة نجاح النتائج يجب استخدام كاميرا بزاوية فتحة أكبر وذلك يفرض أيضاً زيادة تمييزية الكاميرا.

الأبحاث المنشورة

البحث الأول:

تعزيز سمات كائن حراري ضعيف المعالم من أجل الملاحقة والتصنيف.

مجلة جامعة دمشق للعلوم الهندسية. تاريخ الإيداع 2023/12/18، تاريخ القبول: 2024/2/12

البحث الثاني:

تحسين كشف الأشخاص في ظروف الإضاءة السيئة باستخدام نموذج كشف مدرب مسبقاً مع الكاميرات الحرارية.

مجلة جامعة دمشق للعلوم الهندسية. تاريخ الإيداع: 2024/6/30، تاريخ القبول: 2024/9/22

البحث الثالث:

تعريف الأشخاص بالصورة الحرارية.

مجلة جامعة دمشق للعلوم الهندسية. تاريخ الإيداع: 2024/11/25، تاريخ القبول: 2025/1/14

الآفاق المستقبلية Future Work

يمكن تطوير العمل باتجاه محاور متعددة:

- العمل على تصميم مجموعة بيانات حرارية كبيرة خاصة بمشاكل التداخل لتعزيز التعلم وتدريب نموذج الكشف عليه.
- استخدام نماذج كشف حديثة بذات الاجرائيات لاستنتاج قواعد عمل أكبر.
- ان معالجة الخوارزمية لشكل الكائن المتغير هي ميزة من الممكن استخدامها ذاتها أو بقليل من التعديل للمساهمة بتمييز بعض الكائنات الحرارية الأخرى بالإضافة للأشخاص.
- إن استخدام منصة للكاميرا الحرارية، والتحكم بحركتها سيؤدي بالتأكيد لزيادة إمكانية الخوارزمية من ناحية (سرعة زاوية أكبر) للكائنات المراقبة ولكن ذلك بحاجة لتطوير نظام تحريك للملاحقة وإدخال معاملاته ضمن معاملات الخوارزمية المقترحة.

قائمة المراجع

1. Aabidi, M. H., Makrani, A. E. , Jabir, B. , & Zaimi, I. (2023). A Model **Proposal for Enhancing Leaf Disease Detection Using Convolutional Neural Networks (CNN): Case Study**. DOI: https://doi.org/10.3991/ijoe.v19i12.40329
2. Ashrafi, R. Azarbayjania, M. and Tabkhi, H. (2022**). A Fully Annotated Thermal Infrared Face Dataset with Various Environmental Conditions and Distances**. Cornell University, arXiv:2205.02093
3. Bashar, A. (2023). **Object Detection and Classification in Videos using Deep Learning**. ttps://medium.com/
4. Bashar, M., Islam, S., Hussain, K. K., Hasan, B., Rahman, A., & Kabir, H. (2022). **Multiple Object Tracking in Recent Times: A Literature Review**
5. Becker, A. (2023). **Kalman Filter from the Ground Up**. https://www.kalmanfilter.net
6. Bhuiyan, A., Bappee, F. K., Islam, M., & Hasan, M. (2023). **Person Re-Identification with RGB–D and RGB–IR Sensors: A Comprehensive Survey**
7. Borah, B., Saikia, & R, Das. P. (2022) **Animal Motion Tracking in Forest: Using Machine Vision Technology**. DOI: 10.55041/IJSREM15501. Conference: International Journal of Scientific Research in Engineering and Management (IJSREM)
8. Brenner, M., Reyes, N.H., Susnjak, T., & Barczak, A.L.C. (2023). **RGB-d and thermal sensor fusion: A systematic literature review**. IEEE Access 11, 82410–82442 (2023). https://doi.org/10.1109/access.2023.3301119
9. Broyles, D., Hayner, C.R., & Leung, K.: WiSARD, (2022). **A labeled visual and thermal image dataset for wilderness search and rescue**. In: 2022 IEEE/RSJ International. Conference on Intelligent Robots and Systems (IROS). IEEE (oct 2022). https: //doi.org/10.1109/iros47612.2022.9981298
10. Chen, G. and Wang, W. (2022). **Target Recognition in Infrared Circumferential Scanning System via Deep Convolutional Neural**. *Sensors*, *20*(7)
11. **colab**. https://colab.research.google.com
12. Gonzalez, R. C., & Woods, R. E. (2018). **Digital Image Processing**, Fourth Edition, Global Edition. Pearson Education. ISBN 978-0-13-335672-4
13. COWTON, J. ILIAS KYRIAZAKIS2, I. & BACARDIT1, J. (2019). **Tracking And Behaviour Metric Extraction Using Deep Learning.** Citation information: DOI 10.1109/ACCESS.2019.2933060, IEEE Access
14. Devi, N., & Naidu, S. V., (2023). **Leveraging Deep Learning Models for Weld Defect Detection**. Industrial Engineering Journal. ISSN: 0970-2555Volume : 52, Issue 12, No. 4
15. Ding, M. Guan, S. Liu,H. and Yu, K. (2023). TIR-YOLO-ADAS: **A thermal infrared object detection framework for advanced driver assistance systems**. IET Intelligent Transport Systems. DOI: 10.1049/itr2.12471
16. Elliott, J. (2022). **Cross Correlation Computations**. In:Learn Particle Image Velocimetry. College of Engineering UtahState University. http://www.learnpiv.org/correlationComputations/
17. Encord. 2025. **Top 10 Video Object Tracking Algorithms in 2025.** https://encord.com/blog/video-object-tracking-algorithms/

18. Encord1. (2025). **What is YOLO Object Detection**. https://encord.com/blog/yolo-object-detection-guide/

19. Encord2. (2024). **Understanding Object Tracking Metrics**. https://miguel-mendez-ai.com/2024/08/25/mot-tracking-metrics

20. Flir. (2023) **FREE Teledyne FLIR Thermal Dataset for Algorithm Training**. https://www.flir.com/oem/adas/adas-dataset-form/

21. Frank, H. (2023). **Electromagnetic Wave Spectrum**. https://www.e-education.psu.edu/geog160/node/1958 CC-BY-SA 4.0 International

22. Ganguly, S. (2021). **Neural Networks**. In: Quantum Machine Learning: An Applied Approach. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-7098-1_3

23. Guo, S., Wang, S. Yang, Z., Wang, L., Zhang, H., Gao, & Y., Guo, J.(2022**). A Review of Deep Learning-Based Visual Multi-Object Tracking Algorithms for Autonomous Driving**. Applied Sciences 12(21):10741. DOI: 10.3390/app122110741.

24. Gahramanova. (2019). **Locating Centers of Mass with Image Processing,** UJMM Volume 10 | 2019 Fall 2019

25. Gao, X., Wang, Z., Wang, X., Zhuang, S., Wang, H. 2023. **DetTrack: An Algorithm for Multiple Object Tracking by Improving Occlusion Object Detection.** Electronics 13(1):9. DOI: 10.3390/electronics13010091. License CC BY 4.0

26. **github**. https://ikomia-dev.github.io

27. Hardan, H. (2018). **Sharpening Spatial Filters (high pass).** in: Image Processing. https://www.philadelphia.edu.jo/academics/hhardan/uploads/Image_Processing-ch3_part_5.pdf

28. Hande, S.S. (2021). **Image Edge Detection by Mean Difference Thresholding.** In: Nath, V., Mandal, J. (eds) Nanoelectronics, Circuits and Communication Systems. Lecture Notes in Electrical Engineering, vol 692. Springer, Singapore. https://doi.org/10.1007/978-981-15-7486-3_30

29. Hu, Q. Scott, A., Yeung, C., & Fujii1, K. (2024). **Basketball-SORT: An Association Method for Complex Multi-object Occlusion Problems in Basketball Multi-object Tracking**. Multimedia Tools and Applications 83(38):86281-86297. DOI: 10.1007/s11042-024-20360-2

30. Ibrahim, N., Darlis, A. R. Subiantoro, A. ,Yusivar, F., & Kusumoputro, B. (2023). **Online Multi-Object Tracking (Mot) of Sequential Thermal Images Based on Deep Appearance Features Using Yolo-Deepsort**. SSRN Electronic Journal. DOI: 10.2139/ssrn.4364548

31. Jiang, Y·, Li, X., Luo, H., & Yin, S. Kaynak, O. (2022). **Quo vadis artificial intelligence**. *Discov Artif Intell* **2**, 4 (2022). https://doi.org/10.1007/s44163-022-00022-8

32. Johnsen, M., (2024). **Machine Learning**. ISBN 9798332357909

33. Juranek, L., Stastny, J., & Skorpil, V. 2018. **Effect of Low-Pass Filters as a Shi-Tomasi Corner Detector's Window Functions**. DOI: 10.1109/TSP.2018.8441178 Conference: 41st International Conference on Telecommunications and Signal Processing, TSP 2018

34. Kang, J., Tariq, S., Oh, H., Woo, S. (2022). **A Survey of Deep Learning-Based Object Detection Methods and Datasets for Overhead Imagery.** IEEE Access 10:1-1**.** DOI: 10.1109/ACCESS.2022.3149052

35. karen. (2019) **Smart Home Security using Image Detection Neural Nets**. https://medium.com/@karencyen/jolt-and-unbolt-5fcfb68678bc

36. Kaur, R., & Singh, S. (2023). A comprehensive review of object detection with deep learning. Digital Signal Processing. Volume 132, Page 103812. https://viso.ai/deep-learning/object-tracking/

37. Kotai. (2022). **Object Counting Using YOLO Best Practices**. Kotai Electronics Pvt.Ltd.https://medium.com/@eleckotai/object-counting-using-yolo-best-practices-d25d8e55cd26.

38. 35. Lea, S. (2011). **Infrared heating for humans**, Ceramicx. Ireland. https://www.ceramicx.com/information/media/white-papers/infrared-heating-for-humans.

39. Liang, X. (2020), Ascend AI Processor Architecture and Programming. https://www.sciencedirect.com/topics/engineering/perceptron

40. Li, X., Liu, Q., Fan, N., He, Z., & Wang, H. (2018). **Hierarchical Spatial-aware Siamese Network for Thermal Infrared Object Tracking**. Knowledge-Based Systems. https://doi.org/10.1016/j.knosys.2018.12.011

41. Liu, Q., Lua, X., He, Z., Zhang,C., & Chen,W. S. (2017). **Deep Convolutional Neural Networks for Thermal Infrared Object Tracking**. Knowledge-Based Systems. https://www.researchgate.net/publication/318714772. DOI: 10.1016/j.knosys.2017.07.032

42. Ma, L. V., Nguyen, T.T. D., Shim,C., Kim, D, Ha, N., & Jeon, M. (2024). **Visual multi-object tracking with re-identification and occlusion handling using labeled random finite sets**

43. Malche, T., (2024). **Edge Detection in Image Processing**: An Introduction. https://blog.roboflow.com/edge-detection/

44. Manganiello, F. (2021). **Neural Networks**. In: Computer Vision with Maker Tech. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-6821-6_2

45. **matplotlib**. https://pypi.org/project/matplotlib

46. Medium. (2019). **Introduction to FAST (Features from Accelerated Segment Test).** https://medium.com/@deepanshut041/introduction-to-fast-features-from-accelerated-segment-test-4ed33dde6d65

47. Méndez, M. (2024**). Understanding Object Tracking Metrics**. https://miguel-mendez-ai.com/2024/08/25/mot-tracking-metrics

48. Mishra, D. (2022).**Deep Learning based Object Detection Methods: A Review**. Medicon Engineering Themes Volume 2 Issue 4 April 2022

49. **numpy**. https://numpy.org

50. Luiten, J. (2021). **How to evaluate tracking with the HOTA metrics**. https://jonathonluiten.medium.com/how-to-evaluate-tracking-with-the-hota-metrics-754036d183e1

51. Olszewska, J. I. (2016). **Automated Face Recognition: Challenges and Solutions.** DOI: 10.5772/66013. https://www.intechopen.com/chapters/52911

52. **opencv**. https://pypi.org/project/opencv-python

53. **os**. https://docs.python.org/3/library/os.html

54. Parker, J. R. (2010**) Algorithms for Image Processing and Computer Vision**, 2nd Edition. Wiley. ISBN: 978-1-118-02188-0. November 2010. 512 pages

55. **pillow**. https://pypi.org/project/pillow

56. Rudrabhatla, C. (2020). **A Quantitative Approach for Estimating the Scaling Thresholds and Step Policies in a Distributed Microservice Architecture**. IEEE Access 8:180246 – 180254. DOI: 10.1109/ACCESS.2020.3028310

57. Sadli, R. (2023) **Introduction to Multi-Object Tracking** (MOT). in: Machine Learning Space. https://machinelearningspace.com/introduction-to-multi-object-tracking/

58. Saha, S. 2023. **A Guide to Convolutional Neural Networks the ELI5 way**. https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/

59. Sanyam. (2022). **Understanding Multiple Object Tracking using DeepSORT**.learnopencv. https://learnopencv.com/understanding-multiple-object-tracking-using-deepsort/#Evaluating-Performance

60. **scikit**. https://scikit-image.org

61. Serpush, F. & Rezaei, M.(2020). **Complex Human Action Recognition in Live Videos Using Hybrid FR-DL Method**.
SSRN: https://ssrn.com/abstract=3650450 or http://dx.doi.org/10.2139/ssrn.3650450

62. Seseri, R. (2023). **AI Atlas #5 Deep Learning**. https://www.linkedin.com/pulse/ai-atlas-5-deep-learning-rudina-seseri/

63. Shaw, R. N., Ghosh, A., Balas, V. E., & Bianchini, M. (2021**), Artificial Intelligence for Future Generation Robotics.** Paperback ISBN: 9780323854986. eBook ISBN: 9780323857994

64. Shin, U., Park, J., & Kweon, I. S. (2023). **Deep depth estimation from thermal image**. In: 2023. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (jun 2023). https://doi.org/10.1109/cvpr52729.2023.00107

65. Shin. Y, Shin, H. Ok. J, Back, M. Youn, J. and Kim, S. (2024). **DCEF2-YOLO: Aerial Detection YOLO with Deformable Convolution–Efficient Feature Fusion for Small Target Detection**

66. Shin. Y, Shin, H. Ok. J, Back, M. Youn, J. and Kim, S. (2024). **DCEF2-YOLO: Aerial Detection YOLO with Deformable Convolution–Efficient Feature Fusion for Small Target Detection**

67. Simhi, N. and Yovel, G. (2020). **Can we recognize people based on their body-alone? The roles of body motion and whole person context**. Vision Research. Volume 176

68. Syiee. (2023). **Computer Vision: Edge Detection**. https://velog.io/@syiee/ Computer-Vision-Edge-Detection

69. Stanford. (2024). **Image Processing**. https://ai.stanford.edu/~syyeung/cvweb/tutorial1.html

70. Sun, M., Zhang, H., Huang, Z., Luo, Y. & Li, Y. (2021) **Road infrared target detection with I-YOLO.** IET Image Processing 16(4)**.** DOI: 10.1049/ipr2.12331**.** License CC BY 4.0

71. Szeliski, R., (2022). **Algorithms and Applications**. in: Computer Vision. 2nd ed. Springer.

72. Tarabalka, Y., Benediktsson, J. A., Chanussot, J., (2022). **Mathematical morphology**. yuliya.tarabalka@hyperinet.eu

73. Tao, W. (2024). **Analysis the improvements of YOLOv5 algorithms: NRTYOLO, MR-YOLO and YPH-YOLOv5. Proceedings of the 4th International Conference on Signal Processing and Machine Learning** DOI: 10.54254/2755-2721/54/20241466

74. Thapa. M. S. 2022**. Moravec corner detection**. Medium https://mahendrathapa.medium.com/moravec-corner-detector-5191f1c04b30

75. Tian, R., Sun, G., Liu, X., & Zheng, B. (2021). **Sobel Edge Detection Based on Weighted Nuclear Norm Minimization Image Denoising**. *Electronics* 2021, *10*(6), 655; https://doi.org/10.3390/electronics10060655

76. Wirth, M. (2004). **Shape Analysis & Measurement**. *University of Guelph, Computing and Information Science* Amina

77. Wu, H. & Liu, G. (2022) **A dynamic infrared object tracking algorithm by frame differencing**. Infrared Physics & Technology 127(7):104384. DOI: 10.1016/j.infrared.2022.104384

78. Vijayakumar, A. and Vairavasundaram, S. (2024). **YOLO-based Object Detection Models: A Review and its Applications**. https://doi.org/10.1007/s11042-024-18872-y

79. Yani, M. Si, S., Irawan, B., Setianingsih, C. 2019. **Application of Transfer Learning Using Convolutional Neural Network Method for Early Detection of Terry's Nail.** Journal of Physics Conference Series 1201(1):012052**.** DOI: 10.1088/1742-6596/1201/1/012052

80. Yuan, D., Shu, X., Liu,Q., Zhang,X., & He,Z. (2023). **Robust thermal infrared tracking via an adaptively multi-feature fusion model.** Neural Computing and Applications (2023) 35:3423–3434 https://doi.org/10.1007/s00521-022-07867-1

81. Yuheng, S., & Hao, Y. (2017). **Image Segmentation Algorithms Overview**. [cs.CV]. DOI: 10.48550/arXiv.1707.02051. https://www.researchgate.net/publication/318315979_Image_Segmentation_Algorithms_Overview

82. Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). **Object Detection in 20 Years: A Survey**. 2023, IEEE, Volume: 111 Issue: 3, https://ieeexplore.ieee.org/abstract/document/10028728