

Enhancing Localization and Orientation of Dental Tools Using Isolation and Deep Learning Algorithms: A Case Study Utilizing YOLOv5, GrabCut, and PCA

Dr. Iyad Hatem*
Aya Kheirbeq**

(Received 11 / 5 / 2024. Accepted 11 / 7 / 2024)

□ ABSTRACT □

Dental instrument retrieval during surgeries demands precise identification and localization to select the correct tool per the surgeon's requirements. Challenges with human dental assistants include tool misidentification, accidental contact with sharp tips, and potential infection exposure. This research proposes practical solutions for detecting, categorizing, isolating, and localizing specific dental instruments based on the dentist's needs.

We explore the application of the YOLOv5 deep learning algorithm for tool detection and classification, using predefined classes to determine Bounding Boxes (BBs). The Grabcut algorithm is then used to isolate the tool and create a foreground mask. Next, Principal Component Analysis (PCA) is employed for precise localization of the detected instruments.

Our approach leverages deep learning algorithms to accelerate the detection and classification of dental tools, integrating them with background isolation algorithms to obtain an ideal tool mask. Additionally, we determine the tool's orientation using eigenvectors to obtain the general orientation via PCA. The proposed model aims to ensure the lowest position error and minimum calculation time, even as the environment changes.

Keywords: YOLOv5, Tool Localization, Surgical Assistant, PCA, Grabcut Algorithm.

Copyright



:Tishreen University journal-Syria, The authors retain the copyright under a CC BY-NC-SA 04

* Professor, Department of Mechatronics Engineering, Faculty of mechanical and electrical engineering, Tishreen University, Lattakia, Syria. Email: iyadhatem@tishreen.edu.sy

** (Master student) Department of Mechatronics Engineering, Faculty of Mechanical and Electrical Engineering, Tishreen University, Lattakia, Syria. aya.kheirbeq@tishreen.edu.sy

تحسين عملية تحديد مواقع وتوجهات أدوات طبيب الأسنان باستخدام خوارزميات العزل والتعلم العميق: دراسة حالة باستخدام YOLOv5 و GrabCut مع PCA

- د. إياد حاتم
- إيه خيربك

(تاريخ الإيداع 11 / 5 / 2024. قَبْلُ للنشر في 11 / 7 / 2024)

□ ملخص □

تتطلب عملية مناولة أدوات طبيب الأسنان أثناء العمليات الجراحية تحديداً دقيقاً لموضعها وتوجيهها على طاولة العمل مع الاختيار الصحيح للأداة وفق صنفها تبعاً لمتطلبات الطبيب. تشمل التحديات التي يواجهها مساعده طبيب الأسنان البشريون الخطأ في التعرف على الأداة، احتمالية التعرض للأذى بسبب الأطراف الحادة، والتعرض المحتمل للعدوى. يقترح هذا البحث حلاً عملياً لاكتشاف وتصنيف وعزل وتحديد مواضع أدوات طبيب الأسنان وفق الصنف المرغوب والمحدد بناءً على احتياجات طبيب الأسنان لتكون خطوة بداية في عملية المناولة الروبوتية لأدوات الطبيب. يساهم تطبيق خوارزمية التعلم العميق YOLOv5 باكتشاف الأدوات وتصنيفها، باستخدام فئات محددة مسبقاً ليقوم بإنشاء أطر محيطية لها (Bounding Boxes (BBs)). يتم بعد ذلك استخدام خوارزمية Grabcut لإتمام عزل الأداة عن الخلفية ضمن الإطار المحيط السابق وإنشاء قناع عزل خاص بها. ونهايةً، يتم استخدام تحليل المكونات الرئيسية Principle Component Analysis (PCA) لتحديد موضع مركز الأداة وتوجيهها الدقيق للأدوات على مساحة العمل. يستفيد نهجنا من خوارزميات التعلم العميق لتسريع اكتشاف أدوات طب الأسنان وتصنيفها، ودمجها مع خوارزميات عزل الخلفية للحصول على قناع أداة مثالي. بالإضافة إلى ذلك، نحدد اتجاه الأداة باستخدام المتجهات الذاتية للحصول على الاتجاه العام عبر PCA. ويهدف النموذج المقترح إلى ضمان أقل خطأ موضعي وأقل وقت حسابي، حتى مع تغير البيئة.

الكلمات المفتاحية: شبكة التصنيف والتعرف YOLO الإصدار الخامس، تموضع الأدوات، المساعد الطبي، خوارزمية تحليل العناصر الرئيسية، خوارزمية عزل الخلفية Grabcut.



حقوق النشر : مجلة جامعة تشرين- سورية، يحتفظ المؤلفون بحقوق النشر بموجب الترخيص

CC BY-NC-SA 04

* أستاذ، قسم هندسة الميكاترونك، كلية الهندسة الميكانيكية والكهربائية، جامعة تشرين، اللاذقية سورية.
iyadhatem@tishreen.edu.sy
* * مهندس، (طالبة ماجستير)، قسم هندسة الميكاترونك، كلية الهندسة الميكانيكية والكهربائية، جامعة تشرين، اللاذقية، سورية.
aya.kheirbeq@tishreen.edu.sy

مقدمة:

يعد ضمان التحديد الدقيق لأدوات طبيب الأسنان وتحديد موضعها أثناء العمليات الجراحية أمراً بالغ الأهمية لنجاح وسلامة إجراءات طب الأسنان. تقع هذه المسؤولية تقليدياً على عاتق الممرضات ومساعدى الأطباء، لكنها مليئة بالتحديات مثل الإرهاق الجسدي الناتج عن الإجراءات المطولة وارتفاع مخاطر الإصابات، بما في ذلك إصابات الوخز بالإبر الناتجة عن الأدوات الحادة مثل شفرات المشرط والمشابك. هذه التحديات تعكس الحاجة الملحة لحلول متقدمة في تنظيم هذه الأدوات.

الغرض الأساسي من هذا البحث هو تطوير نظام موثوق يدمج عملية الكشف والعزل للأدوات مع تحديد موضعها باستخدام التقنيات المتقدمة. من خلال الاستفادة من خوارزمية التعلم العميق YOLOv5 لاكتشاف الكائنات وتصنيفها في الزمن الحقيقي، وخوارزمية Grabcut لعزل الخلفية بشكل فعال، وتحليل المكونات الرئيسية (PCA) لحساب التموضع الدقيق، يهدف هذا النظام إلى تعزيز سلامة وكفاءة جراحات الأسنان.

يمثل دمج خوارزميات التعلم العميق المتقدمة وتقنيات معالجة الصور في تطبيقات طب الأسنان خطوة إلى الأمام في الأتمتة الطبية والسلامة. من الناحية التكنولوجية، الهدف هو تطوير نظام يمكنه الكشف بدقة عن أدوات طبيب الأسنان وتحديد موضعها في الزمن الحقيقي، مما يقلل من الأخطاء البشرية والإجهاد الجسدي على الطاقم الطبي، ويعمل على تحسين إجراءات السلامة الطبية بشكل عام.

أهمية البحث وأهدافه:

تعد عملية تحديد موضع أداة معينة من بين مجموعة أدوات متشابهة بالبنية والشكل تحدياً يتطلب تقنيات متقدمة قادرة على تحصيل عدد كبير من الخواص والمميزات بأن معاً لإتمام عملية الاكتشاف وتحديد الصنف في آن معاً يضاف إلى ذلك عزل الكائن المحدد والمصنف على مستوى البكسل عن الخلفية المحيطة به مما يتطلب عمليات حاسوبية أكثر ونماذج مدربة على عمليات العزل بكميات أكبر. حددت [1] المناطق ذات الانحناءات الأعلى لأدوات لامعة ضمن بيئة صناعية باستخدام كاميرا متعددة الإضاءة لتحديد الانحناءات ومعرفة الأداة بناء على الانحناء الخاص بها الأمر الذي يصبح معقداً مع وجود عدد كبير من الأدوات. بينما ركزت [2] على تقنية المشاهد المتعددة المستندة إلى البيانات باستخدام واصفات تحويل السمات غير المتغيرة مع المقياس SIFT للتعرف على وضعية الكائن وتقديرها الأمر الذي أعطى انخفاضاً في الأداء عند قلة الواصفات الملتقطة ضمن إطار الصورة بالإضافة إلى محدودية التعامل مع عدد كبير من الأدوات. و عمل [3] على تقدير وضعية الأجسام اللامعة عديمة اللمس باستخدام نماذج عشوائية باستخدام كاميرا واحدة ومصادر إضاءة متعددة.

تناول [4] التعامل الآلي مع الأدوات الجراحية في الصواني المزدهمة، وذلك باستخدام شبكات التعلم العميق YOLO للكشف عن الأشياء وحل مشكلة التراكم. بينما [5] قدم إطار عمل ART-Net للكشف عن الأدوات الجراحية وتقسيمها وتقدير الوضع ثلاثي بمساعدة الكمبيوتر.

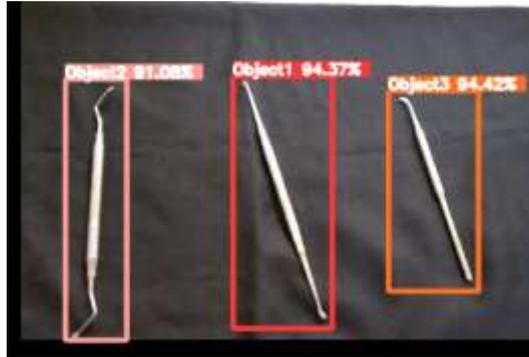
تتضح أهمية المناولة لأدوات طبيب الأسنان من خلال إصابات العمل المتعلقة بإجراءات طبيب الأسنان. حيث عرض [6] ارتفاع معدلات الإصابة بين مساعدى طبيب الأسنان مقارنة بالطلاب في كلية طب الأسنان بجامعة كوينزلاند، مما يسلط الضوء على الحاجة إلى تحسين التدابير والممارسات الوقائية.

في مجال اكتشاف أدوات طبيب الأسنان ومعرفة موضعها، يقترح هذا العمل نهجاً مدمجاً بين التقنيات التقليدية والذكاء الصناعي بالإضافة للتقنيات الرياضية والذي يقترح استخدام شبكة التعلم العميق YOLOv5 لاكتشاف وتصنيف الأدوات واستخراج أطر مستطيلة محيطية بالكائنات وفق أصنافها بالإضافة إلى خوارزمية العزل Grabcut للتأكد من عزل الكائن ضمن إطاره المستطيل عن الخلفية حتى يصار نهاية لإيجاد موضع مركز الأداة الفعلي مع اتجاهها على مستوى العمل بالاعتماد على تحليل العناصر الرئيسية. تشمل مزايا هذه الطريقة الدقة المحسنة باستخدام YOLOv5 مع الزمن القصير الذي تتميز به والذي يعطي النظام ميزة العمل ضمن الزمن الحقيقي، بالإضافة إلى تقليل التوقعات الخاطئة مما يحسن موثوقية النظام. بالإضافة إلى ذلك، يتيح الاستدلال في الزمن الحقيقي الاستجابة في الوقت المناسب أثناء علاجات الأسنان، يسمح التطبيق السريري للنظام المقترح بتحسين إجراءات طبيب الأسنان، وتحسين عمليات غرفة التعقيم، وتعزيز كفاءة تنظيم الأدوات.

باختصار، يقدم هذا العمل حلاً متقدماً ومتخصصاً للكشف عن أدوات طبيب الأسنان وتحديد موضعها، وذلك بالاستفادة من تقنيات YOLOv5، Grabcut و PCA. تشمل فوائد النظام تحسين الدقة وإمكانات الاستدلال في الزمن الحقيقي والتخصيص لأدوات طبيب الأسنان، مما يميزه عن الأساليب التي تمت مراجعتها مسبقاً ويقدم تطورات في اكتشاف أدوات طبيب الأسنان وتحديد موضعها.

طرائق البحث ومواده:

طُبِقَ سيناريو العمل على حاسب Macbook-pro mid 2012 باستعمال نظام تشغيل Linux Ubuntu 2022 وتم استعمال كاميرا Promate Pro-Cam2 Premium بدقة 1080 pixel وقدرة النقاط 30 إطار في الثانية لالتقاط البيانات بدقة عالية وبمعدل إطارات سلس. وأتاح نظام Linux مفتوح المصدر تنفيذ السيناريو بكفاءة وسلاسة مع تخصيص البيئة وتكوينها بالاعتماد على لغة البرمجة بايثون.



الشكل (1): الأدوات المستخدمة مع الوسوم والإطار المحيط ونسبة

الموثوقية الخاصة بها

لتشكيل قاعدة البيانات تم استخدام 3 أصناف أدوات أسنان مختلفة تم تسميتها كالتالي Object1, Object2, و Object3 موضحة في الشكل (1). وتم الاستعانة بالبيئة التطويرية السحابية Google Colaboratory لإتمام عملية التدريب الكلية للبيانات بالاستعانة بميزة وجود خيار المعالجة باستخدام وحدات معالجة الرسومات. أما بالنسبة لوسم الصور فقد

استخدم الموقع [7] والذي يتيح تسمية الأصناف وتحديد شكل الإطار الخاص بالأدوات. تتكامل هذه الأدوات سويةً لإتمام عملية التدريب الكلية وتشغيل النظام كاملاً ضمن الزمن الحقيقي. سنقوم بدايةً بتشكيل قاعدة البيانات والتأكد من ملائمتها ثم الانتقال لتدريب الشبكة لاستحصال الأوزان وتطبيق عملية عزل الخلفية مع الاعتماد على التحليل الرياضي لتحديد موضع الأداة وتوجهها ومن ثم إظهار النتائج.

1. تحصيل البيانات:

تُعد عملية تحصيل البيانات الصحيحة أساساً في عملية التدريب الصحيحة، ولذلك نبدأ من طبيعة الصور المراد التعامل معها. بدايةً، نأخذ بعين الاعتبار تموضع أدوات طبيب الأسنان المراد التقاطها، حيث تتوضع على طاولة عمل بشكل عشوائي من حيث زاوية توجهها وحجمها. بالإضافة إلى ذلك، نأخذ في الاعتبار تنوع ظروف الإضاءة وشكل قاعدة البيانات، بحيث تتضمن صوراً مفردة للأدوات، وصوراً مزدوجة تحتوي على قطعتين مختلفتين، وصوراً بثلاث قطع متنوعة. كذلك، ننتبه لتغير الخلفية بين خلفية داكنة وخلفية فاتحة.

تضمنت عملية اختيار مجموعة البيانات الأنسب للتدريب نموذج خطوات تكرارية لضمان الأداء الأمثل في الكشف عن الأدوات. في البداية، تم جمع مجموعة بيانات مكونة من 100 صورة لأدوات طبيب الأسنان، وتم تقسيمها إلى 80 صورة للتدريب و20 صورة للتحقق من الأداء. ليتم فيما بعد اختبار أوزان الشبكة على منصة بث الفيديو لمشهد طاولة العمل التي يتوضع عليها أدوات طبيب الأسنان. أظهرت النتيجة أداءً غير موثوق به يقدر بنسبة متوسط معيار دقة وسطي (mAP) mean Average Precision 55%، مما دفعنا إلى زيادة عدد العينات المأخوذة للتدريب. تم زيادة عدد العينات تدريجياً لتكون كالتالي 640، 800، 1000 لنستقر في النهاية عند عدد عينات 1620 صورة مقسمة إلى 1296 صورة تدريب و324 صورة للتحقق من الأداء بنسبة 80% إلى 20% بالترتيب حيث أتاحت مجموعة البيانات الشاملة هذه للنموذج تحقيق تقدم ملحوظ في اكتشاف الكائنات في سيناريوهات الزمن الحقيقي، مما أدى إلى تعزيز دقة نظام التعرف وموثوقيته بشكل كبير. أعطت عملية التدريب الأخيرة قيمة 91% لمتوسط معيار الدقة الوسطي عند 100 دورة تدريب.

2. وسم البيانات:

تُعد عملية الوسم الدقيقة خطوة أساسية تسبق عملية التدريب عند تشكيل قاعدة بيانات يدوية، وذلك للتأكد من تعرف النظام على الأدوات المدربة بنفس نسق خرج النظام المرغوب. في هذه الدراسة، تم استخدام عملية وسم لتسمية مجموعة صور قاعدة البيانات بأطر محيطية مستطيلة الشكل، يتم تحديدها بواسطة إحداثيات المركز والارتفاع والعرض. تضمنت عملية الوسم خطوتين أساسيتين:

• الإحاطة:

تتم عملية الإحاطة بدايةً باختيار شكل الإطار المحيط. تأخذ شبكة YOLOV5 أطرًا مستطيلة الشكل، وبالتالي يتم إحاطة الكائنات ضمن الصور باستخدام أطر مستطيلة محددة من خلال موقع مركز الإطار وأبعاده من طول وعرض. تحدد هذه الإحداثيات موقع وحجم الأدوات الطبية الموجودة في الصور بدقة، ليتم الاستفادة منها في تدريب الشبكة على توقع المواقع والأحجام لصور جديدة مدخلة لنظام التعرف.

• التسمية:

سميت الأدوات الثلاث بتسميات محايدة لتأخذ الأسماء Object1, Object2, Object3 بحيث تقابل أرقام أصناف 0، 1، و 2 بالترتيب. هذه الأرقام ترتبط بعملية التدريب بحيث يتم ربط المميزات الخاصة بالإطار المحدد مع رقم الصنف لهذه المميزات، مما يعطي تمييزاً بين الأصناف الثلاث في حالتنا. بعد عملية الوسم، تنتقل إلى عملية التدريب لاستحصال الأوزان التي سيتم تشغيل النظام عليها في النهاية.

3. تدريب الشبكة:

بعد إتمام عملية وسم البيانات، استخدمت البيانات الموسومة لتدريب نموذج الشبكة العصبية YOLOv5. حيث تم اختيار YOLOv5 نظراً لانتمائها إلى فئة كاشفات الكائنات ذات المرحلة الواحدة (single-stage detectors)، والتي تتميز بكفاءتها وسرعة تنفيذها، مما يجعلها ملائمة لعمليات الكشف عن الأدوات في البيئة الطبيعية [8]. تبدأ عملية التدريب بتقسيم قاعدة البيانات إلى مجموعتين: مجموعة التدريب ومجموعة الاختبار. تُستخدم مجموعة التدريب لتعليم النموذج، في حين تُستخدم مجموعة الاختبار لتقييم أداء النموذج بعد التدريب.

تم بناء النموذج باستخدام بنية YOLOv5، التي تتألف من عدة طبقات تلافيفية (convolutional layers) وطبقات تجميع (pooling layers)، بالإضافة إلى طبقات كاملة الاتصال (fully connected layers). كما تستخدم الشبكة تقنيات التحويل (transformations) لزيادة دقة التعرف على الكائنات في الصور.

قبل بدء عملية التدريب يتم إعادة تشكيل حجم الصور مجدداً من خلال تحويل حجم الصور إلى حجم متساوي الأبعاد بالنسبة للطول والعرض 640x640 بكسل لتكون ملائمة لشبكة YOLOv5 بالإضافة لتحسين إمكانية القيام بعملية زيادة حجم قاعدة البيانات من خلال ميزة نموذج YOLOv5 الذي يقوم على زيادة البيانات Data augmentation من خلال تشكيل صور جديدة مطبق عليها بعض العمليات كالانتقالات والقلب والتدوير.

تم إعداد النموذج بمعايير محددة تشمل عدد تكرارات دورة التدريب (epochs)، وحجم نافذة العينات ضمن الدورة الواحدة (batch size)، ومعدل التعلم (learning rate) في هذه الدراسة، ثم تحديد عدد تكرارات دورة التدريب بـ 100، وحجم نافذة العينات ضمن الدورة الواحدة بـ 32، ومعدل التعلم بـ 0.001. يتم تدريب النموذج والاعتماد على تقييم متوسط معيار الدقة الوسطي القائم على التقاطع وفق الاتحاد (Intersection over Union - IoU) لتقييم دقة توقع الأطر المحيطة بالنسبة للإطار الفعلي المأخوذ من بيانات التحقق.

أثناء عملية التدريب، تُستخدم البيانات الموسومة كدخل صور وأطر محيطة (bounding boxes) للنموذج. يتعلم النموذج كيفية تحديد موقع وحجم الأدوات الطبية في الصور بناءً على الإحداثيات والأبعاد المحددة في عملية الوسم. يتم تقييم أداء النموذج باستخدام مجموعة الاختبار، وتُعدّل الأوزان والمعايير لتحقيق أفضل أداء ممكن. في نهاية عملية التدريب، تُستخرج الأوزان النهائية للنموذج. تُستخدم هذه الأوزان عند تشغيل النظام على صور جديدة لتوقع مواقع وأحجام الأدوات الطبية بدقة وكفاءة عالية.

4. قياس أداء الشبكة:

لقياس أداء الشبكة، استخدمت مقاييس التعلم العميق التي توضح مدى فعالية عمل النموذج. يتمثل الهدف الرئيسي في تحقيق دقة عالية في التصنيف. لذلك، يتعين البدء بتحليل مصفوفة الارتباك Confusion Matrix، والتي تقدم لمحة عن أداء الشبكة بعد عملية التدريب.

من خلال مصفوفة الارتباك، يُفحص مدى دقة التصنيفات الصحيحة والخاطئة، وهو أمر يساعدنا في فهم أداء النموذج بشكل أفضل وتحسينه عند الضرورة.

A. Confusion Matrix

هي مصفوفة تستخدم في مجال التعلم العميق لوصف أداء نموذج التصنيف على مجموعة من بيانات الاختبار المعروفة القيم الحقيقية لها. يسمح بتصوير أداء الخوارزمية من خلال مقارنة مسميات الأصناف الفعلية (actual class labels) مع مسميات الأصناف المتوقعة (predicted class labels). فهي توصف كما يلي:

- تمثل أسطر المصفوفة الأصناف المتوقعة، بينما تمثل الأعمدة الأصناف الفعلية.
- تمثل الأرقام ضمن المصفوفة نسبة الكائنات الفعلية التي تم اكتشافها بالنسبة للمتوقعة، وعلى سبيل المثال في الشكل (2) يمثل السطر الثاني العمود الأول كم نسبة الصور التي هي Class 1 فعلياً وتم تنبؤها ك Class 2. يمكن استخدام هذه القيم لحساب معايير أداء أخرى مثل الدقة والانتقاء والحساسية والتحليل الشامل للأداء (F1 Score) ومعدل الدقة المتوسط (mAP)، والتي تعطي فكرة أفضل عن أداء النموذج.

		True Class		
		Class 1	Class 2	Class 3
Predicted Class	Class 1	TP	FP	FP
	Class 2	FN	TN	TN
	Class 3	FN	TN	TN

الشكل (2) مصفوفة Confusion Matrix

من خلال الشكل (2) نلاحظ أن هناك عدة أرقام يجب أخذها بعين الاعتبار لحساب معايير التقييم الأخرى، وتعرف هذه الأرقام بالمسميات TP, FP, FN, TN وهي أنماط التنبؤ الأربعة المحصورة بين 0 و1 والتي سنعتمد عليها ضمن بقية الحسابات.

في الشكل (2) تم الأخذ بعين الاعتبار الدراسة بالنسبة إلى Class 1 كالتالي:

- True Positive (TP): تمثل الحالات التي تم التنبؤ بها بشكل صحيح، حيث يتوافق التصنيف المتوقع مع التصنيف الحقيقي. حيث يمكننا أن نجد قيم TP بالنسبة لكل صنف على القطر الرئيسي للمصفوفة. والتي تعرف بالتناحية (True Class == Predicted Class) كلما كانت نسبة TP أقرب إلى الواحد فإن أداء الشبكة يكون أفضل.
- False Positive (FP): تمثل الحالات التي تم التنبؤ بها بشكل غير صحيح، حيث يختلف التصنيف المتوقع عن التصنيف الحقيقي. تعبر هذه النسبة عن التوقع الخاطئ لصنف معين، في حالتنا ندرس Class 1 فمثلاً إذا اعتبرنا أن الكائن المتنبئ به صنف Class 1 فإن حالة FP له هي أن يكون في الحقيقة Class 2 or Class 3.
- False Negative (FN): تمثل الحالات التي لم يتم التنبؤ بها بشكل صحيح، حيث يختلف التصنيف الحقيقي عن التصنيف المتوقع. تعبر هذه النسبة عن التوقع الخاطئ لصنف معين، في حالتنا إذا اعتبرنا أن الكائن الحقيقي هو صنف Class 1 فإن حالة FN له هي أن يتنبأ به Class 2 or Class 3.

• True Negative (TN) : تمثل الحالات التي تم التنبؤ بها بشكل صحيح كتصنيف سلبي، حيث يتوافق التصنيف المتوقع مع التصنيف الحقيقي للعناصر السلبية. تعبر هذه النسبة عن التوقع الخاطئ لصف معين حيث يكون الكائن الحقيقي والمتوقع مخالفين لحالة Class 1. من خلال تحليل هذه الأنماط، يمكننا تقييم أداء النموذج وفهم الأخطاء التي يرتكبها، وبالتالي تحسينه للحصول على نتائج أفضل.

B. Precision

درجة الدقة هي مقياس تصنيف يقيس قدرة المصنف على عدم تصنيف عينة سلبية على أنها إيجابية. بشكل آخر، يمكن وصف درجة الدقة على أنها نسبة التنبؤات الإيجابية التي تكون فعلاً إيجابية إلى مجموع التنبؤات الإيجابية. يتم التعبير عنها بالعلاقة

$$[1] \text{ دقة العينات الإيجابية} = \frac{\text{عدد التوقعات الصحيحة الإيجابية}}{\text{عدد التوقعات الإيجابية الكلية}}$$

لحساب الدقة (Precision) والتعامل مع كل الأصناف بطريقة متساوية، يتم استخدام ما يُعرف بنسبة الدقة الشاملة (Macro Precision). يتم ذلك عن طريق حساب الدقة لكل صنف على حدة ثم احتساب المتوسط الحسابي للقيم الناتجة. يهدف ذلك إلى ضمان حساب دقة النموذج بشكل متوازن بين جميع الأصناف، دون تحيز. بعد حساب Precision لكل فئة وفق صنفها، يمكننا الحصول على قيمة Precision النهائية من خلال حساب المتوسط الحسابي لجميع القيم المحسوبة للأصناف.

باستخدام مقياس الدقة، يمكننا تقدير كفاءة النموذج في تصنيف العينات بشكل صحيح، مما يساهم في تحسين أداء النظام بشكل عام. تحديد درجة الدقة ونسبة الدقة الشاملة Macro Precision يعطينا فهماً شاملاً لأداء النموذج على جميع الأصناف المختلفة، مما يمكننا من اتخاذ القرارات الأمثل لتحسين النظام وتحسين تصنيف العينات.

C. Recall

الحساسية اتجاه الصنف والتي يتم حسابها من خلال أخذ عمود قيم صنف معين والتي تعبر عن التوقعات الفعلية للصنف ثم تنسب القيم التي تم توقعها بشكل صحيح TP إلى كل قيم الصنف الفعلية (المتوقعة لتكون الصنف نفسه TP أو المتوقعة كصنف آخر FN) وتكون العلاقة المعبرة عن حساسية الصنف كالتالي:

$$[2] \text{ الحساسية} = \frac{\text{عدد التوقعات الصحيحة الإيجابية}}{\text{عدد التوقعات الإيجابية الصحيحة} + \text{عدد التوقعات السلبية الخاطئة}}$$

كما تمت عملية حساب Precision فإن Recall سيتم حسابه بطريقة Macro حيث سيتم حساب Recall كل صنف على حدة ثم حساب Recall النظام كله من خلال أخذ المتوسط الحسابي لكل الأصناف.

D. F1 Score

النقطة الختامية التي تمثل معيار F1 Score تقدم مقياساً شاملاً لأداء النموذج عن طريق جمع معلومات الدقة والحساسية معاً. يقدم هذا المقياس فكرة عامة عن كفاءة النموذج في التعرف على العناصر الإيجابية وتفاذي التصنيف الخاطئ للعناصر السلبية.

على الرغم من أن الدقة والحساسية مهمتان بمفردهما، إلا أن مقياس F1 Score يتيح لنا فهم أداء النموذج بشكل أفضل، خاصة في الحالات التي تكون فيها الفئات متفاوتة الحجم أو غير متوازنة.

العلاقة بين Precision و Recall يمكن تمثيلها بشكل منحني، حيث يتناسب الأداء الجيد للنموذج مع زيادة المساحة تحت المنحني. (Area Under Curve - AUC) كلما زادت قيمة AUC ، كان الأداء والتوازن للنموذج أفضل.

E. Mean Average Precision (mAP)

تستخدم للتعبير عن قيمة Precision و Recall من خلال قيمة واحدة ويتم حسابها بالخطوات التالية:

- المرور على كل ثنائيات Precision/Recall.
- حساب الفرق بين القيمة الحالية واللاحقة من أجل Recall (الوزن).
- ضرب الوزن بقيمة Precision الحالية.
- تكرار آخر خطوة من أجل كل ثنائيات P-R.
- استخلاص كل Scores.

بعد حساب (AP) Average Precision لكل صنف، يتم تقييم النظام ككل عن طريق حساب المتوسط الحسابي لـ AP لكل الفئات المعرفة. هذا يتم عن طريق جمع قيم Average Precision لكل صنف ثم تقسيم النتيجة على عدد الأصناف المعرفة.

5. عمليات ما بعد الشبكة العصبونية:

يتم في هذه المرحلة الاستفادة من خرج إطار الصورة المعالجة بواسطة الشبكة العصبونية، والتي تحتوي على أطر محيطية للكائنات معرفة بمركزها وأبعادها ورقم صنف الأداة التي تحيطها. تحديد مركز الأداة بدقة يختلف عن مركز الإطار المحيط بها، وبالتالي يجب عزل الأداة على مستوى البيكسل ضمن الإطار حتى يصار إلى تحديد مركز الأداة الدقيق. نقوم بتصنيع قناع يتم فيه عزل الكائنات عن الخلفية.

تعد خوارزمية GrabCut من أفضل الطرق لعزل الخلفية مع تغيير خلفيات الكائنات. تعتمد هذه الخوارزمية على مجموعة عمليات تكرارية يتم فيها إيجاد عقدتين، واحدة لتجميع بيانات الخلفية والأخرى لتجميع بيانات الكائنات، ثم يتم تنمية كل بكسل حسب مدى قربه من إحدى العقدتين ليتم فصلهما عن بعضهما [10]. هذه العملية تعد من العمليات التكرارية الجيدة في عملية العزل حتى مع اختلاف بنية الكائنات عن الخلفية وتقارب الألوان، لكنها تعتبر مستهلكة للوقت. لذلك، تم اللجوء إليها على مستوى الكائن في الصورة بعد عملية اقتطاع الأداة المرغوبة.

تبدأ الخوارزمية بتحديد مستطيل خرج شبكة التعلم العميق حول الكائن المراد فصله عن الخلفية في الصورة. يُستخدم نموذج التوزيع الغاوسي المختلط (Gaussian Mixture Model - GMM) لتوزيع القيم اللونية لكل من الكائنات الأمامية والخلفية في الصور. يتم تحويل الصورة إلى مخطط بياني، حيث تمثل البكسلات كعقد وتحدد الأوزان بناءً على احتمالات الانتماء للأمامية أو الخلفية. تُطبق خوارزمية Min-Cut/Max-Flow لفصل البكسلات إلى مجموعات أمامية وخلفية، ويتم تحديث نموذج GMM بشكل تكراري لتحسين الفصل. في النهاية، تُستخدم الأوزان المستقرة لاستخراج الكائن من الصورة بدقة.

6. استحصال قياسات الأداة المرغوبة:

بعد استخراج قناع الكائن، تتم عملية أخذ القياسات له والمحددة وفق مساحة العمل بمركز الكائن وتوجهه على طاولة العمل. استخدمت خوارزمية تحليل العناصر الأساسية (Principal Component Analysis - PCA) لاستخلاص مركز إحداثيات الأداة المحددة بالقناع الناتج عن المرحلة السابقة، بالإضافة إلى إيجاد المتجهات الذاتية بعد تحليل تغاير البيانات، وأخذ المتجه الذاتي المقابل لأكبر قيمة ذاتية، الذي يحدد الاتجاه الرئيسي للكائن [11].

بالإضافة إلى ذلك، يجب الأخذ بعين الاعتبار أن الخرج الناتج عن عملية عزل الخلفية عن الكائنات الأمامية المستخرجة باستخدام GrabCut قد يحتوي على أطراف أدوات أخرى عند تقارب الأدوات من بعضها وتداخل أطرافها المستطيلة مع بعضها. هذا الأمر يقودنا إلى استخدام تقنية انتقاء صحيحة للكائن ضمن الإطار المقطع. لإتمام عملية الانتقاء، تم استخدام حساب المساحة الأكبر ضمن الإطار المستطيل، والتي تعبر عن الأداة المرغوبة ضمن إطارها، ثم تقليل الأبعاد لإتمام عملية إيجاد الحسابات عن طريق إيجاد محيط الكائن ذو المساحة الأعظمية ليتم حساب مركزه وتوجهه بدقة.

عملية تحليل المكونات الرئيسية تتضمن عدة خطوات رياضية تهدف إلى تحليل البيانات بشكل أفضل وتقليل الأبعاد كالتالي:

1. **توسيط البيانات بإزالة المركز إلى منتصف توزع البيانات:** يتم هذا عن طريق حساب متوسط كل ميزة (متغير) في البيانات وطرح قيمة المتوسط من كل قيمة في الميزة. مما يساعد في إزالة المركز إلى الوسط الهندسي لتوزع البيانات.
2. **إيجاد التوجه العام للبيانات:** تحسب مصفوفة التباين (التشتت) للبيانات. هذه المصفوفة تمثل كيفية تباين البيانات في كل اتجاه لإيجاد التوجهات الرئيسية للتغير في البيانات.
3. **حساب القيم والمتجهات الذاتية (Eigenvalues and Eigenvectors) لمصفوفة التباين:** تُعبر القيم الذاتية عن كمية التباين في كل اتجاه، بينما تُعبر المتجهات الذاتية عن الاتجاهات الرئيسية للتباين.
4. **اختيار المكونات الرئيسية:** يتم اختيار عدد من المتجهات الذاتية (المكونات الرئيسية) التي تحتوي على معظم التباين في البيانات عملية تحليل العناصر الرئيسية.
5. **تقليل الأبعاد:** يتم تقليل أبعاد البيانات عن طريق أخذ العينات الموافقة للمكونات الرئيسية المختارة بحيث تُمثل كل نقطة بيانات في الفضاء الجديد توازنًا بين الحفاظ على معظم المعلومات وتقليل الأبعاد. تُساعد هذه الخطوة في تبسيط التحليل وتحسين أداء النماذج التنبؤية والتصنيف.

أولاً، يتم توحيد البيانات من خلال توسيطها حول المبدأ. يتم تحقيق ذلك عن طريق طرح المتجه المتوسط Mean Vector من كل نقطة بيانات. بعد ذلك، يتم حساب مصفوفة التباين Covariance Matrix لتمثيل العلاقات بين المميزات المختلفة في مجموعة البيانات. يتم حساب التباين بين المميزات i و j بواسطة المعادلة:

$$Cov(i, j) = \frac{\sum((X_c[:,i] - \mu[i]) * (X_c[:,j] - \mu[j]))}{(n-1)} \quad 3$$

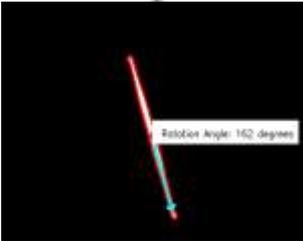
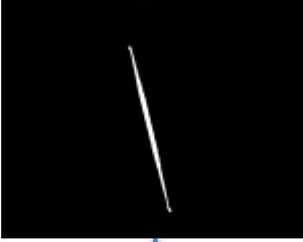
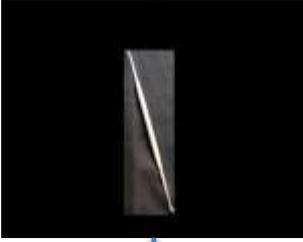
- $Cov(i, j)$ يمثل التباين بين البكسلات i, j
- $X_c[:, i] - \mu[i]$ تقوم على توسيط البكسل i
- $\mu[i]$ يمثل وسيط المتغير i . حيث، يعد طرح المتوسط من كل نقطة بيانات جزءاً من عملية توسيط البيانات.
- $X_c[:, j] - \mu[j]$ تقوم على توسيط البكسل j
- n تمثل عدد نقاط البيانات

تعطينا معادلة التباين مقياساً لكيفية تغير هذه المتغيرات معاً، مع مراعاة المتوسط وحجم مجموعة البيانات. وبعد حساب مصفوفة التباين يتم إيجاد القيم والمتجهات الذاتية لهذه المصفوفة. حيث تحدد القيم الذاتية مقدار التباين الموجود في كل اتجاه، والمتجهات الذاتية تحدد اتجاه هذا التباين.

بعد إيجاد القيم والمتجهات الذاتية يتم ترتيبهم بعدد القيم الذاتية التي حصلنا عليها. ثم يأخذ المتجه الذاتي الذي يحتوي على أكبر قيمة ذاتية والمرتبطة باتجاه أكبر تغير للبيانات، حيث يعد المتجه المرتبط بالقيمة التالية لأكثر قيمة ذاتية مرتبطة باتجاه ثاني أكبر تغير للبيانات.

ونهاية لتقليل الأبعاد يتم أخذ المتجهات الذاتية المرتبطة مع أكبر ثلاث قيم ذاتية لتحديد اتجاه أداة الطبيب المرغوبة.

النتائج والمناقشة:

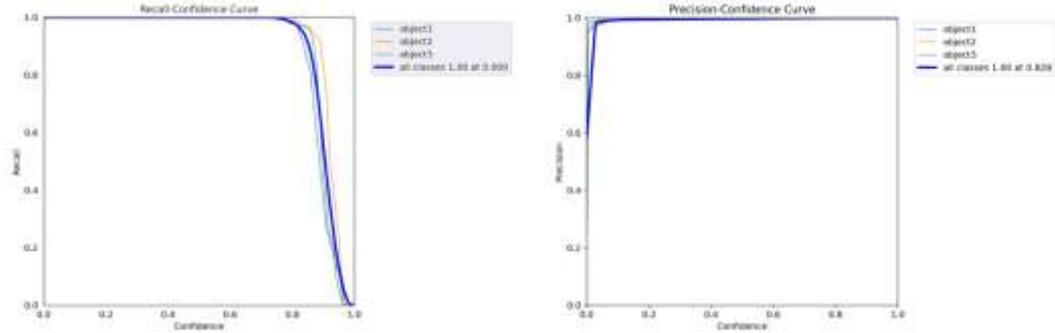


تم بداية عملية استحصال البيانات المشكلة من 1620 صورة قسمت بنسبة 80% تدريب و20% للتحقق حيث أخذت الصورة باعتباريات متنوعة منها توجه الأدوات على طاولة العمل وأحجامها المختلفة مع اختلاف الإضاءة الموجهة على طاولة العمل. قيست الإضاءة باستعمال حساس قياس شدة إضاءة Luxmeter حيث تراوحت القيمة الكلية للإضاءة بين 439-519 Lux. تم ربط كاميرا أحادية العدسة أعلى مشهد طاولة العمل وربطت مع حاسب Macbook pro Mid 2012 حيث تمت عمليات تشغيل نظام الرؤية.

يوضح الشكل (3) خطوات الخوارزمية المقترحة الكلية. بداية يتم التقاط إطار فيديو ومعالجته حيث تزال التشوهات الحاصلة للكاميرا باستخدام مصفوفة التصحيح بمدخلات تحتوي بارامترات تصحيح تشوه الكاميرا مع مصفوفة الكاميرا ليستحصل منها على خرج مصفوفة إطار الفيديو الغير مشوه ثم يمرر الإطار كمدخل لنموذج شبكة التعلم العميق ليتم استخلاص مستطيلات محيطة للكائنات المكتشفة ضمن الإطار محددة بمركز الإطار وأبعاده بالإضافة إلى درجة موثوقية صنف الأداة. في هذه المرحلة نختار الأداة المرغوبة من خلال اختيار رقم الصنف الموافق لها وفق طلب الطبيب. يتم أخذ المستطيل الموافق واقتصاصه من إطار الفيديو ليتم بعدها تطبيق عملية عزل الكائنات الأمامية عن الخلفية باستخدام خوارزمية Grabcut. في هذه المرحلة عند تداخل جزء من أداة أخرى مع مستطيل الأداة المرغوبة فإن عملية تطبيق استخراج القياسات مباشرة عليها سيؤدي إلى خطأ لذلك نوجد الكائن الأكبر ضمن مستطيل الأداة المرغوبة ونقوم بأخذ القناع الخاص به وبعدها نقوم بحساب المحيط الخارجي له ليمرر في النهاية إلى مرحلة إيجاد القياسات باستخدام تحليل العناصر الرئيسية حيث نوجد قناع مساحة الكائن الأكبر ضمن الإطار المستطيل للكائن المرغوب ويتم إيجاد أكبر محور تغاير بيانات لمحيط الكائن المعزول ويحسب منه زاوية توجه الكائن ضمن مستوي مساحة العمل.

الشكل (3) إطار تم التقاطه من بث الفيديو يُظهر خطوات الخوارزمية للأداة التي اختارها طبيب الأسنان

أظهرت نتائج عملية التدريب الموضحة في الشكل (4) دقة بقيمة 100% عند عتبة موثوقية لصنف الكائن مساوية إلى



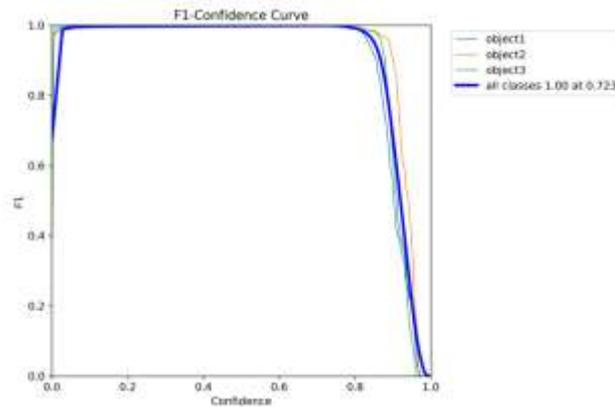
(b) نسبة الحساسية

(b) نسبة الدقة للموثوقية

المهثوقة

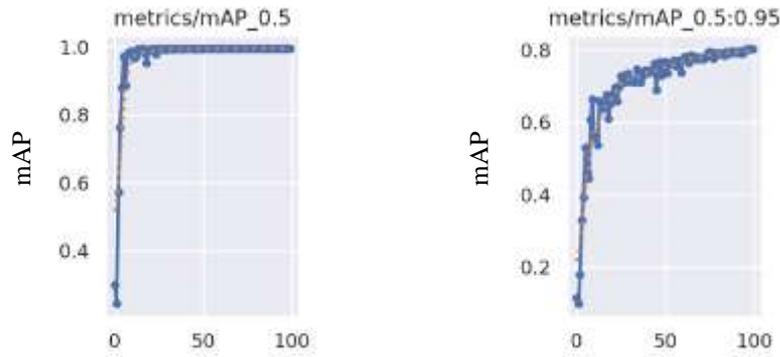
الشكل (4) معياري الدقة والحساسية لكل من الأصناف الثلاث ولكامل الأصناف

82.8% وحساسية اتجاه الأصناف بقيمة 100% حتى درجة موثوقية بقيمة 80%. يعطي منحنى F1score مساحة تحت المنحنى شبه مستطيلة والتي تدل على الأداء الجيد للنموذج من خلال دمج معياري الدقة والحساسية ضمن معيار واحد كما يوضح الشكل (5).



الشكل (5) معيار F1 Score لكل من الأصناف الثلاث ولكامل الأصناف

نهاية يقيم التدريب الكلي بالنسبة لمتوسط معيار الدقة الوسطي لكل من الأصناف الثلاث بقيمة 99.5% والتي تستقر بعد دورة التدريب 53 بالنسبة لقناع 50% من منطقة تقاطع الاتحاد Intersection of Union وتأخذ قيمة 80% بالنسبة لقناع وسطي من 50-95% من منطقة تقاطع الاتحاد عند دورة التدريب 100 كما هو مبين في الشكل (6).



(a) معيار متوسط الدقة المتوسطة عند نسبة تقاطع للاتحاد 50%

(b) معيار متوسط الدقة المتوسطة عند نسبة تقاطع للاتحاد 50-95%

الشكل (6) متوسط معيار الدقة المتوسط خلال دورات التدريب المئة

استخدم عملية التحقق بالتقاطع Cross Validation للتأكد من عدم انحياز الشبكة اتجاه صنف معين من الأدوات حيث قسمت قاعدة البيانات إلى خمس ملفات في كل عملية تدريب يأخذ 4 ملفات منها للتدريب ويبقى الملف الأخير للتحقق. أظهر المتوسط الحسابي قيمة متوسط معيار دقة وسطي لكل الملفات سوية مساوياً 80.6% كما هو موضح في الجدول (1).

جدول (1) قيمة معيار mAP خلال عملية التأكد من عدم الانحياز باستخدام تقنية cross-validation

Average	mAP50-95%					Object
	File 5	File 4	File 3	File 2	File 1	
All	0.803	0.811	0.813	0.8	0.801	All
0.80675	0.772	0.787	0.783	0.774	0.767	Object 1
0.7766	0.853	0.831	0.848	0.835	0.835	Object 2
0.8404	0.784	0.816	0.807	0.79	0.801	Object 3
0.7996						

هذه القيم عكست بتجارب اختبار أعطت زمن التقاط للأدوات خلال دورة تشغيل الخوارزمية 2.58 ثانية لكامل عمليات الاكتشاف والتصنيف مع تحديد المركز والاتجاه. بمقابل زمن اكتشاف وتصنيف فقط يعادل 0.5 ثانية. الأمر الذي يعد مناسباً لإتمام عمليات المناولة لطبيب الأسنان.

في هذه الدراسة، قدمنا نهجاً متكاملًا وفعالاً يدمج خوارزميات YOLOv5 و GrabCut و PCA لتحديد مواضع أدوات طبيب الأسنان على طاولة العمل بدقة عالية، مما يشكل أساساً لعملية المناولة المؤتمتة أثناء العمليات الجراحية. أظهر النموذج المقترح قدرة عالية على اكتشاف وتصنيف وتحديد مواضع الأدوات، حيث حقق دقة متوسطة (mAP50-95) بلغت 80.6%، مع تفوق أداء الكائن 2 على الكائنين 1 و 3.

تساهم خطوات ما بعد المعالجة، التي تشمل العزل باستخدام GrabCut وانتقاء الكائن الأكبر ثم تحليل العناصر الرئيسية، في تحسين دقة تحديد المواضع. يعزز النظام المقترح من كفاءة مساعدي طبيب الأسنان ويقلل من مخاطر الأخطاء الجراحية، مما يحسن من الإجراءات الجراحية ويعزز سلامة المرضى.

يوفر هذا النهج حلاً متخصصاً لإدارة أدوات طبيب الأسنان بفعالية أكبر، مما يساهم في تطوير أنظمة تحديد المواضع في مجال طب الأسنان وتعزيز جودة الرعاية الصحية المقدمة للمرضى. تشير النتائج إلى أن هذا النظام يمكن أن يكون أداة قيمة لتحسين السلامة والكفاءة في البيئات الجراحية، ويفتح آفاقاً جديدة لاستخدام تقنيات الذكاء الاصطناعي في التطبيقات الطبية.

References:

- [1] SHROFF, N. TAGUCHI, Y. TUZEL, O. VEERARAGHAVAN A. RAMALINGAM S. and OKUDA H. "Finding a needle in a specular haystack," in Proc. 2011 IEEE International Conference on Robotics and Automation., 2011, pp. 5963–5970.
- [2] COLLET A., MARTINEZ M., and SRINIVASA S., "The MOPED framework: Object recognition and pose estimation for manipulation," The International Journal of Robotics Research, 2011, pp. 1284–1306.
- [3] RODRIGUES J. J., KIM J. S., FURUKAWA M., XAVIER J., AGUIAR P., and KANADE T., "6D pose estimation of textureless shiny objects using random ferns for bin-picking," 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura-Algarve, Portugal, 2012, pp. 3334-3341.
- [4] XU, Y., MAO, Y., TONG, X., TAN, H., GRIFFIN, W. B., KANNAN, B., and DEROSE, L. "Robotic handling of surgical instruments in a cluttered tray", *IEEE Transactions on Automation Science and Engineering*, 2015, pp. 775–780.
- [5] HASAN, M. K., CALVET, L., RABBANI, N., and BARTOLI, A. "Detection, segmentation, and 3D pose estimation of surgical tools using convolutional neural networks and algebraic geometry", Vol. 70, *Medical Image Analysis*, 2021, 101994.
- [6] MCDONALD, R. I., WALSH, L. J., and SAVAGE, N. W., "Analysis of workplace injuries in a dental school environment", *Australian Dental Journal*, 1997, Vol. 42(2), 109–113.
- [7] "Makesense official website" [Online]. Available: <http://makesense.ai/>. Last date May 2024.
- [8] REDMON, J., DIVVALA, S., GIRSHICK, R., and FARHADI, A., "You only look once: unified, real-time object detection", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, (pp. 779-788).
- [9] MOHANAPRIYA S., MOHANA S., KUMARAVEL T., SUMITHRA P., "Image Detection and Segmentation using YOLO v5 for surveillance", *Applied and Computational Engineering*, 2023, (pp. 160-165).
- [10] WANG Z., LV Y., WU R., ZHANG Y., "Review of GrabCut in Image Processing". *Mathematics*. 2023, Vol. 11(8), pp. 1965.
- [11] MISHRA S., SARKAR U., TARAPHDER S, and Datta S., "Principle Component Analysis", *International Journal of Livestock Research*, 2017, Vol 7(5), (pp. 60-78)