

# Roads Extraction from Satellite Images using Convolution Neural Network Model (Deeplabv3+) A Case Study in Lattakia city

Rose Hwajjeh\*

(Received 2 / 1 / 2023. Accepted 12 / 2 / 2023)

## □ ABSTRACT □

The purpose of this paper is to extract roads from satellite images, based on developing the performance of the deep convolutional neural network model (Deeplabv3+) for roads segmentation, and to evaluate and test the performance of this model after training on our data. This experimental study was applied at Google Colab cloud platform, by software instructions and advanced libraries in the Python. We conducted data pre-processing to prepare ground truth masks, then we trained the model. The training and validation process required (Epochs=4), by (Patch Size=4 images). The Loss function decreased to its minimum value (0.025). Training time was three hours and ten minutes, aided by the advanced Graphics Processing Unit (GPU) and additional RAM. We achieved good results in evaluating the accuracy of the predictions of the trained model (IoU = 0.953). It was tested on two different areas, one of which is residential and the other agricultural in Lattakia city. The results showed that the trained model (DeepLabv3+) in our research can extract the road network accurately and effectively. But its performance is poor in some areas which includes tree shadows on the edges of the road, and where the spectral characteristics are similar to the road, such as the roofs of some buildings, and it is invalid for extracting side and unpaved roads. The research presented several recommendations to improve the performance of the (Deeplabv3+) in extracting roads from high-resolution satellite images, which is useful for updating road maps and urban planning works.

**Keywords:** Roads Extraction, Convolutional neural network, (Deeplabv3+), Semantic segmentation, ground truth masks, Epoch, Patch Size, Loss function, IOU.

---

\*Work Supervisor, Department of Topography- Faculty of Civil Engineering- Tishreen University- Lattakia- Syria. Email: [h.rose80@gmail.com](mailto:h.rose80@gmail.com)

## استخراج الطرق من صور الأقمار الصناعية باستخدام نموذج الشبكة العصبونية الانتقافية (Deeplabv3+): حالة دراسية في مدينة اللاذقية

روز حويجه\*

(تاريخ الإيداع 2 / 1 / 2023. قُبِلَ للنشر في 12 / 2 / 2023)

### □ ملخص □

يهدف هذا البحث إلى استخراج الطرق من صور الأقمار الصناعية، اعتماداً على تطوير أداء نموذج الشبكة العصبونية الانتقافية العميقة (Deeplabv3+) لتوظيفه في مهمة التقسيم الدلالي للطرق، وتقييم واختبار فعالية أداء هذا النموذج بالتنبؤ بالطرق بعد تدريبه على بيانات بحثنا.

تم تطبيق هذه الدراسة التجريبية على منصة غوغل كولااب السحابية، من خلال تنفيذ التعليمات البرمجية والمكتبات المتطورة في لغة البايثون، قمنا بإجراء المعالجة المسبقة لبيانات بحثنا لإعداد صور الحقيقة الأرضية، ثم تدريب النموذج، تطلبت عملية التدريب والتحقق (4) فترات، بحجم دفعة للبيانات (4) صور، تناقصت دالة الخسارة إلى حدودها الدنيا بقيمة (0.025)، واستغرق زمن التدريب ثلاث ساعات وعشرة دقائق، وذلك بالاستعانة بوحدة معالجة الرسومات المتطورة (GPU) وذاكرة وصول عشوائي إضافية. حققنا نتائج جيدة في تقييم الدقة في صحة تنبؤات النموذج المدرب بقيمة (التقاطع إلى الاتحاد=0.953)، فتم اختباره على منطقتين مختلفتين إحداها سكنية وأخرى زراعية في مدينة اللاذقية، أظهرت النتائج أن النموذج (DeepLabv3+) المدرب في بحثنا يمكنه استخراج شبكة الطرق بدقة وفعالية، لكن أدائه ضعيف في بعض المناطق التي تحوي أشجار بسبب تأثير الظلال على حواف الطرق، وحيث تكون الخصائص الطيفية مشابهة للطرق كأسطح بعض المباني، وهو غير صالح لاستخراج الطرق الفرعية وغير المعبدة. قدم البحث عدة توصيات بتحسين أداء النموذج (Deeplabv3+) في استخراج الطرق من صور الأقمار الصناعية عالية الدقة، بما يفيد في تحديث خرائط الطرق وأعمال التخطيط الحضري.

**الكلمات المفتاحية:** استخراج الطرق، الشبكة العصبونية الانتقافية، (Deeplabv3+)، التقسيم الدلالي، صور الحقيقة الأرضية، الفترة، حجم الدفعة، دالة الخسارة، التقاطع إلى الاتحاد.

\* مشرف على الأعمال - قسم الهندسة الطبوغرافية - كلية الهندسة المدنية - جامعة تشرين - اللاذقية - سورية.

بريد الكتروني: [h.rose80@gmail.com](mailto:h.rose80@gmail.com)

**مقدمة:**

تُعدُّ الشبكة العصبونية الالتفافية (Convolutional Neural Network: CNN) أحد أنواع التعلّم العميق (Deep Learning)، وتُستخدم شبكات CNNs عادةً في تطبيقات الرؤية الحاسوبية (Computer vision)؛ فقد تم إثبات فعاليتها على مدار السنوات العديدة الماضية، في اكتشاف وتمييز واستخراج الكائنات والأنماط من الصور ومقاطع الفيديو [1]، ومع ذلك، فإن تطبيقها على صور الأقمار الصناعية أقل تطوراً لأن نماذج الشبكات العصبونية الالتفافية مصممة في الأصل للصور البسيطة وليس لصور الأقمار الصناعية [2].

تحتاج شبكات (CNNs) إلى أن تكون عميقة بدرجة كبيرة لتتعلم تمثيلات تفصيلية كافية لكل صنف من الأصناف في الصورة، بحيث يمكنها التمييز بينها. وتعتبر الشبكة العصبونية الالتفافية بالكامل (fully convolutional network: FCNs) أحد أكثر أنواع بنى الشبكات العصبونية العميقة شيوعاً لتلبية هذه المطالب، والتي يُعدُّ ظهورها أساساً للعمل في التقسيم الدلالي للصور (image semantic segmentation) الذي يتضمن اكتشاف الكائنات (object detection) ثم تطبيق التصنيف (classification)، ليتم العمل على مستوى البكسل لتحديد معالم الكائنات داخل الصورة [3].

تم تطوير عدة نماذج من الشبكات العصبونية الالتفافية العميقة لمهمة التقسيم الدلالي للصور، منها يعتمد استخدام وحدة تجميع الهرم المكاني (Spatial pyramid pooling module)، وشبكات أخرى تعتمد بنية وحدة التشفير وفك التشفير (Encode-decoder structure)، في حين حقق النموذج المُطوّر حديثاً (DeepLabv3+) أداءً مذهلاً ونتائجاً واعدة في التقسيم الدلالي، حيث تم في هذا النموذج تطبيق طريقة التفاف أتروس (Atrous Convolution) الذي يستخدم وحدة تجميع الهرم المكاني مع بنية وحدة فك التشفير، من أجل العمل على تتبع السمات وفصلها على مقاييس متعددة، مما يسمح باستخراج السمات بدقة، ويجعل أداء النموذج أسرع وأقوى [4].

يُعتبر استخراج المعلومات بشكل موثوق من صور الأقمار الصناعية مشكلة صعبة في العديد من التطبيقات العملية. لذلك يتم في الآونة الأخيرة البحث في تطوير أساليب التعلّم العميق بشكل متزايد لاستخراج المعلومات من صور الأقمار الصناعية، نظراً لأدائها الجيد، وقدرتها على التعميم، مقارنة بالطرق التقليدية التي تستغرق وقتاً طويلاً وتحتوي على أخطاء وفيرة يرتكبها العاملون البشريون [5,6,7]. ويعدُّ البحث في مهام استخراج الطرق (Road Extraction) من أهم التحديات للأعمال البحثية في مجال الرؤية الحاسوبية، فقد تم اقتراح مجموعة متنوعة من خوارزميات التعلّم العميق وتدريبها على استخراج الطرق من صور الأقمار الصناعية، ومقارنة أفضلية هذه النماذج مع بعضها من حيث دقتها في الأداء [8,9]. وتمت الإشارة إلى أنه من غير الكافي استخدام معلومات كثافة الصورة لاكتشاف الطرق فقط [10]، ولا تزال هناك حاجة إلى طريقة أكثر عمومية وقوة لتحسين أداء استخراج الطرق [11].

تُقدم في هذا البحث نتائج دراسة تجريبية قُمتُ بإجرائها لاستخراج الطرق من صور الأقمار الصناعية، باستخدام طريقة جديدة تقترح الاعتماد على نموذج الشبكة العصبونية الالتفافية العميقة (deeplabv3+)، وقد قُمتُ بتطوير أدائه لمهمة التقسيم الدلالي للطرق، عبر نهج نقل التعلّم (Transfer the Learning) بناءً على بيانات بحثنا، ومن خلال تنفيذ التعليمات البرمجية والمكتبات المتطورة في لغة البرمجة بايثون (Python)، ثم تقييم فعاليتها واختبار أداء النموذج المدرب في التنبؤ بالطرق.

## أهمية البحث وأهدافه:

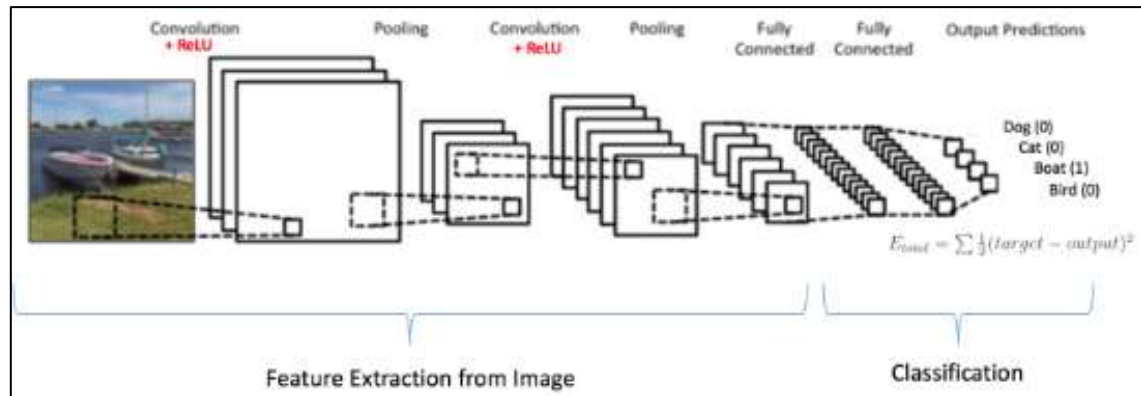
تكمن أهمية البحث في محاولته استخراج الطرق من صور الأقمار الصناعية، بالاعتماد على الشبكة العصبونية الالتفافية العميقة (+deeplabv3) كأحد أهم النماذج المُطَوَّرَة حديثاً في التقسيم الدلالي للتعلم العميق [9,11]، وإعداد النموذج ليصبح أداة يمكن استخدامها لاستخراج الطرق من صور الأقمار الصناعية بدقة وفعالية، مما يسهم في توفير الوقت والجهد اللازم، لدعم اتخاذ القرارات المناسبة بشأن أعمال التخطيط الحضري وإدارة استخدامات الأراضي. إنَّ الهدف الرئيس لهذا البحث هو استخراج الطرق من صور الأقمار الصناعية، اعتماداً على تطوير أداء نموذج الشبكة العصبونية الالتفافية العميقة (+deeplabv3) لتوظيفه في التقسيم الدلالي للطرق، وتقييم واختبار فعالية أداء هذا النموذج بالتنبؤ بالطرق بعد تدريبه على بيانات بحثنا.

## طرائق البحث ومواده:

بداية سنستعرض بنية نموذج الشبكة العصبونية الالتفافية العميقة (+DeepLabv3)، وآلية عملها، بالإضافة إلى البيئة البرمجية والمكتبات المتطورة في لغة البرمجة بايثون، المقترحة لإنجاز هذا البحث:

### بنية نموذج الشبكة العصبونية الالتفافية (+DeepLabv3):

تتكون الشبكات العصبونية الالتفافية من ثلاث أنواع أساسية من الطبقات، هي: طبقة التلافيف (Convolutional Layer)، وطبقة التجميع (Pooling Layer)، وطبقة الارتباط الكامل (Fully-connected Layer). تقوم الطبقتان الأوليتان وهما طبقات التلافيف والتجميع باستخراج المعالم، بينما تقوم الطبقة الثالثة بتعيين وتصنيف السمات المستخرجة في المخرجات النهائية [1,3]. كما يوضح الشكل (1)



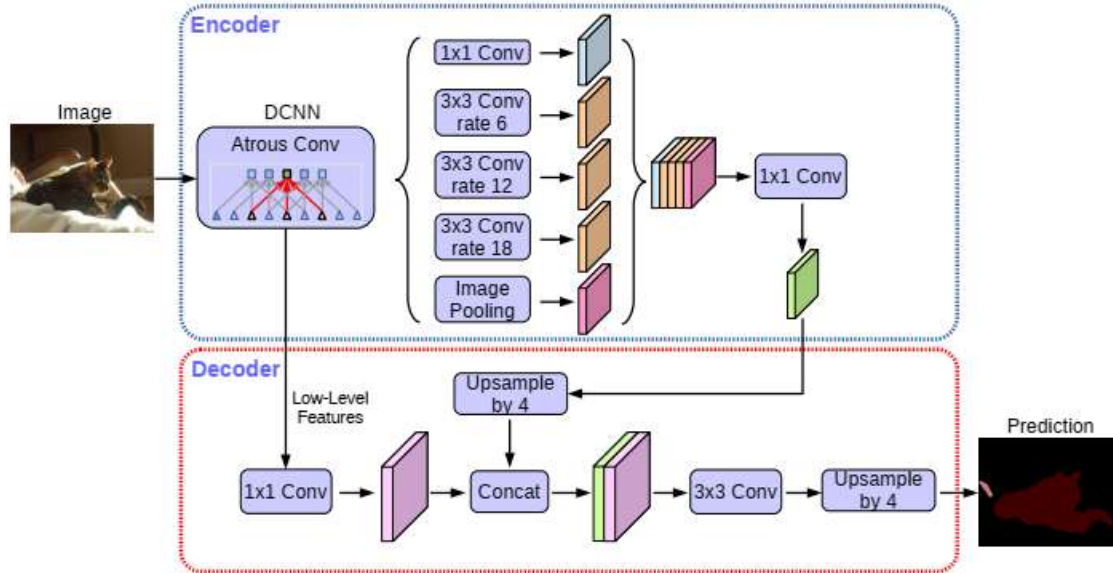
الشكل (1). الطبقات الأساسية المكونة للشبكات العصبونية الالتفافية [3].

تعمل الشبكات العصبونية الالتفافية عن طريق تحريك مرشحات (Filters) على صورة الدخل، كل مرشح له أوزان (weights) مختلفة، بغية استخراج سمات مختلفة، يُمرَّر المرشح على خلايا الصورة كاملةً، وتُطبَّق مجموعة من العمليات الحسابية والتحويلات غير الخطية، بحيث يتناقص حجم الفلاتر عادةً على مدار الشبكة، مع زيادة عددها أيضاً من أجل التعرف على السمات الدقيقة. يُعاد ضبط قيم المرشحات والأوزان خلال عملية التدريب الدورية، يتم تدريب الشبكة لعدد مُعيَّن من التكرارات (Epochs) حتى الوصول إلى الأوزان المثلى التي تجعل تابع الخطأ أقل ما يمكن، لتحقيق التنبؤ الصحيح [3].

أما عن النموذج (DeepLabv3+) المقترح في بحثنا لاستخراج الطرق من صور الأقمار الصناعية، فهو شبكة عصبونية إلتفافية عميقة لمهمة التقسيم الدلالي، مُطوّرة من خلال تطبيق التفاف أتروس (Atrous Convolution) الذي يعمل على تتبع السمات وفصلها على مقاييس متعددة، مما يسمح باستخراج السمات بدقة، ويجعل أداء النموذج أسرع وأقوى. يُبين الشكل (2) مبدأ عمل هذا النموذج المقترح، حيث يستخدم (DeepLabv3+) مشفراً هجيناً مقسماً إلى جزأين [4]، هما:

الجزء الأول (Encoder): يعتمد عمل نموذج الشبكة العصبونية العميقة (auto encoder) التي تستنتج التوزيع الاحتمالي لبيانات الدخل، فمن خلالها يتم استنتاج كثافة البيانات الموجودة في الصورة، وبالتالي نتائجها عبارة عن غمامة كثافة تُعطي ترميز المعلومات الدلالية الغنية في خريطة السمات كنتيجة، إلا أن المعلومات التفصيلية المتعلقة بحدود الكائنات مفقودة بسبب عمليات التجميع أو التلافيف داخل بنية الشبكة.

الجزء الثاني (decoder): يعتمد مبدأ عمله على فك تشفير غمامة الكثافة وتوضيحها كمعالم، أي هو تطوير على عمل الخوارزميات (auto encoder)، حيث تقوم (decoder) بإيجاد وفك غمامة الكثافة من الصور وتحسين نتائج التقسيم الدلالي من خلال النقاط حدود أكثر وضوحاً للكائنات، عن طريق استرداد المعلومات المكانية تدريجياً.



الشكل (2). مبدأ عمل نموذج الشبكة العصبونية الالتفافية العميقة (DeepLabv3+) [4].

#### البيانات المستخدمة في البحث:

تم تأمين البيانات اللازمة للبحث، وهي:

(1) بيانات تدريب نموذج الشبكة العصبونية للتعليم العميق: وهي مجموعة (dataset) تحتوي على مسارات لعدد (8570) صورة، كل منها بأبعاد (1024\*1024\*3)، بدقة (50cm pixel resolution)، بثلاث قنوات لونية (RGB)، فالصورة بالنسبة إلى الحاسوب عبارة عن مصفوفة ثلاثية الأبعاد (العمق×الارتفاع×العرض) من قيم تتراوح بين (0-255)، تغطي مجموعة البيانات هذه كلاً من مناطق حضرية وريفية. تم الحصول على بيانات التدريب هذه من معهد ماساتشوستس للتكنولوجيا (MIT - Massachusetts Institute of Technology)، بمدينة كامبريدج، يضم مراكز أبحاث بارزة معروفة عالمياً في مجالات البحث العلمي والتكنولوجي، لعب هذا المعهد دوراً

أساسيا في تطوير علم البيانات (Data Science)، وتصميم الحواسيب العملاقة (GPU و TPU)، ويتيح مجموعة البيانات لمشاكل عديدة في العالم.

(2) صورتين فضائيتين من غوغل إيرث (Google Earth)، لمنطقتي دراسة مختلفتين، من أجل اختبار أداء النموذج المُدرَّب على عينات جديدة للتنبؤ بالطرق، هما:

- a. صورة غوغل إيرث لجزء من منطقة سكنية في المشروع العاشر في مدينة اللاذقية، الشكل (3-a).  
b. صورة غوغل إيرث لجزء من أرض زراعية في منطقة مشقيتا في ريف اللاذقية، الشكل (3-b).



a- صورة فضائية لمنطقة سكنية في المشروع العاشر b- صورة فضائية لأرض زراعية في منطقة مشقيتا

الشكل (3). منطقتي الدراسة من أجل اختبار أداء النموذج المُدرَّب للتنبؤ بالطرق

#### التطبيق البرمجي:

(1) لغة البرمجة المستخدمة: قمنا بكتابة الكود البرمجي بلغة بايثون، وهي اللغة الأكثر إستخداماً في مجال الذكاء الاصطناعي (Artificial Intelligence)، ولديها العديد من المكتبات الملائمة لتطوير نماذج التعلم الآلي (Machine Learning) والتعلم العميق [12,13].

(2) بيئة العمل البرمجية: بما أننا نحتاج إلى بيئة برمجية متطورة ومواصفات حاسوب متقدمة ليتمكن النظام من تحمل عبء الذاكرة الكبيرة ومعالجة البيانات الضخمة. كان من الصعب تأمين هذه المواصفات العالية في أجهزتنا. ولذلك تم الاعتماد على منصة غوغل كولا ب (Google Colab) التي ساهمت بمساعدة الباحثين في مجال الذكاء الاصطناعي، وهي نسخة سحابية مُحسَّنة من (Jupyter Notebook)، تمكَّنتنا من كتابة وتنفيذ التعليمات البرمجية بلغة بايثون، ويتميز بوجود العديد من المكتبات المُثبتة مُسبقاً، يمكن استخدام هذه المنصة بسهولة وربطها بحساب غوغل، وتوفّر ذاكرة وصول عشوائي (RAM) إضافية، وتتيح القدرة على استخدام وحدات معالجة الرسومات (GPU و TPU)، التي تُستخدم لبناء المشاريع البرمجية التي تتضمن تقنيات التعلم الآلي أو التعلم العميق، وتؤمن الموارد اللازمة لتشغيلها، وخاصةً لمن لا يمتلك وحدة معالجة رسومات متطورة في جهازه [14].

(3) المكتبات المستخدمة: تم الاعتماد على مجموعة من المكتبات المتطورة في لغة البرمجة بايثون [13]، أهمها:

- مكتبة بايتورث PyTorch: هي مكتبة تعلم عميق مفتوحة المصدر مبنية على لغة البرمجة بايثون تُستخدم لتطبيقات الرؤية الحاسوبية، إن تدريب النموذج واختباره يتطلب تنصيب مكتبة بايتورث PyTorch ، بالإضافة إلى أداءها العالي في المعالجة من خلال الاستفادة من وحدة معالجة الرسومات GPU.
- مكتبة الرؤية الحاسوبية (OpenCv): وهي مكتبة مفتوحة المصدر لمهام الرؤية الحاسوبية ومعالجة الصور في التعلم الآلي والعميق، غنية بالخوارزميات والتوابع التي تسهل عمليات التعامل مع الصور ومعالجتها، حيث يمكننا قراءة الصور وإجراء العديد من العمليات عليها كتحويل الصور الملونة إلى صور رمادية، واقتصاص الصور، وتغيير أبعاد الصور، وتحليل الصور.
- مكتبة Numpy: هي مكتبة متخصصة في الحوسبة العلمية بلغة البايثون، للتعامل مع العمليات الرياضية المعقدة مثل الجبر الخطي وتحويل فورييه والميزات التي تعمل مع المصفوفات في بايثون.
- مكتبة Matplotlib: هي إحدى أشهر مكتبات بايثون، مهمتها عمل الرسومات البيانية ذات البعدين وعرضها، تعتمد تلك المكتبة على مكتبة NumPy للتعامل مع المصفوفات الكبيرة لضمان أداء أفضل.
- مكتبة pandas: هي مكتبة مفتوحة المصدر، باستخدام هذه المكتبة يمكنك قراءة البيانات من مجموعة واسعة من المصادر مثل CSV وقواعد بيانات SQL وملفات JSON و Excel. كما أنها تقدم مجموعة واسعة من الأدوات لمعالجة البيانات مثل تنظيف البيانات، وإجراء العمليات الإستكشافية على البيانات، وإمكانية إعادة تشكيل هيكل البيانات، وتسهيل دمج البيانات ببعضها أو تجزئتها.

### النتائج والمناقشة:

- لتحقيق الهدف من البحث، وللوصول إلى النتائج المرجوة، تم إنجاز الدراسة وفق مراحل عمل متسلسلة، كما يأتي:
- أولاً: المرحلة التحضيرية والمعالجة المسبقة للبيانات.
- ثانياً: استدعاء نموذج الشبكة العصبونية العميقة المقترحة، ثم ضبط إعداداته، وتدريب النموذج والتحقق.
- ثالثاً: تقييم دقة أداء النموذج المُدرَّب.
- رابعاً: تطبيق النموذج المُدرَّب على منطقة الدراسة لاختبار أدائه في استخراج الطرق.
- أولاً: المرحلة التحضيرية والمعالجة المسبقة للبيانات:
- شملت المرحلة التحضيرية والمعالجة المسبقة للبيانات عدة خطوات متسلسلة، وهي:
- (1) تنصيب مكتبة (segmentation-models-pytorch) إلى منصة غوغل كولا، وهي المكتبة التي ينتمي إليها نموذج الشبكة العصبونية العميقة المقترحة في بحثنا وهي (deeplabv3+)، كما تم استدعاء المكتبات الأخرى التي سنحتاجها في كل من مراحل المعالجة المسبقة للبيانات وتدريب النموذج واختباره.
  - (2) تحميل مجموعة البيانات (dataset) إلى منصة غوغل كولا.
  - (3) استكشاف مجموعة البيانات التي تحوي (8570 سجل، و 4 أعمدة)، وفي كل سجل بيانات لصورة أبعادها (1024\*1024\*3)، بثلاث قنوات لونية هي (RGB)، بدقة (50cm pixel reselotion).
- عند استكشاف مجموعة البيانات، الشكل (4)، تبين الآتي:

- بعض السجلات تحوي مسارات الصور الأصلية (sat image path) بالإضافة إلى مسارات صور الماسك (mask path) لنفس المنطقة، وهي تتبع للسجلات ضمن العمود (Split=train)، تفيد في تعليم النموذج المقترح لاستخراج الطرق تعليمياً خاضعاً للإشراف (Supervised Learning)، حيث تحتوي الصور الأصلية على كامل تفاصيل الصورة المعقدة (كافة المُسققات والطرق والأراضي المغطاة بالنباتات والأشجار والجبال...)، في حين تحتوي صور الماسك على أصناف للطرق ولأنواع مختلفة من المعابر.
- يوجد سجلات أخرى تحتوي فقط على مسارات الصور الأصلية، ولا تحوي على مسارات صور الماسك، وهي تتبع للسجلات ضمن العمود (Split=test)، لكنها لن تخدم بحثنا في عملية اختبار وتقييم دقة أداء النموذج، الأمر الذي تطلب منا حذف السجلات في حال حوت قيمة (mask path=NaN)، فهي بيانات زائدة عن الحاجة، وهذا ما يسمى بعملية تنظيف مجموعات البيانات.

	image_id	split	sat_image_path	mask_path
0	100034	train	train/100034_sat.jpg	train/100034_mask.png
1	100081	train	train/100081_sat.jpg	train/100081_mask.png
2	100129	train	train/100129_sat.jpg	train/100129_mask.png
3	100703	train	train/100703_sat.jpg	train/100703_mask.png
4	100712	train	train/100712_sat.jpg	train/100712_mask.png
...	...	...	...	...
8565	998829	test	test/998829_sat.jpg	NaN
8566	998928	test	test/998928_sat.jpg	NaN
8567	999518	test	test/999518_sat.jpg	NaN
8568	999651	test	test/999651_sat.jpg	NaN
8569	999703	test	test/999703_sat.jpg	NaN

8570 rows x 4 columns

الشكل (4). استكشاف مجموعة البيانات

من الضروري استخدام التوزيع العشوائي (randomization) أثناء تصميم تجارب التدريب، بحيث لا يكون هناك تحيز في تدريب نموذج الشبكة العصبونية على صور بحالات معينة وإهمال صور بحالات أخرى، الأمر الذي يقلل من نجاح عملية التدريب، فيفشل النموذج في الحصول على تنبؤات صحيحة في بيانات جديدة، ونقع في الحالة المسماة (Overfitting) [13]، لذلك تم تفعيل ميزة خلط البيانات باستخدام البارامتر (shuffle=true)، وهذا يحقق أخذ العينات بطريقة عشوائية (Random Sampling).

ويبين الشكل (5) الآتي مجموعة البيانات الجديدة الناتجة بعد عمليتي تنظيف وخلط البيانات، حيث أصبح لدينا عدد العينات اللازمة لانجاز بحثنا هو (6226) عينة موزعة بشكل عشوائي.



	image_id	sat_image_path	mask_path
656	197397	train/197397_sat.jpg	train/197397_mask.png
5250	854448	train/854448_sat.jpg	train/854448_mask.png
5113	834943	train/834943_sat.jpg	train/834943_mask.png
1984	381744	train/381744_sat.jpg	train/381744_mask.png
5075	829648	train/829648_sat.jpg	train/829648_mask.png
...	...	...	...
5682	919207	train/919207_sat.jpg	train/919207_mask.png
2785	49425	train/49425_sat.jpg	train/49425_mask.png
1817	35917	train/35917_sat.jpg	train/35917_mask.png
3570	608673	train/608673_sat.jpg	train/608673_mask.png
1468	307908	train/307908_sat.jpg	train/307908_mask.png
6226 rows × 3 columns			

الشكل (5). مجموعة البيانات الجديدة

4) ثم قمنا بتقسيم مجموعة البيانات الجديدة لدينا إلى قسمين وفق معيار التقسيم (90:10)، وهي نسبة التقسيم القياسية المستخدمة في بناء الشبكات العصبونية للتعليم العميق، بحيث (5603) من بيانات بحثنا تُستخدم كمجموعة عينات لتدريب النموذج (Training data)، و(623) من البيانات تم وضعها جانباً كمجموعة عينات للتحقق (validation data) من أداء النموذج على بيانات جديدة لم يسبق له أن رآها.

5) تم العمل على استخلاص صور الحقيقة الأرضية (Ground Truth Mask) التي تحتوي على صنفين أساسيين فقط، هما: الطريق وغير الطريق؛ لأن هدفنا هو تعليم الشبكة على استخراج الطرق وإهمال كل المعالم الأخرى. وذلك عبر خطوات تعتمد على توابع التقسيم الدلالي للصور [3]، كالآتي:

(a) إنشاء مصفوفة بصنفين فقط، هما: الطريق (road)، وخلفية الأرض غير الطريق (background). وكذلك مصفوفة الألوان المقابلة لها، كالآتي:

مصفوفة الأصناف: Classes Name :['road', 'background']

مصفوفة الألوان: Class RGB Values :[[255, 255, 255], [0, 0, 0]]

حيث يقابل الطريق اللون الأبيض [255, 255, 255]، ويقابل خلفية الأرض اللون الأسود [0, 0, 0].

(b) تفكيك الصورة الماسك بتطبيق التابع (one hot encode): وهو تابع يقوم بتحويل الصورة (RGB) إلى صورة بنفس الطول والعرض ولكن عمقها يعبر عن عدد الأصناف، عند تمرير مصفوفة الأصناف على صورة الماسك، يكون لدينا أحد الاحتمالات الآتية:

إما [True, False]: أي لدينا صنف طريق، وستكون بالشكل: [1,0]  
أو [False, True]: أي لدينا صنف أرض، وستكون بالشكل: [0,1]  
أو [False, False]: وهنا لدينا صنف آخر غير الأرض والطريق، قد تكون عتبة أحد الأبنية: [0,0].  
أو [True, True]: وهذه مستحيلة، إذ لا يمكن أن يكون لدينا صنف طريق وغير طريق في نفس الوقت.  
أي تم تحويل صور الماسك ( $3 * 1024 * 1024$ ) إلى صورة بعمقين ( $2 * 1024 * 1024$ ) هما الطريق والأرض.  
C ثم تطبيق التابع (reverse one hot encode): لطي المصفوفة الناتجة عن (one hot encode) إلى مصفوفة  
بعمق واحد، عن طريق أخذ (argmax) في كل موضع بكسل، وبذلك حصلنا على صورة ( $1 * 1024 * 1024$ )  
أسميناها (reverse hot encoded mask) قيمها إما (0) أو (1)، أي تحوي دليل صنف الطريق والأرض في  
مصفوفة الأصناف لدينا. كما مٌبين في الشكل (6).  
d) لتحويل الصورة (reverse hot encoded mask) إلى صورة (RGB)، طبقنا التابع (code color  
segmentation)، لاستبدال دليل الصنف بالقيمة اللونية للصنف في مصفوفة الألوان بالأعماق الثلاث، بالنتيجة  
حصلنا على صورة الحقيقة الأرضية (Ground truth mask) التي تحتوي فقط على صنفى الطرقات والأرض  
( $3 * 1024 * 1024$ ) قيمها إما (0) أو (255)، كما مٌبين في الشكل (6).  
وبذلك انتهت مرحلة المعالجة المسبقة لمجموعة البيانات، وأصبحت الصور الأصلية وصور الحقيقة الأرضية  
(Ground Truth Mask) جاهزة لتدريب نموذج الشبكة العصبونية العميقة المقترحة في بحثنا.



الشكل (6). إجراء المعالجة المسبقة لمجموعة البيانات واستخلاص صور الحقيقة الأرضية (Ground Truth Mask)

ثانياً: استدعاء نموذج الشبكة العصبونية العميقة المقترحة، ثم ضبط إعداداته، وتدريب النموذج والتحقق. تم استدعاء نموذج الشبكة العصبونية العميقة المقترحة في بحثنا وهو (deeplabv3+) المبني بواسطة مكتبة (segmentation-models-pytorch)، حيث يؤدي استخدام نموذج تم تدريبه مسبقاً على مجموعات بيانات كبيرة مختلفة أخرى (التمييز بين القطط والكلاب) إلى تقليل وقت التدريب والتجهيز الزائد، وهو نهج يُعرف بإسم نقل التعلم الذي يُمكن من استخدام الأوزان والبنية التي تم الحصول عليها وتطبيق التعلم على بيان المشكلة الجديدة لدينا [3]، وهي استخراج الطرق من الصور الفضائية.

قمنا بإجراء تعديلات على النموذج الموجود مسبقاً (deeplabv3+)، عن طريق ضبط إعدادات النموذج (hyperparameters) بما يُحسن عملية تدريبه على بيانات بحثنا، من أجل تطويره لأداء مهمة استخراج الطرق من الصور الفضائية، حيث تم اتباع الآتي في ضبط إعدادات النموذج:

- تهيئة الوزن (Weight initialization): إن تهيئة أوزان الشبكات العصبونية العميقة له تأثير كبير على تقليص الخطأ، ورفع الدقة، تُستخدم أوزان نماذج (Resnet50) المدربة مسبقاً على (ImageNet) على نطاق واسع لبناء هياكل شبكات عصبونية عميقة مختلفة، نظراً لأنها مبنية على مجموعات بيانات كبيرة بما يقارب 1.2 مليون صورة [15].

تم تحسين أداء (deeplabv3+) من خلال تهيئة الشبكة بأوزان نماذج Resnet50 المدربة مسبقاً على ImageNet.

- ضبط معدل التعلم (Learning rate): للحصول على أداء أفضل وتسريع المعالجة أثناء التدريب تم ضبط معدل التعلم وفق المُحسن آدم (Adam's optimization) الذي يعد أفضل خوارزميات التحسين للتعلم العميق التي تعمل على تحديث الأوزان باستمرار، وتسريع زمن التدريب، حتى يتقارب تابع الخسارة (loss functions) من الصفر [9].
- حجم الدفعة (Batch size): لا يمكن تمرير مجموعة بيانات التدريب والاختبار كاملة إلى الشبكة مرة واحدة، لذلك يتم تقسيم مجموعة البيانات إلى أجزاء يشار إليها بحجم الدفعة، وهي قيمة تجريبية [13]. تم في عملنا اختيار حجم دفعة البيانات (Patch Size=4) أي سيتم تمرير كل 4 صور تلو الأخرى إلى الشبكة لتدريبها.
- عدد الفترات (Number of epochs): هو عدد المرات التي يمكن أن تمر فيها مجموعة البيانات بأكملها ذهاباً وإياباً عبر الشبكة العصبونية مرة واحدة فقط، وهي قيمة تجريبية، حيث نستمر بزيادة عدد الفترات Epochs حتى الوصول في تابع الخسارة loss functions إلى الحد الأدنى [13].

تم إدخال مجموعة بيانات التدريب إلى الشبكة (deeplabv3+)، حيث تم التعامل مع الصور الأصلية كمدخلات، وصور الحقيقة الأرضية كمخرجات، لتدريب النموذج على التمييز بين صنفَي الطريق والأرض تعليماً خاضعاً للاشراف، وقامت الشبكة بتعديل الأوزان بالاعتماد على كُلى من الخرج المرغوب (صور الحقيقة الأرضية) والخرج الذي تعطيه (صور الطرق المُستخرجة المُنتبأ بها)، من خلال حساب تابع الخطأ، والبحث عن النهاية الصغرى لهذا التابع، لتحقيق التنبؤ الصحيح. ثم قمنا بإدخال مجموعة بيانات التحقق لاختبار النموذج المدرب على عينات التحقق الجديدة، وزيادة القدرة التعليمية له.

عند تدريب الشبكة، لاحظنا أن الخطأ تابع الخسارة (Loss Function= 0.1195) لا يقل في الفترات الأولى، ولذلك تم زيادة عدد الفترات لزيادة تدريب الشبكة وإنقاص تابع الخسارة ليصبح أقل ما يمكن.

توقفنا عند عدد فترات هو (Epochs=4)، لأننا لاحظنا ثبات قيمة الخطأ عند الحدود الدنيا لتابع الخسارة (0.0252= Loss Function)، وأصبحت زيادة الفترات غير مجدية في تدريب الشبكة.

إن ذاكرة الوصول العشوائي (RAM) المتاحة على منصة غوغل كولا ب هي 12 غيغا، لكنها لم تكفي لتدريب النموذج في بحثنا، لذلك احتجنا إلى استئجار رام خارجي أون لاين إضافي يعادل 32 غيغا، وتم استهلاك 10 غيغا منه، وبالتالي عملية تدريب النموذج واختباره تطلبت 22 غيغا. بالإضافة إلى أن عملية التدريب تمت بالاستعانة بوحدة معالجة الرسومات المتطورة (GPU) وذلك لتقليل الزمن المستغرق في تدريب النموذج على مجموعات البيانات، حيث استغرق زمن التدريب مدة ثلاث ساعات وعشرة دقائق، في حين لو تم على الحواسيب العادية ذات وحدة المعالجة المركزية (CPU) كان سيستغرق بحدود 11 ساعة أقل ما يمكن.

#### ثالثاً: تقييم دقة أداء النموذج المُدرَّب.

تم تقييم أداء النموذج (deeplabv3+) في استخراج الطرق من صور الأقمار الصناعية، من خلال مقياسي تقييم الدقة الأساسيين للشبكات العصبية العميقة لمهمة التقسيم الدلالي، وهما:

- تابع الخسارة (Loss Function): هو مقياس نسبة الخطأ في التنبؤات، لمعرفة ما إذا كانت الشبكة قد تعلمت بدقة من بيانات التدريب أم لا، فكلما قل تابع الخسارة كلما كانت الشبكة أفضل، تُحسب الخسارة عن طريق حساب الخطأ المتوسط التربيع، وفي نماذج التقسيم الدلالي يُستخدم التابع اللوغاريتمي (Cross-Entropy). تتراوح قيمة تابع الخسارة بين 0 و 1، ويدل الرقم (0) على أن النموذج المُدرَّب مثالي [8,10].

▪ التابع ((Intersection over Union (IOU)): هو مقياس نسبة الصحة في التنبؤات، يُستخدم بشكل أساسي في خوارزميات اكتشاف الكائنات والتقسيم الدلالي [4]، يعطي نسبة مساحة التداخل بين منطقة التنبؤ ومنطقة الحقيقة الأساسية للكائن، ويتم تحديده من خلال التقاطع إلى الإتحاد، وفق المعادلة (1) الآتية [16,17]:

$$IOU = \frac{TP}{(TP + FP + FN)} \quad (1)$$

حيث يشير كل من TP وFP و FN إلى:

TP (true positive): عدد وحدات البكسل المتنبأ بها الإيجابية الحقيقية، وهي منطقة التقاطع بين الحقيقة الأرضية ومنطقة التنبؤ.

FP (False Positive): عدد وحدات البكسل المتنبأ بها الإيجابية الكاذبة، وهي في منطقة التنبؤ خارج منطقة الحقيقة الأرضية.

FN (false negatives): عدد وحدات البكسل المتنبأ بها السلبية الكاذبة، وهي في منطقة الحقيقة الأرضية التي فشل النموذج في التنبؤ بها.

تم تحديد العتبة (IOU(Threshold)=0.5)، ويتراوح مقياس IoU بين 0 إلى 1. وبديل الرقم (1) على أنّ النموذج المُدرّب مثالي [17].

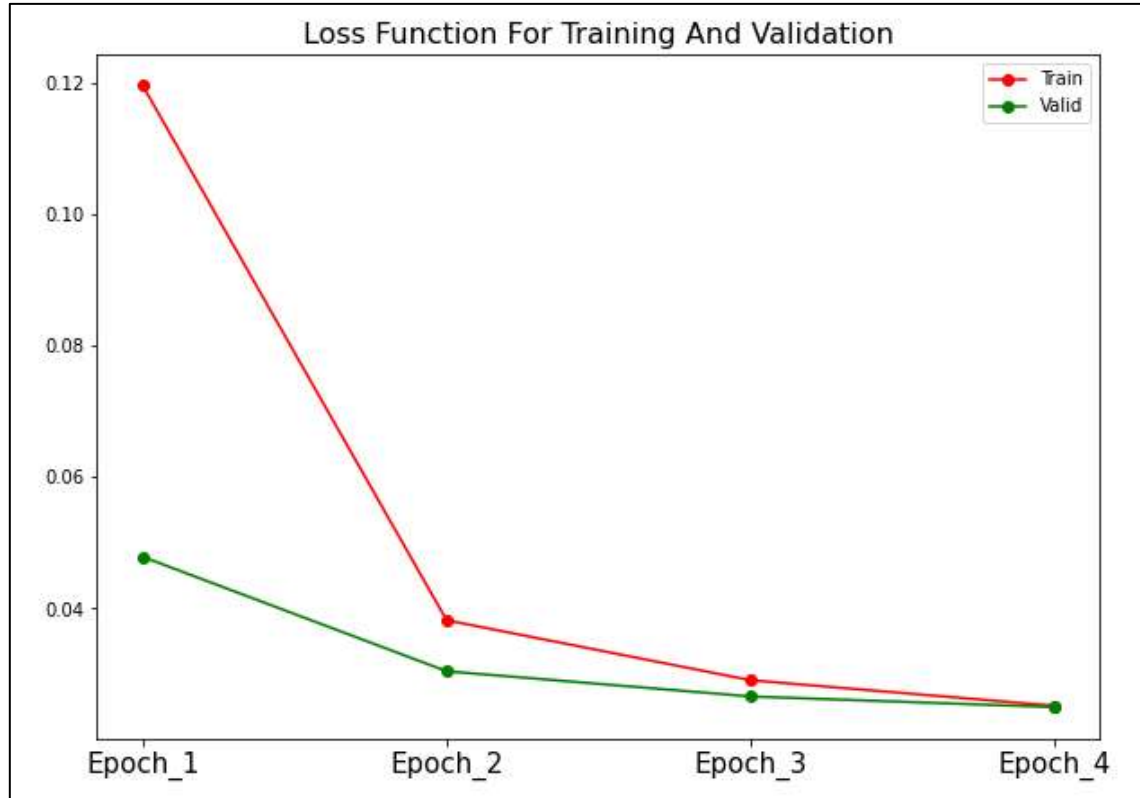
تطلبت عملية التدريب والتحقق في بحثنا عدد فترات هو (Epochs=4)، عند كل فترة سنقوم بعرض دالة الخسارة (Loss Function) ومقياس (IOU)، لكل من مرحلتي التدريب والتحقق.

يبين الجدول (1) دالة الخسارة عند كل فترة من مرحلتي التدريب والتحقق، حيث تم حساب نسبة الخطأ في التنبؤات فيما إذا كان البكسل من صنف الطريق أم الأرض، نلاحظ نقصان الخطأ في مرحلة التدريب من (0.1195) في الفترة الأولى إلى (0.0252) في الفترة الرابعة، وكذلك نقصان الخطأ في مرحلة التحقق من (0.0479) في الفترة الأولى إلى (0.0250) في الفترة الرابعة.

الجدول (1). دالة الخسارة عند كل فترة لكل من مرحلتي التدريب والتحقق.

الفترة (Epochs)	دالة الخسارة (Loss Function)	
	خطأ التدريب (Training Loss)	خطأ التحقق (Validation Loss)
1	0.1195	0.0479
2	0.0382	0.0304
3	0.02909	0.0266
4	0.0252	0.0250

ويبين الشكل (6) مقارنة منحنى تابع الخسارة في مرحلة التدريب مع مرحلة التحقق، يشير المنحني الأحمر إلى منحنى تابع الخسارة في التدريب، والأخضر إلى منحنى تابع الخسارة في التحقق. نلاحظ تلاشي الفجوة بين خطأ التدريب وخطأ التحقق، مع نقصان الخسارة، بزيادة عدد الفترات، مما يثبت أن الشبكة العصبونية الالتفافية العميقة (deeplabv3+) تدرّبت لتؤدي أداءً جيّداً في استخراج الطرق من صور الأقمار الصناعية.



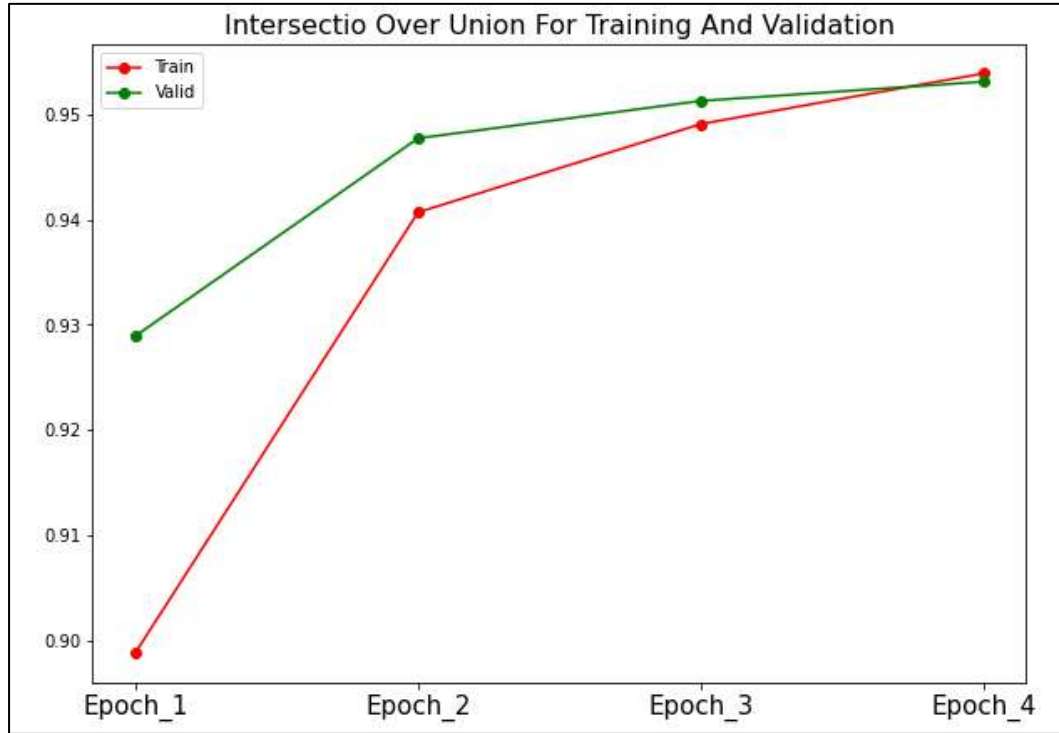
الشكل (6) مقارنة منحنيي تابع الخسارة في مرحلة التدريب مع مرحلة التحقق

يبين الجدول (2) مقياس (IOU) عند كل فترة من مرحلتي التدريب والتحقق، حيث تم حساب متوسط الدقة في صحة التنبؤات فيما إذا كان البكسل من صنف الطريق أم الأرض، نلاحظ زيادة صحة التنبؤ في مرحلة التدريب من (0.8988) في الفترة الأولى إلى (0.9539) في الفترة الرابعة، وكذلك زيادة صحة التنبؤ في مرحلة التحقق من (0.9289) في الفترة الأولى إلى (0.9531) في الفترة الرابعة.

الجدول (2) مقياس IOU عند كل فترة لكل من مرحلتي التدريب والتحقق.

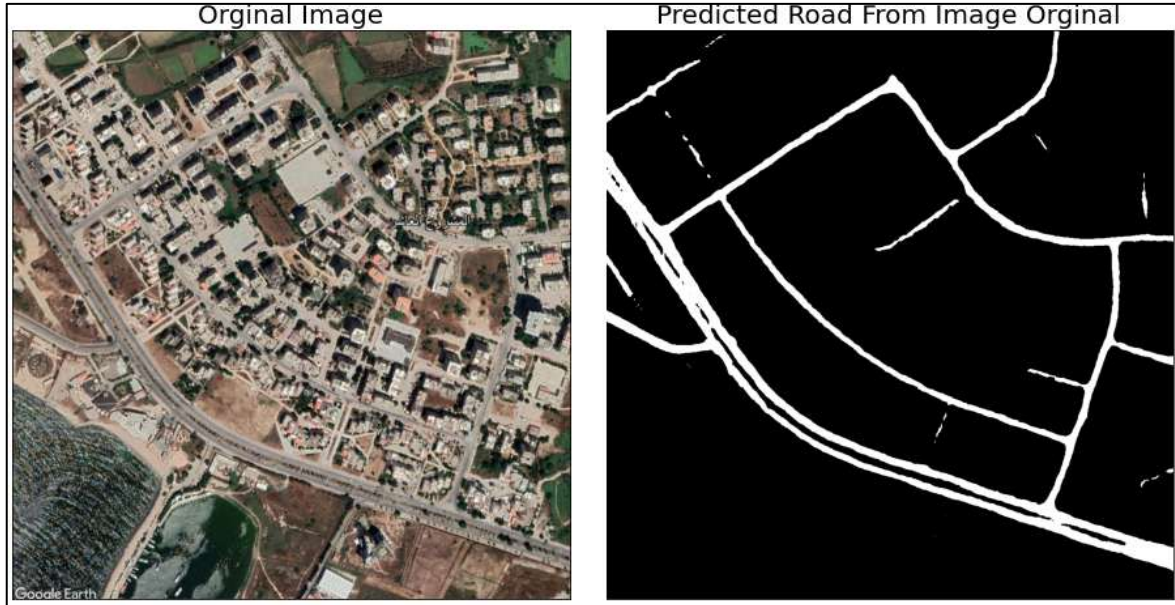
الفترة (Epochs)	مقياس IOU	
	في التدريب (Training IOU)	في التحقق (Validation IOU)
1	0.8988	0.9289
2	0.9407	0.9477
3	0.9491	0.9513
4	0.9539	0.9531

ويبين الشكل (7) مقارنة منحنيي متوسط قياس الدقة في صحة التنبؤ (IOU) في مرحلة التدريب مع مرحلة التحقق. نلاحظ تلاشي الفجوة وزيادة صحة التنبؤ في مرحلة التدريب والتحقق بزيادة عدد الفترات. مما يثبت أن الشبكة العصبية الالتفافية العميقة (deeplabv3+) تدرت لتؤدي أداءً جيدًا في استخراج الطرق من صور الأقمار الصناعية.

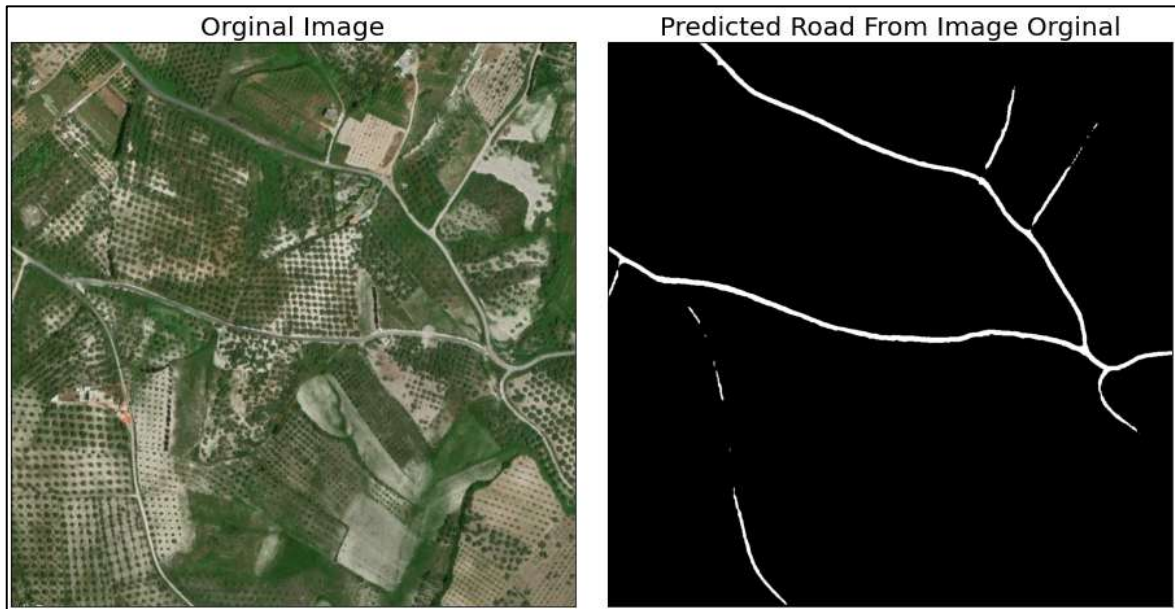


الشكل(7). مقارنة منحنيي مقياس صحة التنبؤ IOU في مرحلة التدريب مع مرحلة التحقق

- كما تدل المنحنيات في الشكلين (6 و 7) إلى تلافي مشكلتين أساسيتين قد تظهرها في حالات التعلم العميق [13]، هما:
- مشكلة الضبط الزائد (Overfitting): عندما يتعلم النموذج على بيانات التدريب بشكل مثالي فتكون نسبة الخطأ صغيرة في مرحلة التدريب، لكنها عالية في مرحلة التحقق، أي يفشل في الحصول على تنبؤات صحيحة في بيانات جديدة، عندها سيكون هناك فجوة بين منحنيي التدريب والتحقق.
  - مشكلة الضبط الناقص (Underfitting): عندما تكون نسبة الخطأ عالية في كل من مرحلة التدريب والتحقق معاً، عندها نقول أن النموذج لم يستطع التعلم من بيانات التدريب بشكل جيد.
- رابعاً: تطبيق النموذج المُدرَّب على منطقة الدراسة لاختبار أدائه في استخراج الطرق.
- من أجل اختبار وتوضيح أداء النموذج المُدرَّب للتنبؤ بالطرق على عينات جديدة تمثل مناطق حضرية وأخرى ريفية، تم اختيار منطقتي دراسة مختلفتين، هما:
- المنطقة الأولى: جزء من منطقة سكنية في المشروع العاشر في مدينة اللاذقية.
  - المنطقة الثانية: جزء من أرض زراعية في منطقة مشقيتا في ريف اللاذقية.
- وكانت نتائج الاختبار كما يبين الشكلان (8) و(9)، حيث يمثل العمود الأول الصورة الأصلية، ويمثل العمود الثاني التنبؤ بالطرق بالطريقة المقترحة.



الشكل (8). نتائج تطبيق النموذج المُدرَّب (deeplabv3+) على منطقة الدراسة الأولى



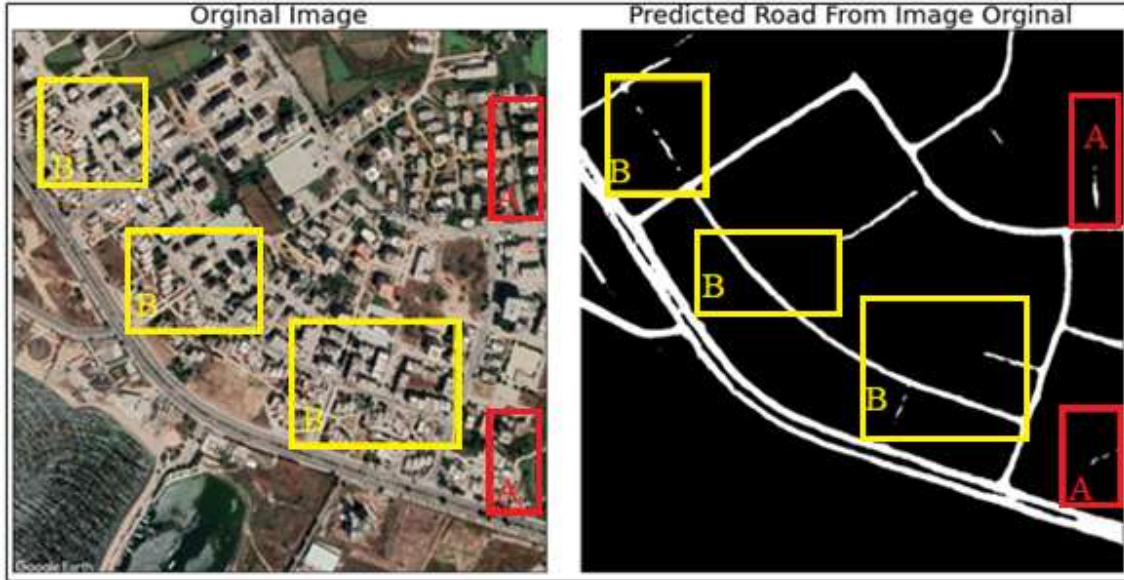
الشكل (9). نتائج تطبيق النموذج المُدرَّب (deeplabv3+) على منطقة الدراسة الثانية

أظهر نموذج الشبكة العصبوية الالتفافية العميقة (deeplabv3+) المقترح في بحثنا أداءً جيداً في استخراج الطرقات من الصور الفضائية في منطقتي الدراسة المختاريتين، إلا أن نتائج صور التنبؤ بالطرق كشفت عن بعض السلبيات في هذه الطريقة المقترحة، فقد كان أداء النموذج ضعيفاً في استخراج الطرق في بعض المناطق الأتية:

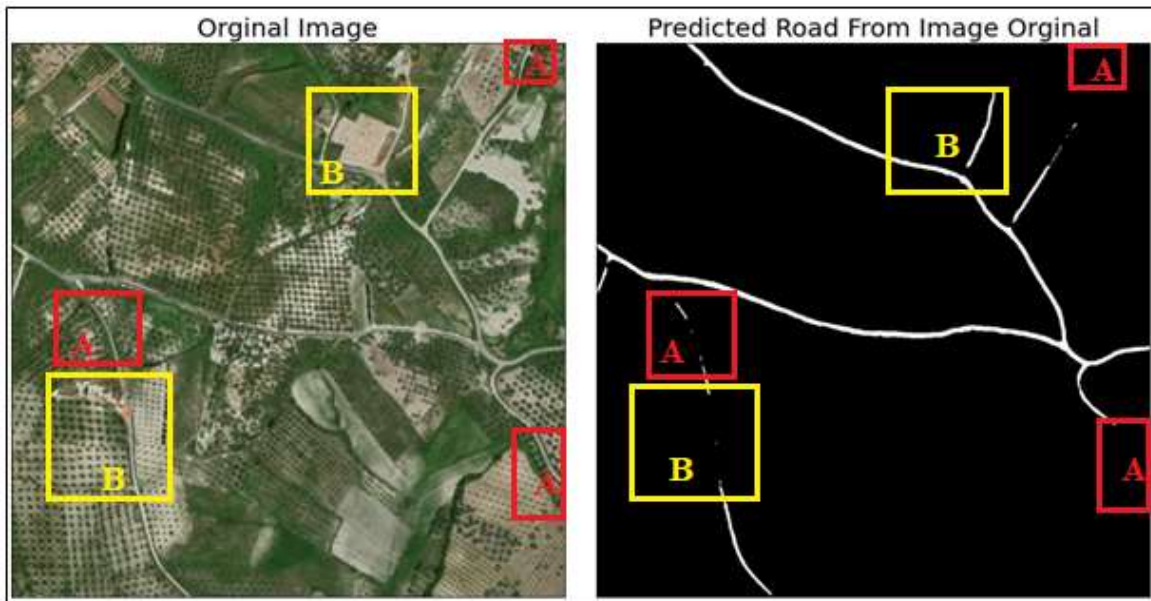
- المناطق التي تحوي أشجاراً بسبب تأثير ظلال الأشجار على جوانب الطرق، كما تبيين الأجزاء ذات الرمز (A) في كل من الشكلين (10) و (11).
- المناطق التي تحوي كائنات خصائصها الطيفية مشابهة للطريق كأسطح بعض المباني... الخ، كما تبيين الأجزاء ذات الرمز (B) في كل من الشكلين (10) و (11).



- كما لم يتمكن النموذج (deeplabv3+) من استكشاف الطرق الفرعية قليلة العرض، والطرق غير المعبدة؛ فهياكل الطرق معقدة في الواقع، وأنواع الطرق المختلفة لها أبعاد مختلفة (ليست بعرض واحد).



الشكل (10). الأجزاء (A,B) من منطقة الدراسة الأولى التي كان أداء النموذج (deeplabv3+) ضعيفاً في استخراج الطرق فيها



الشكل (11). الأجزاء (A,B) من منطقة الدراسة الثانية التي كان أداء النموذج (deeplabv3+) ضعيفاً في استخراج الطرق فيها

### الاستنتاجات والتوصيات:

- نتيجة لهذه الدراسة، تم التوصل إلى مجموعة من الاستنتاجات، يمكن إيجازها بالآتي:
- أدى اتباع نهج نقل التعلم إلى تطوير نموذج الشبكة العصبونية الالتفافية (deeplabv3+) لأداء مهمة استخراج الطرق من الصور الفضائية، وذلك من خلال تدريبه على مجموعة بيانات بحثنا الخاصة بالطرق، وتهيئة بأوزان نماذج

Resnet50، واستخدام خوارزمية التحسين آدم، مما ساهم في تسريع المعالجة، وتحسين الدقة، حيث تطلبت عملية التدريب والتحقق عدد فترات هو (Epochs=4)، وصلت دالة الخسارة إلى (Loss Function. =0.025).  
2. وفرت منصة غوغل كولا ببيئةً برمجيةً متطورةً لتدريب النموذج في بحثنا، فقد تطلبت عملية التدريب رامات خارجية بحجم 22 غيغا، بالإضافة إلى الاستعانة بوحدة معالجة الرسومات المتطورة GPU، فاستغرقت زمن تدريب ثلاث ساعات وعشرة دقائق، في حين لو تم على وحدة المعالجة المركزية CPU للحاسب المحمول كان سيستغرق زمن بحدود 11 ساعة أقل ما يمكن.

3. وصل متوسط قياس الدقة في صحة تنبؤات النموذج المدرب (التقاطع إلى الاتحاد) إلى (IOU =0.953)، مما يثبت أن الشبكة (deeplabv3+) تدريباً تؤدي أداءً جيداً في استخراج الطرق.

4. كشفت نتائج الدراسة عن ضعف أداء النموذج في استخراج الطرق بدقة، في الحالات الآتية:

a. المناطق التي تحوي أشجار بسبب تأثير ظلال الأشجار على حواف الطرق.

b. المناطق التي تحوي كائنات خصائصها الطيفية مشابهة للطريق كأسطح بعض المباني.

c. الطرق الفرعية قليلة العرض، وكذلك الطرق غير المعبدة.

مما يدل على حساسية النموذج (deeplabv3+) للضجيج، وتأثير السمات الهندسية ومعلومات النسيج للطرق على دقة استخراج الطرق بالطريقة المقترحة.

كما تم التوصل إلى عدة توصيات يمكن أخذها في الحسبان، على الشكل الآتي:

1. نقترح إيجاد طرق لإدخال تمثيل سمات النسيج والحواف في صور الحقيقة الأرضية لتتكامل مع إستراتيجية نموذج التقسيم الدلالي (deeplabv3+) بهدف استخراج الطرق قليلة العرض وغير المعبدة، واستكشاف معلومات السياق التي تكشف عن العلاقات بين كائنات الصورة (الطرق، أسطح المباني، الأشجار، المركبات على الطرق...)، وبالتالي تقليل فقد المعلومات المكانية.

2. نوصي بتطوير المنهجية المقترحة وتطبيقها على بيانات صور أقمار صناعية عالية الدقة المكانية ومتعددة الأطياف، ودراسة تأثير أهمية الدقة المكانية والدقة الطيفية على استخراج الطرق.

3. تطوير شبكة عصبية عميقة أكثر تحسناً لاستخراج الطريق بتنسيق متجه (vector format)، والتي يمكن استخدامها مباشرة للتطبيقات العملية في تحديث خرائط الطرق وأعمال التخطيط الحضري.

4. نوصي الباحثين بالاستفادة من البيئة السحابية غوغل كولا ب، المتاحة جنباً إلى جنب مع وحدة معالجة الرسومات GPU، لمعالجة البيانات الضخمة وتحمل عبء الذاكرة الكبير وتحسين سرعة تدريب نماذج الشبكات العصبونية، مما يوفر عليهم الحاجة إلى استخدام حواسيب بمواصفات عالية.

## خاتمة

تبني هذا البحث بنية نموذج الشبكة العصبونية الالتفافية العميقة (deeplabv3+) الذي حقق حديثاً أداءً مذهلاً في التقسيم الدلالي للصور، وتم توظيفه في بحثنا لمهمة استخراج الطرق من الصور الفضائية، فحققنا نتائج جيدة في تقييم دقة أداء النموذج بعد تدريبه، لكن لا تزال هناك حاجة إلى تحسين أداء النموذج (deeplabv3+) في استخراج الطرق من صور الأقمار الصناعية عالية الدقة، لتقليل فقد المعلومات المكانية، وبما يفيد في تحديث خرائط الطرق وأعمال التخطيط الحضري.

## References:

1. Keiron, O.; Ryan, N. *An Introduction to Convolutional Neural Networks*, cornel university, arXiv:1511.08458 [cs.NE], 2015,pp10.
2. Christopher, S.; Christopher, H. *Deep Learning Neural Networks for Land Use Land Cover Mapping*. IGARSS - IEEE International Geoscience and Remote Sensing Symposium, 2018, pp. 3445–3448.
3. *A Beginner's Guide to Segmentation in Satellite Images: Walking through Machine Learning Techniques for Image Segmentation and Applying Them to Satellite Imagery*. <https://www.gsitechnology.com/Beginners-Guide-to-Segmentation-in-Satellite-Images> (Accessed 15-10-2022)
4. Chen, L.; Qianli, Z.; Papandreou, G.; Schroff, F.; Adam, H. *Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation*, Computer Vision – ECCV, 2018, pp 833–851.
5. Darwishe, D.; Mohammad, A.; Chaaban, F. *Developing a Model of Deep Learning by ANNs for Urban Areas Extraction from Remote Sensing Images -Study Area: HomsTartous*, Al-Baath University Journal, V. 43, NO. 7, 2021, PP.11-42.
6. Yunfeng, H.; Zhu, Y.; Yunzhi, Z.; Huimin, Y. *Deep Convolution Neural Network Method for Land Cover Mapping: A Case Study of Qinhuangdao, China*, remote sensing, 2018, pp.16.
7. Shouji, D.; Shihong, D.; Liu, B.; Xiuyuan, Z. *Incorporating DeepLabv3+ and object-based image analysis for semantic segmentation of very high resolution remote sensing images*, International Journal of Digital Earth, 2021, V. 14, NO. 3, 357–378.
8. Ghandorh, H.; Boulila, W.; Masood, S.; Koubaa, A. .; Ahmed, F.; Ahmed, J. *Semantic Segmentation and Edge Detection—Approach to Road Detection in Very High Resolution Satellite Images*, remote sensing, 2022, pp.22.
9. Yongyang, X.; Zhong, X.; Yaxing, F.; Zhanlong, C. *Road Extraction from High-Resolution Remote Sensing Imagery Using Deep Learning*, remote sensing, 2018, pp.16.
10. Volodymyr, M.; Geoffrey, H. *Learning to Detect Roads in High-Resolution Aerial Images*. European conference on computer vision- Springer, 2010, pp.14
11. Yun, W.; Suye, W.; XueBin, H. *Road Extraction using High Resolution Satellite Images based on Receptive Field and Improved Deeplabv3+*, Journal of Physics: Conference Series, 2022, pp.9.
12. Ravishankar, C.; Sridevi, P. *Image Processing and Acquisition using Python*, CRC Press, First edition published 2021, PP.453.
13. Singh, H. *Practical Machine Learning and Image Processing - For Facial Recognition, Object Detection, and Pattern Recognition Using Python*, A Press Springer Science and Business Media Finance Inc, 2019, PP.169.
14. *Machine Learning-Google Colab Tutorial*, 2021. [https://speech.ee.ntu.edu.tw/~hylee/ml/ml2021-course-data/hw/Colab/Google\\_Colab\\_Tutorial.pdf](https://speech.ee.ntu.edu.tw/~hylee/ml/ml2021-course-data/hw/Colab/Google_Colab_Tutorial.pdf) (Accessed 1-8-2022)
15. Heryadi, Y.; Irwansyah, E.; Miranda, E. *The Effect of Resnet Model as Feature Extractor Network to Performance of DeepLabV3 Model for Semantic Satellite Image Segmentation*, IEEE Asia-Pacific Conference on Geoscience, Electronics and Remote Sensing Technology, 2020, PP.4.
16. Rahman, M.; Wang, Y. *Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation*, advances in visual computing,2016,volume10072.

17. *Intersection over Union (IoU) in Object Detection and Segmentation*, 2022.  
<https://learnopencv.com/intersection-over-union-iou-in-object-detection-and-segmentation>  
(Accessed 1-9-2022)