مجلة جامعة دمشق للعلوم الهندسية
المجلد42 العدد (3) – 2026: 326- 336

Damascus University Journal for engineering sciences
Vo 42 number (3) -2026: 326-336

تحسين التقطيع الدلالي باستخدام الشبكات العصبونية العميقة

محمد أنس محمد غسان مخللاتي*[1] رؤوف حمدان[2]

[1*] طالب ماجستير في قسم هندسة الحواسيب والأتمتة–جامعة دمشق.
anas28.mokhallalti@damascusuniversity.edu.sy
[2] مدرس، دكتور، مهندس في قسم هندسة الحواسيب والأتمتة– جامعة دمشق–الإبصار الحاسوبي ومعالجة الصور الرقمية–(raouf.hamdan@damascusuniversity.edu.sy

الملخص:

تعتبر الشبكات العصبونية العميقة (التعلم العميق) إحدى أهم الطرق المستخدمة في حل مشكلات الرؤية الحاسوبية، وتعتبر خوارزميات التقطيع الدلالي من الخوارزميات المتقدمة في الرؤية الحاسوبية التي تقدم أداءً عالياً كونها تعطي تفاصيل دقيقة عند استخدامها بالشكل المناسب، تعتبر خوارزمية الشبكات العصبونية (U_Net) من الطرق جيدة الأداء في مشكلات التقطيع الدلالي إلا أنها تعاني من بعض المشكلات وذلك لعدم قدرتها على التعامل مع مشكلات تجزئة المثيل.

تم في هذه الدراسة تعديل ودمج مجموعة من الخطوات بهدف منح U_Net القدرة على التعامل مع تجزئة المثيل حيث تمت إضافة عمليتين عند المعالجة المسبقة للبيانات وهما تقطيع الصور إلى شرائح وترميزها بحسب الأصناف الموجودة بها باستخدام(one hot encoding) بالإضافة إلى تجزئة الصور إلى أجزاء صغيرة متساوية للتمكن من تدريب النموذج لاحقاً.

كما تم استخدام تابع كلفة خاص عند تدريب النموذج (soft dice-loss) ومقارنة أدائه مع أهم توابع الكلفة الشائع استخدامها عادةً (cross-entropy و mae).

تمت التعديلات المذكورة على عينات من الصورة الجوية لمناطق محددة واختبار قدرة النموذج المعدل على تحديد أنواع المناطق الجغرافية المتضمنة في الصورة الجوية والتي تم تقسيمها إلى خمسة أصناف.

الكلمات المفتاحية: التعلم العميق، التقطيع الدلالي، تجزئة المثيل، الصور الجوية.

تاريخ الإيداع: 2023/9/26
تاريخ القبول: 2023/10/22

http://journal.damascusuniversity.edu.sy/index.php/index/index
ISSN (online): 2789-6854

Enhancing Semantic Segmentation Using Deep Neural Networks

Mhd Anas Makhalalati*[1] Raouf Hamdan[2]

*[1.] Master student in Computer and Automation Engineering Department Damascus University.
anas28.mokhallalati@damascusuniversity.edu.sy

[2.] Lecturer, Dr.Eng in Computer and Automation Engineering Department Damascus University- computer vision and digital image processing.
raouf.hamdan@damascusuniversity.edu.sy

Received: 26/9/2023
Accepted: 22/10/2023

Abstract:

Deep neural networks (deep learning) are considered as one of the most important methods used in solving computer vision problems.

Semantic segmentation algorithms are among the advanced algorithms in computer vision that provide high performance by giving them accurate details when used appropriately.

(U_Net) is a well-performing method in semantic segmentation problems, but it suffers from some problems due to its inability to deal with the instance segmentation problems.

In this study, a modification and integration of a set of steps have been done in order to give U_Net the ability to deal with (instance segmentation), where two operations were added to the data pre-processing stage, namely cutting images into slices and encoding them according to the categories in them, in addition to segmenting images into Small equal parts to be able to train the model later
(patchifying).

Also a special cost function when training the model (soft dice-loss) has been used and compared its performance with the most commonly used cost functions (cross-entropy and mae). The mentioned modifications were made to samples of the aerial images of specific areas and the ability of the modified model to identify the types of geographical areas included in the aerial image were tested, which were divided into five categories.

Keywords: Deep Learning, Semantic Segmentation, Instance Segmentation, Areal images.

المقدمة (Introduction):

أدت زيادة قوة الحوسبة مع تطوير تقنية الإنتشار الخلفي (المعروفة بأشكال مختلفة منذ الستينيات، ولكن لم يتم تطبيقها بشكل كامل حتى الثمانينيات) أدت إلى تجدد الإهتمام بالشبكات العصبونية، فالآن لا نملك أجهزة حوسبة قادرة على تدريب شبكات أعقد فحسب، وإنما لدينا أيضاً التقنيات اللازمة لتدريب شبكات أعمق بكفاءة عالية.

أول الشبكات العصبونية التلافيفية جمعت هذه الأفكار مع نموذج التعرف البصري من أدمغة الثدييات، مما يؤدي لأول مرة إلى شبكات يمكنها التعرف على الصور المعقدة بكفاءة مثل الأرقام المكتوبة بخط اليد والوجوه.

تقوم الشبكات التلافيفية بذلك عن طريق تطبيق نفس "الشبكة الفرعية" لمواقع مختلفة من الصورة وتجميع نتائج هذه الشبكات لتحديد ميزات عالية المستوى(Douwe Osinga,2018,9).

تعتبر الشبكات العصبونية التلافيفية نواة أي شبكة عصبونية مهمتها الأساسية فهم وتحليل الصور أي إضافة القدرة على الرؤية الحاسوبية إلى الشبكة العصبونية. ولعل التقطيع الدلالي من أهم الخوارزميات التي تم تطوريها بناءً على الشبكات التلافيفية.

يتعلق التقسيم الدلالي بتسمية كل بكسل فردي في صورة بفئته التي ينتمي إليها(Gerhard Neuhold et al., 2016).

خوارزمية التقطيع الدلالي هي واحدة من أصناف الخوارزميات التي تهدف إلى استخراج المعلومات من الصورة واستخدامها لإنجاز مهمة محددة(Olaf Ranneberger et al,2016).

تعتبر خوارزمية الـ U-Net من أشهر الخوارزميات المستخدمة في التقطيع الدلالي والتي تم تطويرها بهدف تحليل الصور الطبية، تعتمد في بنيتها على نموذج المرمز–مفكك الترميز Encoder–Decoder Networks وهو ما يمنحها الدقة في تحليل الصور، تعاني هذه الخوارزمية من المحدودية في مشاكل تجزئة المثيل فهي غير قادرة على فصل بيكسلات الصورة التي تنتمي إلى صنف وحيد مما يجعلها غير قادرة على تقديم الحل المناسب عند التعامل مع أكثر من صنف في نفس الصورة والتعرف عليهم في آنٍ معاً.

1 . الدراسات المرجعية (Literature Review):

تم معالجة التقطيع الدلالي في العديد من الدراسات السابقة ونذكر منها: طرح كل من (Olaf Ranneberger, Philipp Fischer, and Thomas Brox,2016) نموذج الشبكة العصبونية U-Net للمرة الأولى وكانت الغاية منها تطبيق التقطيع الدلالي على الصور الطبية فقط(Olaf Ranneberger et al,2016). بينما اعتمد كل من (Mingxing Li et al,2021) على نموذج U-Net لبناء نموذج معدل يقوم بعملية تجزئة التقطيع الدلالي وذلك أيضا في معالجة الصور الطبية حيث قدمت الدراسة نموذجاً معدلاً Res-UNet لتحليل صور الميتاكوندريا. كما قدم كل من (Fang Chen et al,2022) نموذجاً معدلاً Res2-Unet عن U-Net لكشف الأبنية في الصور الجوية. يوضح الشكل(1) طبيعة الصور الجوية وخريطة الميزات المحدودة بعملية الفصل الثنائي لشبكة Res2-Unet:

الشكل (1) صورة جوية مع خريطة الميزات

2. الهدف:

يهدف هذا البحث إلى تحسين خوارزمية التقطيع الدلالي U-Net لمنحها القدرة على تحليل الصور الجوية للمناطق الجغرافية وتمييز محتوياتها (أرض عشبية – أرض ترابية – مباني – مسطحات مائية – طرقات) حيث تم تحقيق تقدم في هذا الخصوص عن طريق إضافة خطوات في التحضير المسبق

http://journal.damascusuniversity.edu.sy/index.php/index/index
ISSN (online): 2789-6854

للصور كتقسيمها إلى شرائح وتجزيئها بالاضافة إلى استخدام تابع كلفة خاص واختيار عدد الطبقات والروابط بينها بشكل يعطي أعلى دقة ممكنة.

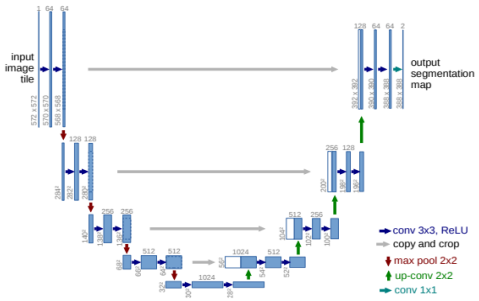
3. خوارزمية U-Net:

تم تدريب شبكة في إعداد نافذة منزلقة للتنبؤ بتسمية فئة كل بكسل من خلال توفير منطقة محلية حول هذا البكسل (Olaf Ranneberger et al,2016).

وهو المعنى المقصود عملياً بالتقطيع الدلالي، طورت هذه الخوارزمية في البداية بشكل خاص لمعالجة الصور الطبية 2016 قبل أن تعمم لاحقاً ليتم تطبيقها في مجالات مختلفة.

يوضح الشكل(2) بنية U-Net (كمثال 32x32 بيكسل كأدنى دقة)، حيث تمثل الصناديق باللون الازرق استجابة خريطة الميزات متعددة القنوات، وعدد القنوات تم توضيحه أعلى كل صندوق، حجم x-y تم توضيحه أسفل الصناديق، والصناديق البيضاء تمثل خريطة الميزات التي تم نسخها أما

فيما يخص الأسهم فهي موضحة أسفل الشكل لقيامها بعمليات متعددة، كما يوضح الشكل(2).



الشكل(2) بنية U-Net

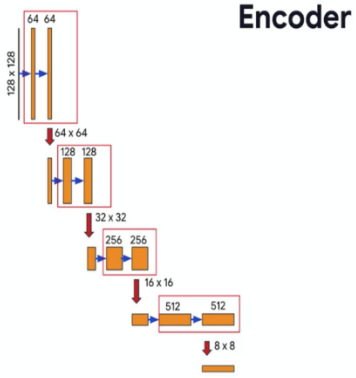
4. المرمز–مفكك الترميز(Encoder-Decoder):

تنقسم آلية العمل في U-Net إلى قسمين أساسيين هما: استخراج الميزات وإعادة بناء خريطة البيكسلات (خريطة الميزات) وهذه الآلية تتميز بأنها عملية متناظرة أي تتساوى فيها طبقات كل مرحلة.

1.4. المرمز(Encoder):

يمثل مرحلة استخراج الميزات من صورة الدخل وبالتالي هو عبارة عن مجموعة من طبقات شبكة عصبونية تلافيفية(CNN) بعد الاستغناء عن الطبقات كاملة الاتصال في نهايتها، يتم في كل طبقة استخراج مستوى معين من الميزات (في الطبقات المبكرة يتم استخراج الميزات البسيطة كالحواف والزوايا بينما في الطبقات المتقدمة يتم استخراج مزايا الغرض المراد تحديده بشكل أدق) وتسمى هذه المرحلة في شبكات U-Net بـ (down sampling).

يوضح الشكل (3) بنية المرمز(Encoder) في شبكات الـ(U-Net).



الشكل(3) المرمز في U-Net

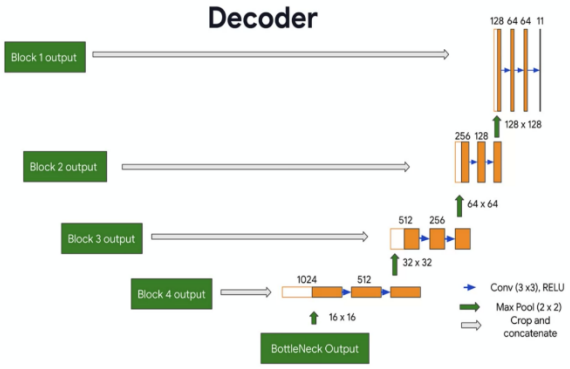
2.4. مفكك الترميز (Decoder):

يتم في هذه المرحلة إعادة بناء الصورة أي استنتاج خريطة البيكسلات (خريطة الميزات) وفي كل خطوة تتم عملية مطابقة مرحلية مع مرحلة الترميز (down sampling) لضمان تساوي أبعاد الصورة بعد إعادة بناء خريطة الميزات.

ISSN (online)
http://journal.damascusuniversity.edu.sy/index.php/index/index

في المرحلة الأخيرة من مفكك الترميز نلاحظ إضافة بعد جديد هو عدد الأصناف في خريطة الميزات.

يوضح الشكل(4) بنية مفكك الترميز (Decoder) في شبكات (U-Net).



الشكل(4) مفكك الترميز في U-Net

5. طرق إعادة البناء في مفكك الترميز:

تتمثل الصعوبة الرئيسية في مهمة التجزئة الدلالية في أن الصور تفقد الدقة المكانية عندما تمر عبر شبكة تلافيفية CNN(بسبب الطبقات ذات الخطوات الأكبر من 1)، لذلك قد ينتهي الأمر بشبكة CNN العادية إلى معرفة أن هناك شخصاً في الصورة، في مكان ما في أسفل يسار الصورة، لكنها لن تكون أكثر دقة من ذلك بكثير(Aurélien Géron, 2019,487) تم تقديم العديد من الحلول لهذه المشكلة ولعل عملية (Up-Sampling) بأنواعها هي الأبرز، على سبيل المثال يمكن استخدام بعض الطرق مثل الجوار الأقرب والتي فيها يتم تكرار لقيمة كل بيكسل من بيكسلات الصورة في الطبقات الأدنى للوصول إلى البعد المطلوب في المرحلة الأعلى، كما يمكن استخدام (Bed-of-Nails) والتي فيها تتم عملية الحشو الصفري للقيم المفقودة وبالتالي الإنتقال إلى بعد أكبر في حجم الصورة مع الحفاظ على قيم البيكسلات، وهناك طريقة أخرى تسمى(Max UnPooling) والتي فيها يتم استخدام الموضع الخاص بالبيكسلات في طبقة الـ (pooling) في مرحلة استخراج الميزات.

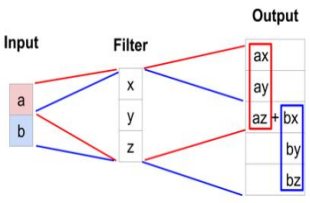
ما ينقص هذه الطرق هو عدم المرونة وفقدانها القدرة على ملاءمة طبيعة الصور التي يتم التدريب عليها فهي تنتقل إلى بعد أكبر بطرق نمطية لا تضمن الحفاظ على دقة الصورة.

تم الاعتماد في هذه الدراسة على طريقة نوعية جديدة في إعادة ترميم الصورة وهي عكس (منقول) الالتفافية.

عكس (منقول) الالتفافية (Transpose Convolution):

وهي عملية معاكسة للجداء النقطي التي تتم في مرحلة استخراج الميزات في الشبكة التلافيفية حيث تتم عملية الجداء بين القيم في شعاع الميزات وقيم المرشح على أن يؤخذ مجموع الجداءات عند حدوث تقاطع (overlapping).

ما يميز هذه الطريقة هو تعديل قيم المرشح أثناء عملية التدريب مما يقلل من ضياع الميزات وفقدان الصورة لوضوحها أثناء إعادة البناء(Up-Sampling). يوضح الشكل (5) طريقة عمل ال(Transpose-Convolution).



الشكل(5) عكس الالتفافية

6. بناء النموذج:

ISSN (online)
http://journal.damascusuniversity.edu.sy/index.php/index/index

اعتماداً على بنية الشبكة العصبونية U-Net كما تم توضيحه وعلى طرق الترميز وفك الترميز، تم بناء نموذج مكون من تسعة طبقات متناظرة بين (Down-Sampling) و (Up-Sampling) واستخدام توابع تفعيل في الطبقات الخفية من نوع (Relu) وطريقة (same padding) في الحشو المرحلي.

أما في مرحلة التدريب فقد تم اعتماد تابع التحسين (adam).

8. خوارزمية التحسين(ADAM):

التي تعني تقدير العزم التكيفي، وتجمع بين أفكار تحسين العزم momentum optimization) و (RMSProp): تماماً مثل تحسين العزم (، فإنه يتتبع متوسط الإنحدار الأسي للتدرجات السابقة، ومثل (RMSProp)، فإنه يتتبع متوسط الإنحدار الأسي للتدرجات التربيعية السابقة– $1.m \leftarrow \beta 1m$ [4].

$(1-\beta 1)\nabla_\theta J(\theta)$

$2.s \leftarrow \beta 2s + (1-\beta 2)\nabla_\theta J(\theta) \otimes J(\theta)$

$3.\widehat{m} \leftarrow \frac{m}{1-\beta_1^t}$

$4.\hat{s} \leftarrow \frac{s}{1-\beta_2^t}$

$5.\theta \leftarrow \theta + \eta\hat{s} \oslash \sqrt{\hat{s}+\varepsilon}$

تمثل t عدد الدورات بدءاً من الرقم 1.

9. تابع الكلفة:

تم في هذه الدراسة استخدام تابع كلفة خاص في حالة التقطيع الدلالي وتجزئة المثيل وهو تابع (Soft Dice-Loss)، أتت التسمية من (Dice Score) وهو معيار يقيس دقة أداء النموذج كنسبة ما تم توقعه بشكل صحيح نسبةً إلى ما توقعه بشكل كامل.

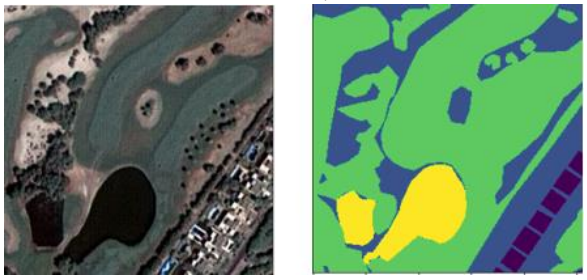
وتعطى معادلة تابع الكلفة (Soft Dice-Loss) كما يلي:

$$J(w,b) = 1 - \frac{\sum y_{pred} \cdot y_{True}}{\sum y_{pred}^2 + \sum y_{True}^2}$$

وللتحقق من دقة أداء النموذج تمت مقارنة أداء تابع الكلفة مع بعض التوابع الشهيرة لحساب الكلفة مثل (cross entropy) و(mean average error) كما سيتم توضيحه.

10. الصور الجوية:

تم استخدام مجموعة من الصور الجوية لمدينة دبي أخذت بتقنية (MBRCS) وهي عبارة عن 72 صورة تمت عملية تجزئتها يدوياً وتصنيف محتواها إلى الاصناف التالية: بناء–أرض عشبية–مسطحات مائية–طرق معبدة–أبنية. يوضح الشكل(6) عينة من الصور الجوية المستخدمة في عملية التدريب:



الشكل(6) الصور الجوية

11. مراحل الخوارزمية المعدلة:

الدخل: هو مجموعة من الصور الجوية تمت عملية تقسيمها إلى شرائح(slicing) وترميز كل صنف فيها ومن ثم استخدام الصورة المجزئة لتدريب النموذج على عدد أكبر من الصور المتاحة، أبعاد صورة التدريب بعد هذه الخطوات هو 3x256x256 وعدد الصور الإجمالي المستخدم في التدريب هو 1000 عينة.

النموذج: هو بنية الشبكة U-Net.

تابع التحسين: (ADAM Optimizer).

تابع الكلفة: تم استخدام تابع الكلفة الخاص (Soft Dice-Loss) ومقارنة النتائج مع توابع كلفة أخرى.

معايير تقييم الاداء: تم الاعتماد على الدقة والخطأ كمعيارين أساسيين في كل مرحلة من مراحل الاختبار، حيث تم الإجراء التجريبي على أربعة مراحل.

12. الاختبار التجريبي ونتائج كل مرحلة:

تمت عملية الاختبار على أربعة مراحل، حيث تم في كل مرحلة قياس الدقة (Accuracy) والكلفة (Error) لمقارنة أي من التعديلات يعطي نتيجة أفضل، تم استخدام لغة البرمجة بايثون

ISSN (online)
http://journal.damascusuniversity.edu.sy/index.php/index/index

وإطار عمل (Tensorflow) وواجهة التخاطب البرمجية (Keras) لبرمجة النموذج، كما تم استخدام (Tensorboard) لقياس الأداء ورسم المنحنيات.

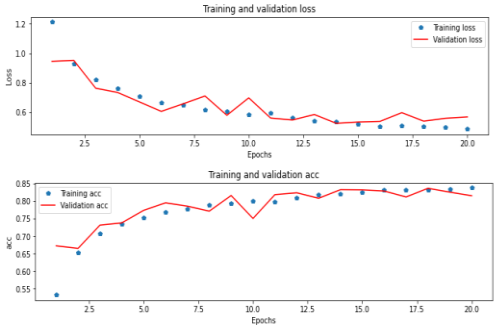
وكانت الاختبارات كمايلي:

1.12. الاختبار الأول:

باستخدام تابع الكلفة (Cross-Entropy) وبتقسيم المعطيات إلى 80% عينات (تدريب وتقييم) و 2% عينات اختبار كانت النتائج كالتالي:

دقة التدريب 0.837 ودقة التقييم 0.8147 بينما كانت خسارة التدريب 0.484 وخسارة التقييم 0.566.

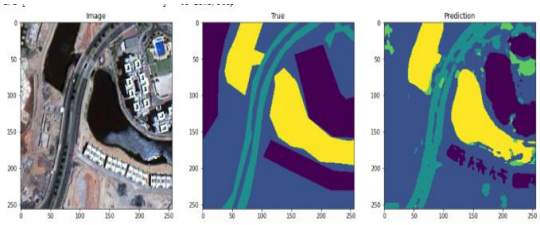
يوضح الشكل(7) منحنيات الدقة والخطأ في الاختبار الأول:



الشكل (7) منحنيات الدقة والخطأ – الاختبار الاول

يعاني الاختبار الأول كما توضح منحنيات تقييم الأداء من بعض الموائمة الزائدة (overfitting) وهو مايجعل النموذج يواجه بعض مشاكل التعميم.

يوضح الشكل(8) مقارنة نتائج الاختبار الأول:



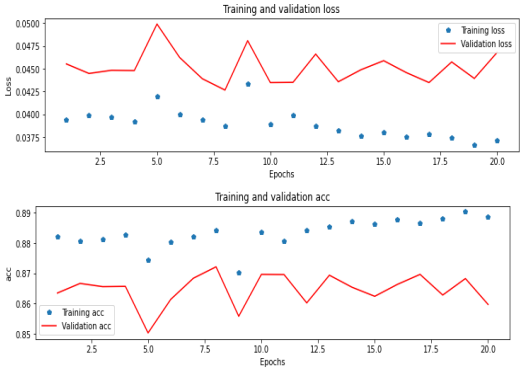
الشكل (8) نتائج الاختبار الأول

2.12. الاختبار الثاني:

باستخدام تابع الكلفة (Cross-Entropy) ولكن بالاستغناء عن عينات التقييم وهذا لجعل عينات التدريب اكثر وبالتالي التخلص من مشكلة الموائمة الزائدة للنموذج كانت النتائج كالتالي:

دقة التدريب: 0.888 دقة الاختبار: 0.85 بينما كانت خسارة التدريب:0.037 وخسارة الاختبار: 0.0468.

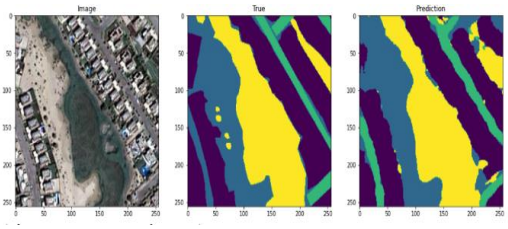
يوضح الشكل(9) منحنيات الدقة والخطأ في الاختبار الثاني:



الشكل (9) منحنيات الدقة والخطأ – الاختبار الثاني

على الرغم من التخلص من مشكلة الموائمة الزائدة في هذا الاختبار إلا أنه يوجد بعض الاهتزاز او الاضطراب في عملية التعلم خلال دورات التدريب.

يوضح الشكل(10) مقارنة نتائج الاختبار الثاني:



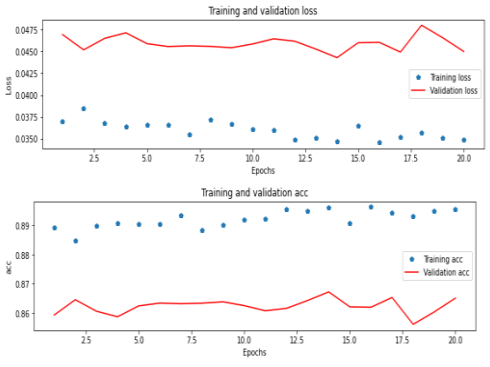
الشكل (10) نتائج الاختبار الثاني

3.12. الاختبار الثالث:

باستخدام تابع الكلفة (Mean Average Error) وبتقسيم المعطيات إلى 80% لتدريب والتقييم و20% للاختبار كانت النتائج كالتالي:

دقة التدريب: 0.895 ودقة التقييم: 0.865 بينما كانت خسارة التدريب: 0.034 وخسارة التقييم:0.045.

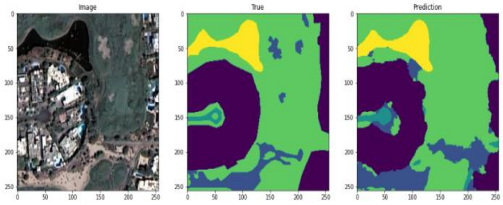
يوضح الشكل(11) منحنيات الدقة والخطأ في الاختبار الثالث:



الشكل(11) منحنيات الدقة والخطأ – الاختبار الثالث

على الرغم من التخلص من الاهتزازات في الاختبار الثاني إلا أنه هناك تباين واضح بين مرحلتي التدريب والاختبار في كل من منحنيات الدقة والخطأ.

يوضح الشكل(12) مقارنة نتائج الاختبار الثالث:



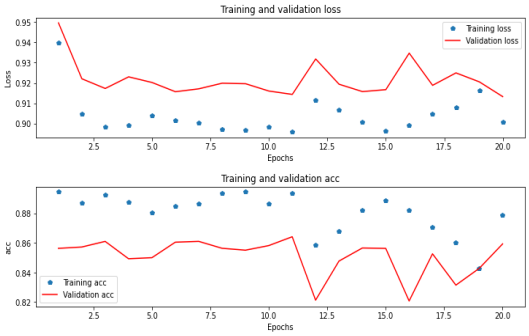
الشكل (12) نتائج الاختبار الثالث

4.12. الاختبار الرابع:

باستخدام تابع الكلفة الخاص (Soft Dice-Loss) وبتقسيم المعطيات إلى 80% للتدريب والتقييم و20% للاختبار، كانت النتائج كالتالي:

دقة التدريب: 0.878، دقة التقييم: 0.859 بينما كانت خسارة التدريب: 0.91 وخسارة التقييم: 0.913.

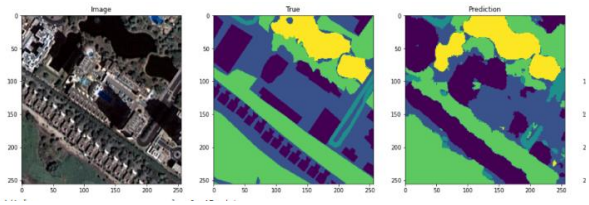
يوضح الشكل(13) منحنيات الدقة والخطأ في الاختبار الرابع:



الشكل(13) منحنيات الدقة والخطأ – الاختبار الرابع

من الملاحظ في هذا الاختبار أنه تم التخلص من الموائمة الزائدة والاهتزازات الكبيرة بنسبة مقبولة وكذلك من التباين بين مرحلتي التدريب والاختبار.

يوضح الشكل(14) مقارنة نتائج الاختبار الرابع:



الشكل (14) نتائج الاختبار الرابع

يمكن تلخيص نتائج الاختبارات الاربعة في الجدول التالي:

الجدول(1) نتائج الاختبارات

	Acc	Loss	Val.Acc	Val.Loss
Test1	0.837	0.484	0.8147	0.566
Test2	0.888	0.037	0.85	0.0468
Test3	0.895	0.034	0.865	0.045
Test4	0.878	0.91	0.859	0.913

نلاحظ من جدول النتائج السابق أن الاختبار الرابع أدى إلى الوصول إلى أعلى دقة ممكنة للتدريب والتقييم مقارنة بماسبقه من الاختبارات الثلاثة السابقة ويعود هذا التحسن إلى استخدام تابع الكلفة الخاص المقترح في هذا البحث مقارنةً باستخدام بعض توابع الكلفة الشهيرة، وكما تم ذكره سابقاً فإنه تم التخلص من بعض مشكلات النموذج كالموائمة الزائدة والتباين بين منحنيات التدريب والاختبار والوصول إلى دقة مقبولة بالنسبة لمشكلة البحث.

13.الاستنتاجات والتوصيات:

قدم هذا البحث تحسيناً على خوارزمية U-Net الخاصة بالتقطيع الدلالي حيث تم الانتقال من المحدودية في الفصل الثنائي كما تظهر الدراسات المرجعية السابقة إلى فصل متعدد عن طريق استخدام التقسيم إلى شرائح وتجزئة الصور في مرحلة المعالجة المسبقة للبيانات، وكذلك استخدام تقنية خاصة أثناء عملية إعادة بناء خريطة الميزات في مرحلة مفكك الترميز وهي عكس الالتفافية (Transpose Convolution) واستخدام تابع كلفة خاص عند تدريب النموذج ومقارنة أدائه بعد هذه التعديلات مع بعض توابع الكلفة الشهيرة. يمكن العمل مستقبلا على تعميم النموذج بشكل أكبر وتحسين دقة الاختبار والتقيم.

التمويل: هذا البحث ممول من جامعة دمشق وفق رقم التمويل (501100020595).

ISSN (online)
http://journal.damascusuniversity.edu.sy/index.php/index/index

References:

1- Aurélien Géron, 2019- Hands-on Machine Learning with Scikit-Learn, Keras and Tensorflow Concepts, Tools and Techniques to Build Intelligent Systems (3d edition).

2- Douwe Osinga, 2018- Deep Learning Cookbook: Practical Recipes to Get Started Quickly.

3-Fang chen, Ning Wang, Bo Yu, Lei wang, 2022- Res2-Unet, a new Deep Learning, architecture for building detection from high spatial resolution images. IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

3-Gerhard Neuhold, BSc, 2016 - Semantic Segmentation with Deep Neural Networks. Graz University of Technology.

4-Ilyes Batatia, 2020 - A DEEP LEARNING METHOD WITH CRF FOR INSTANCE SEGMENTATION OF METAL-ORGANIC FRAMEWORK SINSCANNING ELECTRON MICROSCOPY IMAGES, University Paris Saclay, ENS Paris-Saclay. Mingxing Li, Chang Chen, Xiaoyu Liu, Wei Huan, Yueyi Zhang, Zhiwei Xiongy,2022- Advanced Deep Networks For 3D Mitochondria Instance Segmentation, UNIVERSITY OF SCINCE AND TECHNOLOGY OF CHINA.

5- Olaf Ranneberger, Philipp Fischer, and Thomas Brox, 2016-U-Net: Convolutional Networks for Biomedical Image Segmentation. Computer Science Department and BIOSS Centre for Biological Signaling Studies, University of Freiburg, Germany.

6-Tsung-Yi Lin1,2, Piotr Dollar´ 1, Ross Girshick1, Kaiming He1, Bharath Hariharan1, and Serge Belongie2, 2017- Feature Pyramid Networks for Object Detection.