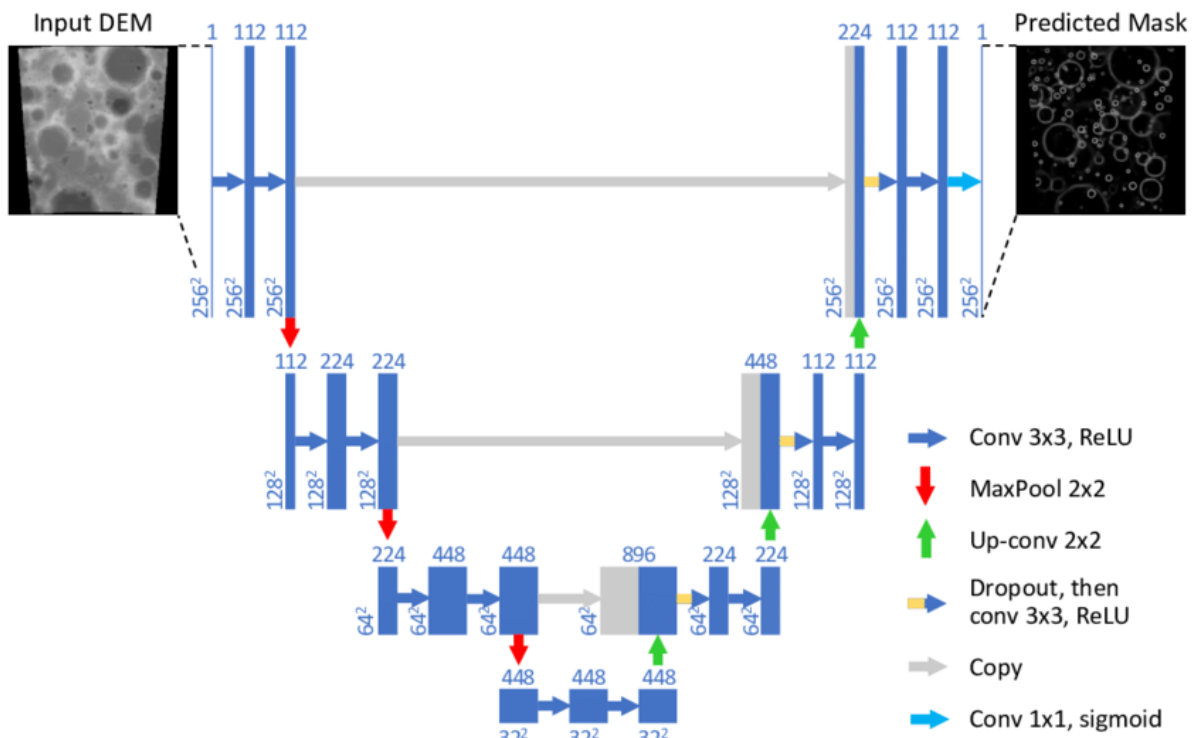


البيئة والأدوات

1. Python 3.8
2. TensorFlow 2.10
3. PyTorch 1.13
4. NumPy
5. Pandas
6. JSON

بنية الشبكة العصبونية المستخدمة UNet [1]



تم استخدام البنية ذاتها في تدريب كلا النموذجين (Fire Detection & Depth Estimation)

تم أيضا تدريب عدة بنى لشبكات عصبونية ذات عدد بارامترات أكبر مثل Mask RCNN التي تستخدم عادة في مسائل التجزيء [2]، ولكن الإمكانيات الحسابية المتوافرة تجعل من الصعب بناء نموذج يعمل بالزمن الحقيقي لتجزيء صورة بحثاً عن النار بالتوازي مع توقع العمق من أجل كل بكسل في الصورة من خلال نموذج آخر حيث أن المواصفات الحسابية المتوافرة لتدريب وتجريب النماذج تتمثل في معالج رسومات بالموصفات التالية:

GPU: Nvidia GEFORCE RTX 3050

VRAM: 4 GB

CUDA Cores Count: 3840

في حين أن المساحة الذاكرة المطلوبة لتخزين كل من النماذج السابقة تتجاوز الـ 3 GB مما يعني استحالة استخدام نموذجين في نفس اللحظة دون التفريط بالكثير من الدقة في أحدهما

قواعد البيانات المستخدمة:

مسألة تجزئ الصورة Image Segmentation:

Fire Segmentation Image Dataset (FSID): قاعدة بيانات تم جمعها عبر الانترنت واستضافتها على موقع KAGGLE لاستخدامها من قبل الباحثين في مسائل تجزئة الصور. [3]

Fire-Detection-UAV-Aerial-Image-Classification-Segmentation-Unmanned Aerial Vehicle (FDUAV): قاعدة بيانات تحتوي على صور مبنية على مشاهد video scenes تم التقاطها بواسطة طائرات مسيرة Drones، جميع المشاهد هي مشاهد خارجية Outdoor Imagery. [4]

مسألة تقدير العمق Depth Estimation:

NYU Depth: هي مجموعة بيانات تتألف من تسلسلات فيديو من مجموعة متنوعة من المشاهد الداخلية التي تم تسجيلها باستخدام كاميرات RGB وعمق من Microsoft Kinect. تتضمن المجموعة عدة مكونات بما في ذلك مجموعة فرعية معلمة من بيانات الفيديو ترافقها تصنيفات متعددة وكثيفة، وبيانات RGB والعمق الخام وبيانات المعايرة المقدمة من Kinect، ولوحة أدوات بها وظائف مفيدة لمعالجة البيانات والتصنيفات.

يمكن استخدام هذه المجموعة لمهام مثل تقسيم المشاهد الداخلية والتعرف على الكائنات. إنها توفر مصدر غني للبيانات للباحثين العاملين على تطبيقات رؤية الحاسوب والتعلم الآلي. [5]

KITTI

تم إصدار مجموعة البيانات KITTI لأول مرة في عام 2012 من قبل جيجر وآخرون. تم جمعها باستخدام سيارة فولكس واجن باسات عند القيادة في وحول كارلسروه في ألمانيا. تم تجهيز السيارة بكاميرات (رمادية وملونة) وليدار ونظام تحديد المواقع العالمي (GPS) ووحدة القياس الداخلي (IMU) للسماح بجمع البيانات متعددة الوسائط.

تستفيد مجموعة بيانات KITTI Vision Benchmark Suite من منصة قيادة ذاتية تسمى Anniway لتطوير مقاييس تحديثية جديدة ومستحدثة في رؤية الحاسوب العالمية الحقيقية التحديات. تشمل المهام المهمة الاستريو، وتدفق الضوء البصري، والتصوير البصري السريع، وكشف الكائنات ثلاثية الأبعاد وتتبعها.

تعد هذه المجموعة من بين أكثر المجموعات شهرة للاستخدام في الروبوتات المحمولة والقيادة الذاتية. إنها توفر مصدرًا غنيًا للبيانات للباحثين العاملين على تطبيقات رؤية الحاسوب والتعلم الآلي. [6]

التجارب:

تجزئ الصور:

تم تجريب بناء النموذج من الصفر بناءً على قاعدة البيانات FSID وكانت النتائج مقبولة عند النظر إلى معيار Dice Coefficient والذي يقيس معدل التطابق بين خرائط التجزئ التي تم توقعها النموذج المدرب مع قاعدة الحقيقة Ground Truth في الDataset (0.8 on Dice Coefficient Metric)

تم تجريب تدريب النموذج بشكل مسبق على قاعدة البيانات الثانية FDUAV ثم على الأولى لتحسين النتائج لكن تمت ملاحظة أن قدرة النموذج على تجزئ النار تراجعت (0.73 on Dice Coefficient) ويعود ذلك إلى الفرق في التوزيع Distribution Difference بين قاعدتي البيانات حيث أن إحداها صور لمشاهد قليلة في الغابات بينما الأخرى تحوي مشاهد متنوعة بدقة أقل.

تقدير العمق:

تم تدريب نموذج من أجل كل قاعدة بيانات وتمت ملاحظة أن النموذج الذي تم تدريبه على قاعدة بيانات NYU Depth أظهر نتائج أفضل مقارنة بنظيره الذي تم تدريبه على KITTI حيث حقق النموذج الأول معدل خطأ تربيعي Mean Squared Error مقداره 0.9 في حين حقق الأخير معدل خطأ تربيعي قدره 2.4

References:

1. Li, X., Chen, H., Qi, X., Dou, Q., Fu, C. W., & Heng, P. A. (2018). H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE transactions on medical imaging*, 37(12), 2663-2674.
2. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).
3. [Fire Segmentation Image Dataset | Kaggle](#)
4. Shamsoshoara, A. (2020). Fire-Detection-UAV-Aerial-Image-Classification-Segmentation-UnmannedAerialVehicle. GitHub repository, GitHub.
5. Couprie, C., Farabet, C., Najman, L., & LeCun, Y. (2013). Indoor semantic segmentation using depth information. arXiv preprint arXiv:1301.3572.
6. Geiger, A., Lenz, P., Stiller, C., & Urtasun, R. (2013). Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, 32(11), 1231-1237.