

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

المدرسة العليا للإعلام الآلي - 08 ماي 1945 - بسبدي بلعباس
Ecole Supérieure en Informatique -08 Mai 1945- Sidi Bel Abbès



Graduation Thesis

To obtain the diploma of **State Engineering**

Field of Study : **Computer Science**

Specialty : **Artificial Intelligence & Data Science (IASD)**

Thème

**On-Board Learning-Based Scheduler with Vision-Based
Look-Ahead for Agile Earth Observation**

Presented by :

Mlle Sanaa Bensaid

Defended on : 07/07/2026 In front of the jury composed of :

Mme ALLAL Lamia

Président

Mr Meziane Iftene

Supervisor

Mr Chaib Souleyman

Co-Supervisor

Mr Mohammed El Amin Larabi

Co-Supervisor

Mme NAIB Rabiaa

Jury

Academic Year : 2025 / 2026

Abstract

Earth observation satellites in low Earth orbit are scheduled on the ground six to twenty-four hours before each orbital pass, with no real-time awareness of cloud conditions at acquisition time. Cloud contamination renders sixty to eighty percent of downlinked imagery unusable, representing a fundamental waste of onboard storage, downlink bandwidth, and orbital opportunity. For ALSAT-2A, operated by ASAL from a sun-synchronous orbit at 670 km, this problem is acute : each target window lasts only thirty to ninety seconds, and ground-computed schedules predict cloud-free conditions with at most sixty to seventy percent accuracy.

This thesis presents ALSAT-EO-1, an onboard scheduling system integrating two tightly coupled innovations. A lightweight Convolutional Neural Network (CNN) processes forward-looking scout camera imagery to estimate cloud probabilities seventy to one hundred forty seconds before acquisition. A Proximal Policy Optimisation (PPO) reinforcement learning agent uses these real-time estimates to revise the imaging schedule at each decision point. The decision problem is formalised as a Semi-Markov Decision Process (SMDP), correctly accounting for variable-duration intervals between target windows. The agent is pre-initialised through behavioural cloning from a greedy oracle and refined through a four-phase curriculum of progressively increasing difficulty, trained over approximately 125,000 environment steps (320 logged episodes) across 45 global target locations.

The training trajectory exhibits a characteristic staged progression : the reward stabilises during static-target-only phases, then rises when dynamic emergency events (wildfires, floods, seismic events within Algeria) are introduced as an additional reward source, reaching a stable plateau of 105.0 ± 11.6 mean episode reward. In a high-fidelity Basilisk simulation environment configured to ALSAT-2A parameters, and in the final evaluation regime with 75% cloud cover and two events per hour, the system achieves 96.0% cloud-free imaging rate and $95.4\% \pm 2.7\%$ dynamic event success rate (mean of the final 50 evaluation episodes, 320-episode training log), with the battery state of charge never dropping below the safety threshold across the full training and evaluation run. Deployment characterisation on the Intel Myriad X VPU projects a 58 ms cognitive cycle and 0.034 J per decision, well within ALSAT-class constraints.

Keywords : ALSAT-2A ; Earth observation ; onboard autonomy ; reinforcement learning ; Semi-Markov Decision Process ; cloud detection ; agile satellite scheduling ; Basilisk ; edge AI ; curriculum learning ; behavioural cloning ; SHAP ; Intel Myriad X.

Résumé

Les satellites d'observation de la Terre en orbite basse sont planifiés au sol six à vingt-quatre heures avant chaque passage orbital, sans connaissance en temps réel des conditions de nébulosité au moment de l'acquisition. La contamination nuageuse rend inutilisable de soixante à quatre-vingts pour cent des images reçues au sol, représentant un gaspillage fondamental de la capacité de stockage embarquée, de la bande passante de téléchargement et des opportunités orbitales. Pour ALSAT-2A, opéré par l'ASAL depuis une orbite héliosynchrone à 670 km, ce problème est particulièrement aigu : chaque fenêtre de visibilité ne dure que trente à quatre-vingt-dix secondes, et les planifications calculées au sol prévoient des conditions sans nuages avec une précision d'au plus soixante à soixante-dix pour cent.

Cette thèse présente ALSAT-EO-1, un système de planification embarqué intégrant deux innovations étroitement couplées. Un réseau de neurones convolutif (CNN) léger traite des images provenant d'une caméra de reconnaissance orientée vers l'avant pour estimer les probabilités de nuages soixante-dix à cent quarante secondes avant l'acquisition. Un agent d'apprentissage par renforcement basé sur l'optimisation de politique proximale (PPO) utilise ces estimations en temps réel pour réviser le programme d'imagerie à chaque point de décision. Le problème de décision est formalisé comme un Processus de Décision Semi-Markovien (SMDP), prenant correctement en compte les intervalles de durée variable entre les fenêtres de visibilité des cibles. L'agent est pré-initialisé par clonage comportemental depuis un oracle glouton et affiné à travers un curriculum à quatre phases de difficulté progressive, entraîné sur environ 1 000 000 pas d'environnement répartis sur 45 sites cibles mondiaux.

La trajectoire d'entraînement présente une progression par étapes caractéristique : la récompense se stabilise pendant les phases de cibles statiques seules, puis augmente lorsque des événements d'urgence dynamiques sont introduits comme source de récompense supplémentaire, atteignant un plateau stable de $10,47 \pm 2,21$ de récompense moyenne par épisode. Dans le régime d'évaluation final avec 75% de couverture nuageuse et deux événements par heure, dans un environnement de simulation Basilisk haute fidélité configuré selon les paramètres d'ALSAT-2A, le système atteint $95,5\% \pm 2,4\%$ de taux d'imagerie sans nuages et $98,9\% \pm 1,4\%$ de taux de succès sur les événements dynamiques (moyenne sur 150 épisodes d'évaluation), avec zéro événement d'épuisement de batterie sur 200 épisodes d'évaluation.

Mots-clés : ALSAT-2A ; observation de la Terre ; autonomie embarquée ; apprentissage par renforcement ; SMDP ; détection de nuages ; planification de satellite agile ; curriculum ; clonage comportemental ; SHAP ; Intel Myriad X.

ملخص / Résumé (Arabe)

مَجْدَلُ تَعْلِي مُدَجِّجٍ لِلأَقْمَارِ الاصطناعية للاستشعار عن بُعْد: ALSAT-EO-1

تُجَدِّدُ أَقْمَارُ الاصطناعية للاستشعار عن بُعْدٍ في المدار الأرضي المنخفض من الأرض قبل كل مرور مداري بست إلى أربع وعشرين ساعة، دون أي إدراك فوري للغطاء السحابي لحظة الالتقاط. يُفسد التلوث السحابي ما بين ستين وثمانين بالمئة من الصور المستقبلية، مما يُمثل هدراً جوهرياً في طاقة التخزين المتاحة، وعرض نطاق الإرسال، وفرص الرصد. بالنسبة للقمر الجزائري ALSAT-2A الذي تشغله الوكالة الفضائية الجزائرية من مدار متزامن مع الشمس على ارتفاع 670 كيلومتراً، تبقى نافذة رؤية كل هدف بين ثلاثين وتسعين ثانية فقط.

تُقَدِّمُ هذه الأطروحة نظام ALSAT-EO-1 وهو منظومة جدولة ذكية مُدمجة تجمع بين شبكة عصبية تلافيفية خفيفة الوزن تعالج صور كاميرا استطلاع أمامية لتقدير احتمالية وجود السحب قبل سبعين إلى مئة وأربعين ثانية من الالتقاط، وعامل تعلم معزز مبني على خوارزمية PPO يستخدم هذه التقديرات الفورية لتعديل جدول التصوير. تم صياغة مسألة القرار كعملية قرار شبه ماركوفية SMDP تأخذ بعين الاعتبار الفترات الزمنية المتغيرة بين نوافذ رؤية الأهداف. تدرَّب العامل عبر منهج متدرج من أربع مراحل متصاعدة الصعوبة يشمل مليون خطوة تقريباً، موزعة على خمسة وأربعين موقع هدف عالمي.

تكشف مسيرة التدريب عن تقدم متدرج بالمراحل: تستقر المكافأة خلال مراحل الأهداف الثابتة فقط، ثم ترتفع عند إدخال الأحداث الطارئة الديناميكية كالحرائق والفيضانات والزلازل داخل الجزائر بوصفها مصدراً إضافياً للمكافأة، لتصل إلى هضبة مستقرة بمتوسط 11.6 ± 105.0 . في نظام التقييم النهائي، وباستخدام بيئة محاكاة Basilisk عالية الدقة مُهيأة وفق معايير ALSAT-2A، يحقق النظام نسبة تصوير خالٍ من السحب 96.0%، ونسبة نجاح في الأحداث الديناميكية $95.4\% \pm 2.7\%$ (بمتوسط آخر 50 حلقة من سجل تدريب من 320 حلقة)، مع بقاء شحن البطارية دائماً فوق حد الأمان.

الكلمات المفتاحية: ALSAT-2A الاستشعار عن بُعْد؛ الذكاء الاصطناعي المُدَجِّج؛ التعلم المعزز؛ عملية القرار شبه الماركوفية؛ كشف السحب؛ جدولة الأقمار الرشيقية؛ X. Myriad Intel