

الجمهورية الشعبية الديمقراطية الجزائرية
People's Democratic Republic of Algeria
وزارة التعليم العالي والبحث العلمي
Ministry of Higher Education and Scientific Research
المدرسة العليا للإعلام الآلي 8 ماي 5491 - سيدي بلعباس
Higher School of Computer Science
8 Mai 1945 - Sidi Bel Abbas



Thesis

To obtain the diploma of **Engineering Degree**
Field of Study: **Computer Science**
Specialization: **Ingénierie des Systèmes Informatiques (ISI)**

Theme

AI Driven Fault Detection in Solar Farm Systems

Presented by
Khelif Ibtissem Lidia
Submission Date: **08th July, 2025**
In front of the jury composed of

Dr. Nassima DIF
Pr. Sidi Mohammed BENSLIMANE
Dr. Oussama SERHANE
Mr. Imene NEGGAZ
Mr. Mohamed KECHAR

President of the Jury
Supervisor
Supervisor
Examiner
Incubator Representative

Academic Year: 2024/2025

Abstract

Solar photovoltaic (PV) systems play a pivotal role in the global transition to renewable energy, especially in sun-rich regions like Algeria. Yet their performance is often restrained by fluctuating output driven by changing weather component aging, and faults such as partial shading or inverter failures. Conventional detection and forecasting relying on manual checks and fixed thresholds struggle to keep pace on utility-scale installations, resulting in avoidable energy losses and steep maintenance bills.

That's where deep learning steps in. By feeding continuous streams of NASA's POWER data hourly meteorological records spanning 2001-2025 into advanced neural architectures, we can both forecast tomorrow's solar yield and flag emerging faults before they spiral. In this thesis, we put four contenders through their paces: LSTM, bidirectional LSTM, Transformer, and the Temporal Fusion Transformer (TFT) after cleaning, scaling, and enriching the raw inputs with cloud-cover ratios and weather tags.

We measure success with MAE, RMSE, and R^2 , and find that Transformer-based models, specifically TFT, in particular cut forecast errors by nearly half compared to vanilla LSTM, thanks to their knack for capturing intricate temporal patterns. Even better, deviations between predicted and actual output become reliable early-warning signals for system anomalies, empowering maintenance teams to act preemptively.

Beyond headline numbers, we tackle the real-world headaches: patchy local data, hefty compute requirements, and the tight timeline of live deployment. We spotlight gaps like the need for Algeria-tuned models, uniform benchmarking protocols, and seamless hooks into on-site SCADA and sketch out next steps: real-time inference engines, ensemble blends of these architectures, and fusion with local PV measurements.

By zeroing in on Algeria's unique climate and harnessing state-of-the-art deep learning, this work doesn't just compare models; it charts a path toward truly autonomous, scalable solar farm management, promising greener power, fewer outages, and a sunnier future.

Keywords— artificial intelligence, solar photovoltaic systems, deep learning, LSTM, bidirectional LSTM, Transformer, Temporal Fusion Transformer, PV power forecasting, fault detection, renewable energy, Algeria

Résumé

Les systèmes photovoltaïques (PV) occupent une place clé dans la transition mondiale vers les énergies renouvelables, notamment dans des régions très ensoleillées comme l'Algérie. Cependant, leur rendement demeure souvent limité par la variabilité de la production liée aux conditions météorologiques, par le vieillissement des composants et par des défauts tels que l'ombrage partiel ou les pannes d'onduleur. Les méthodes classiques de détection et de prévision reposant sur des contrôles manuels et des seuils fixes peinent à suivre le rythme des installations à grande échelle, entraînant des pertes d'énergie évitables et des coûts de maintenance élevés.

Le deep learning apporte une réponse puissante. En ingérant en continu les données POWER de la NASA relevés météorologiques horaires couvrant la période 2001-2025 dans des architectures neuronales avancées, on peut à la fois anticiper la production solaire du lendemain et détecter les anomalies naissantes avant qu'elles ne s'aggravent. Dans cette thèse, nous confrontons quatre modèles de pointe LSTM, LSTM bidirectionnel, Transformer et Temporal Fusion Transformer (TFT) après avoir nettoyé, normalisé et enrichi les données brutes avec des taux de couverture nuageuse et des étiquettes météorologiques.

Nous évaluons la performance via le MAE, le RMSE et le R^2 , et constatons que les modèles basés sur Transformer en particulier le TFT réduisent les erreurs de prévision de près de moitié par rapport au LSTM standard, grâce à leur capacité à capturer des schémas temporels complexes. Mieux encore, les écarts entre les valeurs prévues et mesurées deviennent des indicateurs d'alerte fiables pour repérer les anomalies, permettant aux équipes de maintenance d'intervenir de manière proactive.

Au-delà des chiffres, nous abordons les défis concrets : données locales parfois incomplètes, besoins computationnels importants et contraintes de mise en œuvre en temps réel. Nous identifions les lacunes, la nécessité de modèles adaptés au climat algérien, protocoles de benchmarking unifiés et intégration fluide dans les SCADA existants et esquissons les prochaines étapes : moteurs d'inférence en temps réel, approches d'ensembles hybrides et fusion avec les mesures PV locales.

En ciblant le climat spécifique de l'Algérie et en exploitant les dernières avancées du deep learning, ce travail ne se limite pas à comparer des architectures : il trace les voies d'une gestion pleinement autonome et évolutive des fermes solaires, promettant une énergie plus verte, moins d'interruptions et un avenir plus ensoleillé.

Mots-clés—intelligence artificielle, systèmes photovoltaïques, deep learning, LSTM, LSTM bidirectionnel, Transformer, Temporal Fusion Transformer, prévision PV, détection de défauts, énergies renouvelables, Algérie.

ملخص

تلعب أنظمة الطاقة الشمسية الكهروضوئية (PV) دوراً محورياً في الانتقال العالمي إلى الطاقة المتجددة، وخصوصاً في المناطق المشمسة مثل الجزائر. ورغم ذلك، غالباً ما يتأثر أداؤها بتقلب الإنتاج الناتج عن تغير الأحوال الجوية، ويتقدم عمر المكونات، وظهور أعطال مثل التظليل الجزئي أو فشل العواكس. تعجز أساليب الكشف والتنبؤ التقليدية – المعتمدة على الفحوصات اليدوية والعتبات الثابتة – عن مواكبة متطلبات المحطات واسعة النطاق، مما يؤدي إلى فقدان طاقة يمكن تجنبها وتكاليف صيانة مرتفعة.

يخترط التعلم العميق هنا كحل فعال. عن طريق حقن البيانات المستمرة من مجموعة POWER التابعة لوكالة ناسا – سجلات جوية كل ساعة للفترة 2001-2025 – في معماريات عصبية متقدمة، يمكن تنبؤ إنتاج الغد الشمسي والكشف المبكر عن الأعطال قبل تفاقمها. في هذه الأطروحة، نختبر أربعة نماذج متطورة – LSTM، LSTM ثنائي الاتجاه، Transformer، Temporal Fusion Transformer (TFT) – بعد تنقية البيانات، وموازنتها، وإثرائها بنسبة الغيوم وتسميات الحالة الجوية. نقيس الأداء باستخدام متوسط الخطأ المطلق (MAE) وجذر متوسط مربع الخطأ (RMSE) ومعامل التحديد (R^2)، ونكتشف أن النماذج القائمة على Transformer – لا سيما TFT – تقلص أخطاء التنبؤ بنحو النصف مقارنة بـ LSTM القياسي، بفضل قدرتها على التقاط أنماط زمنية معقدة. والأفضل من ذلك، أن الفجوة بين القيم المتوقعة والفعلية تتحول إلى مؤشرات إنذار مبكر موثوقة لعزل الشذوذات، مما يمكن فرق الصيانة من التدخل استباقياً.

نخوض أيضاً في التحديات الميدانية: تدفق البيانات المحلية المتفاوت، ومتطلبات الحوسبة العالية، وضيق الجدول الزمني للنشر في الزمن الفعلي. نسلط الضوء على الثغرات – كالحاجة إلى نماذج مخصصة للمناخ الجزائري، وبروتوكولات تقييم موحدة، وتكامل سلس مع أنظمة SCADA القائمة – ونحدد الخطوات القادمة: محركات استدلال في الوقت الحقيقي، وأساليب التجميع (ensemble) المهيجنة، ودمجها مع قياسات PV المحلية.

من خلال التركيز على المناخ الجزائري الفريد والاستفادة من أحدث أساليب التعلم العميق، لا يكفي هذا العمل بمقارنة النماذج فحسب؛ بل يرسم خارطة طريق نحو إدارة ذاتية شاملة وقابلة للتطوير لمزارع الطاقة الشمسية، واعدة بطاقة أنظف وانقطاعات أقل ومستقبل أكثر إشراقاً.

الكلمات المفتاحية: --- الذكاء الاصطناعي؛ أنظمة الطاقة الشمسية الكهروضوئية؛ التعلم العميق؛ LSTM ثنائي الاتجاه؛ Temporal Fusion Transformer؛ Transformer؛ تنبؤ إنتاج PV؛ اكتشاف الأعطال؛ الطاقة المتجددة؛ الجزائر