الجمهورية الشعبية الديمقراطية الجزائرية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للإعلام الآلي 08٠ ماي 5491٠ بسيدي بلعباس
École Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbès

THESIS

To obtain the diploma of **Engineering**
Field: **Computer Science**
Specialty: **Artificial Intelligence and Data Sciences (AISD)**

Theme

Multiview Anomaly Detection using Deep Neural Networks

Presented by:
ABDELAZIZ Chourouk Halah

Submission Date: **Sept, 2024**
In front of the jury composed of:

Dr. KHALDI Belkacem	President
Pr. BENABDESLEM Khalid	Supervisor
Dr. CHAIB Souleyman	Co-Supervisor
Dr. BEKKOUCHE Mohammed	Examiner

Academic Year : 2023/2024

Abstract

The growth of complex systems across diverse domains, such as healthcare, finance, intelligent transportation, and social media, has resulted in a flood of data. This data often manifests in heterogeneous formats, offering multifaceted "views" of the underlying phenomena. For instance, in healthcare, a patient's data may encompass medical images, sensor readings from wearable devices. This diversity makes anomaly detection in contemporary datasets particularly challenging.
Traditional anomaly detection methods often rely on a singular data view or data type, which is very limiting. Real-world systems, however, typically generate data from multiple sources or views as we mentioned, creating multi-view datasets. While multi-view data holds immense potential for enhancing anomaly detection, it remains a significant challenge(heterogeneous data structures, varying scales, and the need for integration methods…etc).
Current research in this domain lacks comprehensive approaches that effectively integrate multi-view data for anomaly detection.This thesis aims to bridge this gap by thoroughly examining existing methods and introducing key aspects of anomaly detection within the framework of multiview learning.

Keywords— Anomaly detection, Multiview learning, deep learning, Multiview deep AD

Résumé

La croissance des systèmes complexes dans divers domaines tels que la santé, la finance, les transports intelligents et les médias sociaux a entraîné une prolifération de données. Ces données se manifestent souvent sous des formats hétérogènes, offrant des "vues" multifacettes des phénomènes sous-jacents. Par exemple, dans le domaine de la santé, les données d'un patient peuvent comprendre des images médicales et des relevés de capteurs provenant de dispositifs portables. Cette diversité rend la détection des anomalies dans les ensembles de données contemporains particulièrement difficile.

Les méthodes traditionnelles de détection des anomalies reposent souvent sur une seule vue ou un seul type de données, ce qui est très limitatif. Les systèmes réels, cependant, génèrent généralement des données provenant de multiples sources ou vues, créant ainsi des ensembles de données multi-vues. Bien que les données multi-vues aient un potentiel immense pour améliorer la détection des anomalies, elles posent des défis considérables (structures de données hétérogènes, échelles variables, et la nécessité de méthodes d'intégration, etc.).

Les recherches actuelles dans ce domaine manquent d'approches exhaustives qui intègrent efficacement les données multi-vues pour la détection des anomalies. Cette thèse vise à combler cette lacune en examinant en profondeur les méthodes existantes et en introduisant des aspects clés de la détection des anomalies dans le cadre de l'apprentissage multi-vues.

Keywords— Détection des anomalies, apprentissage multivue, apprentissage profond, détection des anomalies profonde multivue.

الملخص

إن نمو الأنظمة المعقدة في مجالات متنوعة مثل الرعاية الصحية، والتمويل، والنقل الذكي، ووسائل التواصل الاجتماعي، قد أدى إلى تدفق كبير من البيانات. غالبًا ما تتجلى هذه البيانات في أشكال غير متجانسة، مما يوفر "مناظر" متعددة الأوجه للظواهر الأساسية. على سبيل المثال، في مجال الرعاية الصحية، قد تشمل بيانات المريض صورًا طبية وقراءات من أجهزة استشعار قابلة للارتداء. تجعل هذه التنوع عملية اكتشاف الشذوذ في مجموعات البيانات المعاصرة تحديًا كبيرًا.

تعتمد طرق اكتشاف الشذوذ التقليدية غالبًا على منظور بيانات واحد أو نوع واحد من البيانات، مما يعد محدودًا للغاية. ومع ذلك، فإن الأنظمة الواقعية تولد عادةً بيانات من مصادر أو مناظر متعددة، مما يؤدي إلى إنشاء مجموعات بيانات متعددة المناظر. رغم أن البيانات متعددة المناظر لديها إمكانات هائلة لتحسين اكتشاف الشذوذ، إلا أنها تشكل تحديات كبيرة (مثل هياكل البيانات غير المتجانسة، المقاييس المتغيرة، والحاجة إلى طرق تكامل، إلخ).

تفتقر الأبحاث الحالية في هذا المجال إلى نهج شامل يدمج بشكل فعال البيانات متعددة المناظر لاكتشاف الشذوذ. تهدف هذه الأطروحة إلى سد هذه الفجوة من خلال دراسة شاملة للطرق الموجودة وتقديم جوانب رئيسية لاكتشاف الشذوذ في إطار التعلم متعدد المناظر.

الكلمات المفتاحية: اكتشاف الشذوذ، التعلم متعدد المناظير، التعلم العميق، اكتشاف الشذوذ العميق متعدد المناظير.