

PEOPLES'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTER OF HIGH EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY OF FERHAT ABASS 1
FACULTY OF SCIENCES
Computer Sciences Department



MASTER THESIS

Presented with a view to obtain the Master's degree in Data Engineering and
Web Technologies

Machine Learning and Deep Learning for Autism Spectrum Disorder (ASD) detection

Supervisor:
Dr. Semchedine Moussa

By:
Mahdaoui Nour elhouda
Belkefoul kaouther

Defended on June 2, 2025, before a jury composed of:

PrESIdEnt: Dr. Belbal Samir UFA, SETIF1

ExAMInER: Dr. Slimani Yacine UFA, SETIF1

2024 / 2025

Abstract

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by difficulties with social interaction, communication, and behavior. Early identification ensures timely intervention and improves quality of life. The present study proposes a two-modality binary classification approach for ASD detection based on facial image data and movement data. The image dataset consists of facial photos of children labeled as ASD or non-ASD. The movement data, captured using a Kinect v2 sensor, comprises 1,259 features per subject derived from 3D joint positions and gait metrics, with a total of 800 samples.

A DenseNet121 convolutional neural network (CNN) was employed for images, achieving 89% accuracy. Grad-CAM was used to provide visual explanations by highlighting important regions in the images. For the movement modality, a Multi-Layer Perceptron (MLP) trained on features learned via an autoencoder achieved 99.38% accuracy, with Shapley Additive exPlanations (SHAP) applied to identify key features influencing model decisions. Finally, a late fusion mechanism combining both models was evaluated, resulting in 88.12% accuracy. Results highlight the effectiveness of unimodal solutions, particularly the MLP with autoencoder, and indicate that multimodal fusion requires further optimization for enhanced overall performance.

Keywords: Autism Spectrum Disorder, Early detection, Deep learning, Facial images, Movement.

Résumé

Le trouble du spectre de l'autisme (TSA) est un trouble neurodéveloppemental caractérisé par des difficultés dans les interactions sociales, la communication et le comportement. Une identification précoce permet une intervention rapide et améliore la qualité de vie. La présente étude propose une approche de classification binaire à deux modalités pour la détection du TSA, basée sur des données d'images faciales et des données de mouvement. Le jeu de données d'images comprend des photos de visages d'enfants diagnostiqués TSA ou non. Les données de mouvement, capturées à l'aide d'un capteur Kinect v2, comportent 1 259 caractéristiques par sujet, extraites des positions 3D des articulations et des métriques de la démarche, pour un total de 800 échantillons.

Un réseau de neurones convolutionnel DenseNet121 a été utilisé pour les images, atteignant une précision de 89 %. Grad-CAM a permis de fournir des explications visuelles en mettant en évidence les régions importantes des images. Pour la modalité mouvement, un perceptron multi-couches (MLP) entraîné sur des caractéristiques extraites via un autoencodeur a atteint une précision de 99.38%, avec l'application des explications additives de Shapley (SHAP) pour identifier les caractéristiques clés influençant les décisions du modèle. Enfin, un mécanisme de fusion tardive combinant les deux modèles a été évalué, aboutissant à une précision de 88.12 %. Ces résultats soulignent l'efficacité des solutions unimodales, en particulier le MLP avec autoencodeur, et indiquent que la fusion multimodale nécessite une optimisation supplémentaire pour améliorer la performance globale.

Mots Clée: Trouble du Spectre de l'Autisme, Détection précoce, Apprentissage profond, Images faciales, Mouvement.

المخلص

اضطراب طيف التوحد (ASD) هو اضطراب نمائي عصبي يتميز بصعوبات في التفاعل الاجتماعي، والتواصل، والسلوك. يضمن التشخيص المبكر التدخل المبكر ويحسن جودة الحياة. تقترح هذه الدراسة نهج تصنيف ثنائي يعتمد على نوعين من البيانات: صور الوجه وبيانات الحركة. تتكون مجموعة بيانات الصور من صور لوجوه أطفال تم تصنيفهم على أنهم مصابون بالتوحد أو غير مصابين. تحتوي بيانات الحركة، التي تم التقاطها باستخدام جهاز Kinect v2، على 1259 ميزة لكل فرد مشتقة من مواقع المفاصل ثلاثية الأبعاد ومقاييس المشي، بمجموع 800 عينة.

تم استخدام شبكة عصبية التلافيفية من نوع DenseNet121 لتحليل الصور، وحققت دقة بلغت 89%. تم استخدام تقنية Grad-CAM لتوفير تفسيرات بصرية من خلال تسليط الضوء على المناطق المهمة في الصور. بالنسبة لبيانات الحركة، تم تدريب شبكة Perceptron متعددة الطبقات (MLP) على ميزات مستخرجة عبر التشفير التلقائي (autoencoder)، وحققت دقة بلغت 99.38%، مع تطبيق تفسيرات Additive Shapley (SHAP) لتحديد الميزات الرئيسية التي تؤثر على قرارات النموذج. أخيراً، تم تقييم آلية دمج متأخر تجمع بين النموذجين، وحققت دقة بلغت 88.12%. تسلط النتائج الضوء على فعالية الحلول الأحادية النمط، خاصة MLP مع التشفير التلقائي، وتشير إلى أن دمج الأنماط المتعددة يحتاج إلى تحسين لتعزيز الأداء الكلي.

الكلمات المفتاحية: اضطراب طيف التوحد، الكشف المبكر، التعلم العميق، صور الوجه، الحركة

TABLE OF CONTENTS

List of Figures	iv
List of Tables	v
List of abbreviations	vi
General Introduction	1
Chapter 1 : Autism spectrum disorder (ASD)	3
Introduction	4
1 Definition and overview	4
1.1 Historical perspective on autism	5
2 Prevalence and epidemiology of ASD	6
3 Causes and risks	6
3.1 Genetic factors	6
3.2 Environmental risk factors	7
3.3 Parental age and premature birth	7
4 Signs and symptoms of autism spectrum disorder (ASD)	8
5 Diagnosis of autism spectrum disorder	9
6 Importance of early diagnosis.....	11
7 Comorbidities of autism spectrum disorder (ASD).....	11
7.1 Common comorbidities	11
7.2 Psychiatric conditions.....	12
8 Interventions and treatments for ASD	12

9	The role of technology in autism spectrum disorder.....	13
	Conclusion	14
Chapter 2: Background and Literature Review		15
	Introduction.....	16
1	Machine learning (ML)	16
	1.1 Definition and basic concepts of ML	16
	1.2 Types of machine learning.....	17
2	Deep learning (DL)	19
	2.1 Definition and basic concepts of DL.....	19
	2.2 Evolution of deep learning models.....	20
	2.3 Types of artificial neural networks	21
	2.4 Evaluation metrics for deep learning and machine learning models	25
	2.5 Deep learning vs. traditional machine learning.....	26
3	ML and DL for ASD detection	26
	3.1 Advantages of ML/DL over traditional methods	27
	3.2 Analyzing complex patterns in large datasets.....	27
4	Machine learning techniques for ASD detection	27
	4.1 Neuroimaging-based approaches.....	28
	4.2 Behavioral data and feature-based models	30
	4.3 Hybrid and ensemble learning approaches.....	31
5	Deep learning techniques for ASD detection.....	32
	5.1 Convolutional neural networks for ASD classification.....	33
	5.2 Transfer learning for ASD detection using facial features.....	34
	5.3 Hybrid and multi-modal learning approaches	36
6	Data sources for machine learning and deep learning in autism spec- trum disorder	36
7	Challenges and limitations in ML/DL for ASD detection	37
	7.1 Unavailable MRI neuroimaging datasets of ASD patients	38
	7.2 Challenges in artificial intelligence algorithms in diagnosing autism spectrum disorder	38
	7.3 Challenges in hardware	38
	Conclusion	39
Chapter3: Contribution,experiments and results		40
	Introduction.....	41
1	Development environment.....	41
	1.1 Hardware.....	41
	1.2 Software	41
2	Proposed methodology	43
	2.1 Phase 1: Motion data analysis.....	43

2.2	Phase 2: Analysis of image data using deep learning.....	47
2.3	Phase 3: Multimodal learning via late fusion	51
3	Validation Strategy	53
3.1	Dataset Preparation.....	53
3.2	Model Architectures	53
3.3	Experimental Settings	56
3.4	Results and Analysis	57
	Conclusion	59
	Conclusion and future work	60
	Bibliographie	75

