تطوير نظام ذكي للكشف المبكر عن مرض التوحد

[1]ريمان الحاج أحمد [2]د.خالد عمر

الملخص

يُعد اضطراب طيف التوحد (Autism Spectrum Disorder – ASD) من الاضطرابات النمائية العصبية المعقدة التي تؤثر على وظائف الدماغ والسلوك، ولا سيما مهارات التواصل والتفاعل الاجتماعي. ونظرا لأهمية الكشف المبكر في تحسين فرص التدخل العلاجي ورفع جودة الحياة لدى الأطفال المصابين، شهدت الأبحاث الحديثة توجها متزايدا نحو توظيف تقنيات الذكاء الاصطناعي والتعلم العميق في دعم عمليات التشخيص الطبي. يهدف هذا البحث إلى تطوير نظام ذكي للكشف المبكر عن اضطراب طيف التوحد اعتماداً على تحليل صور الوجه باستخدام نموذج Xception العميق. ركزت الدراسة على تحسين مراحل المعالجة المسبقة للبيانات وتصميم البنية المعمارية للنموذج من أجل تعزيز الأداء التشخيصي. تم تدريب النموذج باستخدام مجموعة بيانات حقيقية تضم صور أطفال بخصائص متنوعة، مع تطبيق تقنيات تحسين إضافية أثناء عملية التدريب. جرى تقييم أداء النموذج باستخدام عدة مؤشرات، من بينها الدقة والحساسية والقدرة التمييزية. أظهرت النتائج كفاءة عالية للنظام المقترح في التمييز بين الفئات، مما يؤكد إمكانيته كأداة داعمة للكشف المبكر عن التوحد والمساهمة في تحسين التدخل العلاجي وفهم السمات الوجهية المرتبطة بهذا الاضطراب.

الكلمات المفتاحية: اضطراب طيف التوحد، الأنظمة الذكية، الشبكات العصبية العميقة، التعلم العميق.

1 ريمان الحاج احمد طالبة ماجستير دراسات عليا في الجامعة الافتراضية، برنامج علوم الويب .
2 د. خالد عمر رئيس قسم الذكاء الإصطناعي في كلية الهندسة المعلوماتية، جامعة دمشق.

Developing an intelligent system for early detection of autism

Reman Alhaj Ahmad1, Dr. Khaled Omar2
[1] Master Student, Web Science Academic Program, Syrian Virtual University SVU, Syria (E-mail: reman_206339@svuonline.org).
[2] Head of the Artificial Intelligence Department at the Faculty of Informatics Engineering, Damascus University, Syria (E-mail: kh.om.mail@gmail.com)

Abstract

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder that affects brain function and behavior, particularly social communication and interaction skills. Given the importance of early detection in improving therapeutic interventions and enhancing the quality of life of affected children, recent research has increasingly focused on the use of artificial intelligence and deep learning techniques to support medical diagnostic processes. This study aims to develop an intelligent system for the early detection of Autism Spectrum Disorder based on facial image analysis using the deep Xception model. The research emphasizes improving data preprocessing stages and optimizing the model's architectural design to enhance diagnostic performance. The model was trained using a real-world dataset comprising facial images of children with diverse characteristics, with additional optimization techniques applied during the training process. The system's performance was evaluated using multiple metrics, including accuracy, sensitivity, and discriminative ability. The results demonstrate high effectiveness of the proposed system in distinguishing between classes, confirming its potential as a supportive tool for early autism detection and for improving therapeutic intervention and understanding facial features associated with Autism Spectrum Disorder.

KEYWORDS: Autism Spectrum Disorder, Intelligent Systems, Deep Neural Networks, Deep Learning.

1. المقدمة:

في ظل التطورات المتسارعة في قطاع الرعاية الصحية الحديثة، أصبح الكشف المبكر عن اضطرابات النمو العصبية، وعلى رأسها اضطراب طيف التوحد (ASD)، ضرورة ملحة لما له من دور أساسي في تحسين جودة حياة الأطفال المصابين وتعزيز فرص التدخل العلاجي الفعال على المدى الطويل. يُعد اضطراب طيف التوحد من أكثر الاضطرابات النمائية العصبية تعقيداً وانتشاراً، إذ يؤثر بشكل مباشر على مهارات التواصل الاجتماعي والسلوك والقدرات الإدراكية [1]. وغالبا ما يُظهر الأطفال المصابون بالتوحد أنماطا سلوكية وتعلمية غير نمطية، مثل السلوكيات المتكررة أو الاهتمام المفرط بأشياء محددة. وتشير الإحصاءات الحديثة الصادرة عن مركز السيطرة على الأمراض والوقاية منها (CDC) إلى تزايد ملحوظ في انتشار الاضطراب، حيث سُجّلت حالة واحدة من كل 36 طفلا بعمر 8 سنوات في الولايات المتحدة [2]. كما تؤكد بيانات شبكة رصد التوحد والإعاقات النمائية (ADDM) تضاعف معدلات التشخيص بين عامي 2000 و2021 [3]. وعلى الرغم من هذا الانتشار المتزايد، لا يزال التشخيص المبكر يمثل تحدياً بسبب دقة الأعراض الأولية وتباين مظاهرها، فضلًا عن اعتماد الطرق التقليدية على التقييمات السريرية السلوكية. ومع التقدم في تقنيات الذكاء الاصطناعي والتعلم العميق، برزت إمكانات واعدة لتحليل البيانات البصرية، ولا سيما صور الوجه، باعتبارها مصدراً غنياً بالمؤشرات المرتبطة باضطراب طيف التوحد. وفي هذا السياق، يسعى هذا البحث إلى تطوير نظام ذكي يعتمد على تحليل صور الوجه لدعم الكشف المبكر عن اضطراب طيف التوحد وتحسين دقة وموضوعية التشخيص مقارنة بالأساليب التقليدية.

2. هدف البحث:

يهدف هذا البحث إلى تطوير نظام ذكي للكشف المبكر عن اضطراب طيف التوحد اعتماداً على تحليل صور الوجه باستخدام تقنيات الذكاء الاصطناعي والتعلم العميق. يعتمد النظام على نموذج Xception كشبكة أساسية مع دمج وحدة الانتباه CBAM لاستخلاص السمات الوجهية الدقيقة ذات الدلالة التشخيصية. كما تم تحسين جودة البيانات وتطبيق تقنيات تدريب متقدمة لضمان استقرار النموذج ورفع

دقته. جرى تقييم النظام باستخدام مقاييس إحصائية شاملة على مجموعة اختبار مستقلة للتحقق من موثوقية الأداء.

3. مواد وطرق البحث:

اعتمدت هذه الدراسة على لغة البرمجة Python لتطوير النظام المقترح، نظراً لمرونتها وانتشارها الواسع في تطبيقات الذكاء الاصطناعي والتعلم العميق. تم تنفيذ النموذج باستخدام بيئة Google Colab، لما توفره من بنية حوسبة سحابية تفاعلية ودعم لوحدات المعالجة الرسومية (GPU) لتسريع تدريب النماذج العميقة. استُخدم إطارا TensorFlow و Keras لبناء وتدريب نموذج Xception ودمج وحدة الانتباه CBAM وضبط خوارزميات التحسين. كما تم توظيف مكتبات OpenCV و PIL لمعالجة الصور وتنفيذ مراحل المعالجة المسبقة، إلى جانب NumPy و Pandas للمعالجة العددية وإدارة البيانات. أُجري تحليل النتائج وتمثيلها بصرياً باستخدام Matplotlib، مع ضمان التوافق بين جميع الأدوات لتحقيق الاستقرار وقابلية إعادة التجربة.

4. الدراسات السابقة:

قام الباحثون في الدراسة [4] بالاستفادة من خصائص وجوه الأطفال المصابين باضطراب طيف التوحد والأطفال الطبيعيين للكشف عن التوحد باستخدام نماذج التعلم العميق VGG16 و Xception القائمة على التعلم بالنقل. اعتمدت المنهجية على مجموعة بيانات من منصة Kaggle تضم 2940 صورة وجه، تم تقسيمها إلى مجموعات تدريب وتحقق واختبار بعد تطبيق المعالجة المسبقة والتقييس. استُخدمت الشبكات العصبية التلافيفية (CNN) لاستخلاص السمات الوجهية الدقيقة تلقائياً عبر طبقات الالتفاف والتجميع والاتصال الكامل. تم تحسين التدريب باستخدام Batch Normalization والإيقاف المبكر ومُحسن RMSprop مع استخدام دالة التنشيط Softmax في طبقة الإخراج لأغراض التصنيف. أظهرت نتائج التقييم تفوق نموذج Xception بدقة بلغت 91% مقارنة بنموذج VGG16 الذي حقق دقة قدرها 78%.

في الورقة البحثية [5] تم تطوير نموذج يعتمد على التعلم بالنقل مبني على معمارية AlexNet لتشخيص اضطراب طيف التوحد من خلال تحليل ملامح الوجه باستخدام صور أطفال مصابين وغير مصابين

بالتوحد من منصة Kaggle. استُخدم نموذج AlexNet المدرب مسبقاً على ImageNet مع إعادة تدريب طبقاته الأخيرة لتلائم مهمة تصنيف التوحد، مع الإبقاء على الطبقات الأولى لاستخلاص السمات العامة. تم تدريب النموذج باستخدام 80% من البيانات والتحقق باستخدام 20%، مع اعتماد خوارزمية SGDM ومعدل تعلم قدره 0.0001 .أظهرت النتائج أداء مرتفعاً للنموذج بدقة وصلت إلى 91%، متفوقا على أساليب تقليدية مثل PCA و CNN التقليدي..

في الدراسة [6] ركز الباحثون على توظيف تقنيات التعلم الآلي في تحليل اضطراب طيف التوحد باستخدام بيانات ديموغرافية وطبية وتصويرية، مع مقارنة عدة خوارزميات تصنيف. استُخدمت خوارزميات مثل SVM و Random Forest و Decision Tree والانحدار اللوجستي، إلى جانب تقنيات التعلم العميق لتحليل بيانات التصوير بالرنين المغناطيسي. تم تطبيق اختيار الميزات وتحليل المكونات الرئيسية (PCA) مع استخدام أسلوب الموازنة العشوائية للبيانات لتحسين أداء المصنفات. أظهرت النتائج تفوق مصنف الفضاء المتجه (SVM) بدقة تصنيف بلغت 96% مقارنة ببقية الخوارزميات. كما بينت الدراسة أن الجنس والعمر واليرقان تُعد من العوامل المؤثرة في اضطراب طيف التوحد، مؤكدة فعالية التعلم الآلي في دعم التشخيص المبكر.

في [7] يقترح هذا البحث أداة فحص إلكترونية سريعة عبر الإنترنت للكشف المبدئي عن اضطراب طيف التوحد اعتماداً على صور وجوه الأطفال.

تعتمد الأداة على نموذج DenseNet من الشبكات العصبية الالتفافية، والمُدرب باستخدام خوارزمية التدرج الفرعي المتحكم به (CSA) لتحسين معدل التعلم ورفع دقة التصنيف.

تم تطوير النظام باستخدام إطار Django بلغة بايثون واستضافته سحابيا، مع آلية بسيطة لرفع الصور ومعالجتها ثم عرض نتيجة التنبؤ ونسبة الثقة.

أظهرت النتائج أداءً مرتفعاً للنموذج بدقة اختبار بلغت 98%، مع قيم متوازنة لمقاييس Precision و Recall و F1 لكلتا الفئتين.

يخلص البحث إلى أن الأداة تمثل وسيلة عملية داعمة للتشخيص المبكر، مع التأكيد على أنها لا تغني عن التقييم الطبي المتخصص.

5. الشبكات العصبية:

1.5 الشبكات العصبية الاصطناعية (ANN):

تُعرف الشبكات العصبية الاصطناعية بأنها نماذج حسابية تحاكي – إلى حدٍ ما – آلية عمل الدماغ البشري من خلال عصبونات مترابطة تعالج المعلومات بصورة مشابهة للخلايا العصبية [8] .

تتكون هذه الشبكات من طبقة إدخال تستقبل البيانات الأولية، وطبقات مخفية مسـؤولة عن معالجة البيانات واستخلاص الأنماط، وطبقة إخراج تُنتج التنبؤ النهائي.

تعتمد الشبكات العصبية على دوال تنشـيط مثل Sigmoid و Tanhو ReLUلإضفاء خاصية اللاخطية وتمكين النموذج من تمثيل العلاقات بين المدخلات والمخرجات.

2.5 الشبكات العصبية العميقة DNN والتعلم العميق Deep Learning :

الشبكات العصبية العميقة هي امتداد للشبكات العصبية الاصطناعية التقليدية، حيث تحتوي على عدد كبير من الطبقات المخفية، مما يسمح لها بتمثيل العلاقات المعقدة داخل البيانات.

يُعد التعلم العميق أحد الفروع المتقدمة للتعلم الآلي، ويعتمد على استخدام الشبكات العصبية الاصطناعية العميقة .

يتميز التعلم العميق بقدرته العالية على التعامل مع كميات ضخمة من البيانات غير المهيكلة مثل الصور،الصوت، والنصوص، وقد أدى ذلك إلى تطور كبير في تطبيقات الذكاء الاصطناعي، خاصة في مجالات الرؤية الحاسوبية ومعالجة اللغة الطبيعية.

بخلاف خوارزميات التعلم الآلي التقليدية التي تعتمد على هندسة الميزات يدوياً، تقوم نماذج التعلم العميق باستخلاص الميزات تلقائياً من البيانات الخام، حيث تتعلم الطبقات الأولى الخصائص البسيطة بينما تتعلم الطبقات العميقة الأنماط الأكثر تعقيداً [9] [20].

3.5 الشبكات العصبية التلافيفية CNNs :

تُعد الشبكات العصبية التلافيفية من أهم نماذج التعلم العميق المستخدمة في مجال الرؤية الحاسوبية Computer Vision ، حيث صُممت خصيصاً لمعالجة البيانات ذات البنية الشبكية مثل الصور والفيديو[9] [10]، وهو ما يجعلها محور هذه الدراسة .

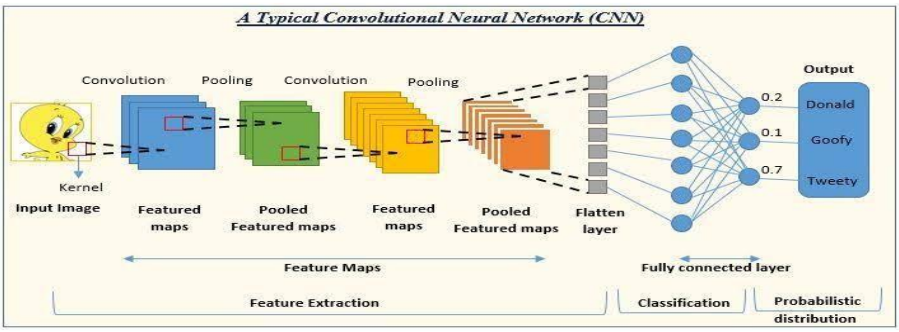
1.3.5 المدخلات والفلاتر في الشبكات العصبية التلافيفية :

تُمثَّل الصور المدخلة إلى الشبكة العصبية التلافيفية على شكل مصفوفات رقمية، حيث تعبر كل قيمة عن شدة بكسل لونية، وتُطبق عليها عمليات رياضية متتالية لاستخلاص الخصائص المهمة.

تُستخدم الفلاتر (Filters) وهي مصفوفات صغيرة الحجم مثل 3×3 أو 5×5تتحرك فوق الصورة لإجراء عمليات الالتفاف، بهدف استخراج الميزات الأساسية كالحواف، الزوايا، والأنماط المتكررة، والتي تُعد مدخلات أساسية لمرحلة التصنيف اللاحقة .

2.3.5 الطبقات الرئيسية في الشبكات العصبية التلافيفية CNN:

تتكون CNN من عدة طبقات أساسية، أبرزها طبقة الالتفاف (Convolutional Layer) التي تُعد جوهر الشبكة، حيث تطبق فلاتر متعددة على البيانات المدخلة. تليها طبقة التجميع (Pooling Layer) التي تُستخدم لتقليل الأبعاد المكانية لخرائط الميزات، مما يقلل التعقيد الحسابي ويحد من الإفراط في التعلم مع الحفاظ على الخصائص المهمة[11]، ومن أشهر أنواعها Max Pooling و Average Pooling. أما الطبقات الكثيفة (Fully Connected Layer) فتقع في نهاية الشبكة، حيث يتم ربط جميع العقد معا لاتخاذ القرار النهائي للتصنيف [12] وفقا للشكل (1) :



الشكل(1) : بنية شبكة عصبية تلافيفية CNN

3.3.5 آلية تدريب الشبكات العصبية التلافيفية CNN:

تتم عملية تدريب الشبكات العصبية التلافيفية من خلال تغذية الشبكة بصور مُسماة (Labeled Images) [8]، حيث تمر الصور عبر طبقات الالتفاف والتجميع لاستخراج الميزات والأنماط المهمة. بعد ذلك، يتم حساب الفرق بين المخرجات المتوقعة والمخرجات الحقيقية باستخدام دالة الخسارة وبالاعتماد على خوارزمية الانتشار العكسي (Backpropagation) وخوارزميات التحسين المبنية على مبدأ الانحدار المتدرج (Gradient Descent) ، يتم تحديث الأوزان والفلاتر تدريجياً بهدف تقليل قيمة الخطأ وتحسين أداء النموذج.

4.5 نماذج التعلم العميق المتقدمة:

مع تطور تقنيات التعلم العميق، ظهرت نماذج متقدمة مثل Xception المعتمدة على مبدأ الالتفاف القابل للفصل[13] (Depthwise Separable Convolutions) ، والذي يفصل بين الالتفاف المكاني والالتفاف عبر القنوات، مما يقلل عدد المعاملات والتعقيد الحسابي مع الحفاظ على دقة عالية. وقد أثبت هذا النموذج كفاءة كبيرة في مهام تصنيف الصور عالية الدقة، لذلك تم اعتماده في هذه الدراسة كعمود فقري للنموذج المقترح لقدراته العالية على استخلاص السمات البصرية الدقيقة المرتبطة بالمؤشرات التشخيصية لاضطراب طيف التوحد، مع تحقيق توازن مناسب بين الدقة والاستقرار الحسابي.

6. منهجية العمل:

يعتمد النظام المقترح على بنية متكاملة لأتمتة الكشف المبكر عن اضطراب طيف التوحد عبر تحليل السمات الوجهية الدقيقة باستخدام التعلم العميق. تبدأ العملية بإدخال صور الأطفال الخام، ثم تمر الصور بمرحلة معالجة مسبقة تشمل اكتشاف الوجه، مواءمته، توحيد الأبعاد، وتحسين الجودة البصرية لتقليل العوامل غير المهمة وتعزيز السمات التشخيصية. بعد ذلك، تُمرر الصور إلى نموذج Xception لاستخلاص السمات الوجهية العميقة تلقائياً، مع دمج وحدة الانتباه CBAM لإعطاء أولوية للمناطق الأكثر دلالة. تُستخدم طبقات النموذج النهائية مع دالة Softmax لإجراء التصنيف الثنائي وتحديد احتمالية انتماء الطفل لفئة المصابين أو غير المصابين.

1.6 مجموعة البيانات (Dataset):

تُعد مجموعة البيانات عنصراً أساسياً في أنظمة التعلم العميق، خاصة في التطبيقات الطبية التي تتطلب دقة ووثوقية عاليتين، إذ يرتبط أداء النموذج المقترح بشكل مباشر بجودة البيانات وتنوعها وطريقة معالجتها. لذلك، تم اختيار مجموعة بيانات مناسبة للكشف المبكر عن اضطراب طيف التوحد اعتماداً على تحليل صور الوجه.



الشكل(2) : عينة من مجموعة البيانات (– سليم، – مصاب)

وهي بيانات حقيقية منشورة عبر منصة Kaggle، مما يعزز قابلية إعادة إنتاج النتائج والشفافية العلمية، حيث تتضمن صور وجوه لأطفال مصنفين إلى فئتين: أطفال مصابون باضطراب طيف التوحد (Autistic) وأطفال غير مصابين (Typical) وقد استُخدم الترميز الثنائي لتمثيل الفئات، بما يتوافق مع مهام التصنيف الثنائي.

ولضمان تقييم موثوق لأداء النموذج، قُسمت البيانات إلى ثلاث مجموعات مستقلة: التدريب، التحقق، والاختبار، مع اعتماد التقسيم الطبقي (Stratified Splitting) للحفاظ على التوزيع النسبي للفئات في جميع المجموعات. تم تحميل البيانات وتنظيمها مباشرة داخل بيئة Google Colab باستخدام Kaggle API، ثم إخضاعها لعمليات المعالجة المسبقة قبل إدخالها إلى نموذج التعلم العميق.

رابط مجموعة البيانات على Kaggle :

https://www.kaggle.com/datasets/remoha/autismdataset

2.6 المعالجة المسبقة للبيانات (Data Preprocessing):

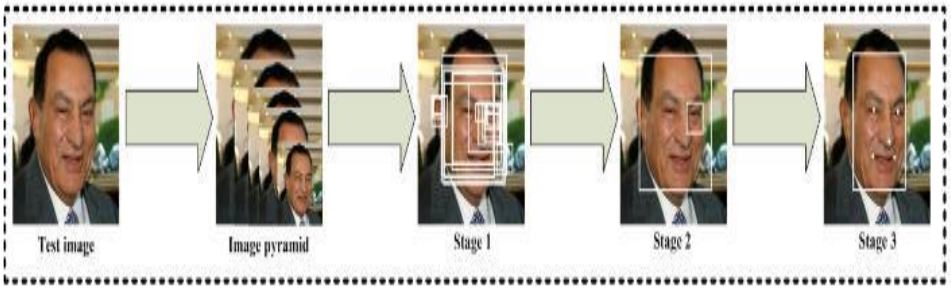
تُعد مرحلة المعالجة المسبقة للبيانات من أهم المراحل في أنظمة التعلم العميق المعتمدة على الصور [14]، إذ تؤثر بشكل مباشر على جودة السمات المستخرجة وعلى الأداء النهائي للنموذج. ونظراً لدقة الفروقات الوجهية المرتبطة باضطراب طيف التوحد، فقد تم تنفيذ سلسلة من الخطوات المنهجية لتحويل الصور الخام إلى مدخلات معيارية مناسبة لنموذج Xception المستخدم في هذه الدراسة، بما يضمن استقرار عملية التعلم وتحسين دقة التصنيف.

في البداية، تم تحميل الصور من المجلدات المخصصة لكل فئة Autistic و Typical، ومن ثم تطبيق خطوات المعالجة المسبقة عليها.

1.2.6 اكتشاف الوجه ومواءمته:

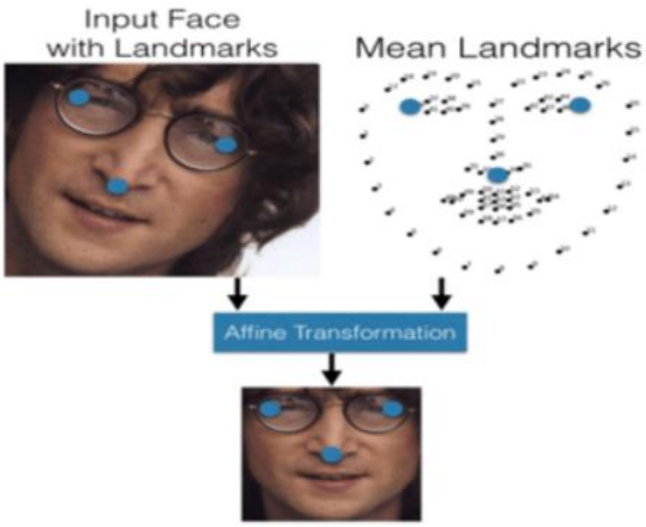
نظراً لاحتمال احتواء الصور الأصلية على خلفيات أو عناصر غير ذات صلة، تم تطبيق مرحلة اكتشاف الوجه ومواءمته كخطوة أساسية في المعالجة المسبقة، لضمان تركيز النموذج على الملامح الوجهية فقط. ولتحقيق ذلك، تم اعتماد خوارزمية MTCNN (Multi-Task Cascaded Convolutional Neural Network)، والتي تُعد من أكثر الخوارزميات دقة وانتشاراً في مجال اكتشاف الوجوه وتحديد المعالم الوجهية الأساسية [15]، مثل مواضع العينين، الأنف، الفم وحدود الوجه. تعتمد خوارزمية MTCNN على بنية متعددة المراحل تتكون من ثلاث شبكات عصبية تلافيفية متتابعة تعمل بشكل تسلسلي. في المرحلة الأولى، تقوم شبكة P-Net بتوليد مقترحات أولية لمواضع الوجوه المحتملة داخل الصورة باستخدام نوافذ منزلقة وعلى مقاييس متعددة، مما يسمح باكتشاف الوجوه الصغيرة والكبيرة. وفي المرحلة الثانية، تتولى شبكة R-Net تصفية هذه المقترحات من خلال استبعاد

المربعات الخاطئة وتحسين دقة تحديد موضع الوجه. أما في المرحلة الثالثة، فتقوم شبكة O–Net بإجراء التحديد النهائي لمواضع الوجوه إلى جانب استخراج المعالم الوجهية الأساسية بدقة عالية وفقاً للشكل (2):



الشكل(3) آلية عمل خوارزمية MTCNN

بعد استخراج المعالم الوجهية، تم تطبيق عملية مواءمة الوجه (Face Alignment) باستخدام تحويل تآلفي (Affine Transformation) ، وذلك من خلال الاستفادة من إحداثيات المعالم الوجهية المكتشفة. تهدف هذه العملية إلى تصحيح ميلان الوجه داخل الصورة، بحيث يتم تدوير الوجه rotation ، تحجيمه scaling وإزاحته translation ليصبح موضع العينين على خط أفقي ثابت وبمواقع نسبية متقاربة بين جميع الصور وفقاً للشكل (3). يساهم ذلك في تقليل الاختلافات الناتجة عن زاوية التصوير أو ميلان الرأس، مما يؤدي إلى توحيد التمثيل الهندسي للوجوه وتحسين اتساق المدخلات قبل تمريرها إلى نموذج Xception ، مما ينعكس إيجاباً على دقة السمات المستخرجة وأداء النظام ككل.



الشكل(4) مواءمة الوجه باستخدام تحويل تآلفي

2.2.6 توحيد أبعاد الصور:

بعد إتمام مرحلة مواءمة الوجه، تم توحيد أبعاد جميع الصور لتتوافق مع متطلبات نموذج Xception، حيث يستقبل النموذج مدخلات بحجم 299×299 بكسل. ممايضمن توافق الصور مع بنية النموذج، وتوحيد شكل المدخلات، ومنع الأخطاء البرمجية أثناء التدريب.

3.2.6 تقييس قيم البكسل (Normalization):

تعد عملية تقييس قيم البكسل خطوة أساسية لتحسين استقرار عملية التعلم. لم يُستخدم في هذه الدراسة التقييس التقليدي بتحويل القيم إلى المجال[0–1] ، بل تم اعتماد دالة preprocess_input الخاصة بنموذج Xception ، التي تقوم بتحويل قيم البكسل من المجال [0–255] إلى [−1,+1] بما يتوافق

مع آلية التقييس المعتمدة أثناء التدريب المسبق للنموذج على مجموعة بيانات ImageNet ، مما يضمن الاستفادة المثلى من الأوزان المسبقة وتحسين سرعة التقارب ودقة النموذج.

4.2.6 زيادة البيانات (Data Augmentation):

نظراً لمحدودية حجم البيانات في التطبيقات الطبية، تم اعتماد تقنيات زيادة البيانات بهدف تعزيز قدرة النموذج على التعميم وتقليل مشكلة فرط التعلم. وقد طُبقت هذه التقنيات حصرياً على بيانات التدريب، مع الإبقاء على بيانات التحقق والاختبار دون أي تحويلات لضمان نزاهة التقييم. شملت الزيادة التقليدية عمليات مثل الانعكاس الأفقي، التدوير المحدود، الإزاحات البسيطة، والتغيير الطفيف في المقياس، وذلك لمحاكاة التنوع الواقعي في ظروف التصوير دون التأثير على السمات الوجهية التشخيصية الأساسية.

إضافة إلى ذلك، تم اعتماد تقنيات زيادة بيانات متقدمة على مستوى الدفعات التدريبية، تمثلت في MixUp و CutMix [16] .

تعتمد **تقنية MixUp** على المزج الخطي بين عينتين عشوائيتين داخل الدفعة نفسها مع دمج تسمياتهما وفق معامل عشوائي λ مستمد من توزيع Beta كما هو موضح في الشكل (4) ، مما يؤدي إلى إنشاء عينات هجينة وتوسيع فضاء البيانات بشكل فعال، يساهم هذا الاسلوب في تقليل فرط التعلم وتنعيم حدود القرار.

الصياغة الرياضية :

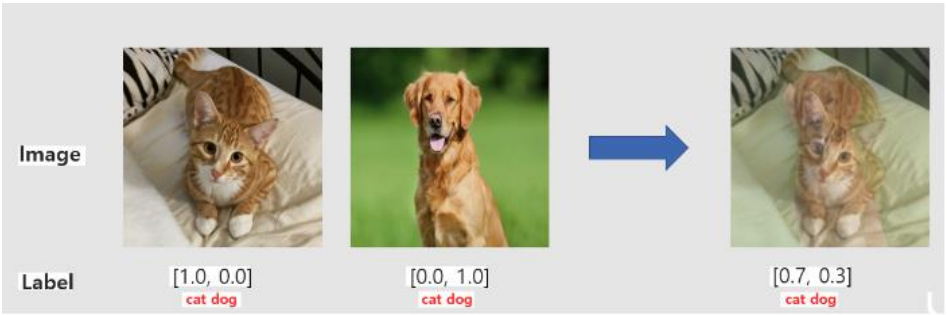
لتكن : $(xi,yi),(xj,yj)$ عينتين عشوائيتين مأخوذتين من نفس الدفعة التدريبية، حيث x تمثل الصورة و y تمثل التسمية المقابلة لها.

يتم توليد معامل المزج: $\lambda \sim Beta(\alpha,\alpha), \lambda \in [0,1]$

ثم تُحسب العينة الجديدة كما يلي:

$$x\sim = \lambda xi + (1-\lambda)xj$$

$$y\sim=\lambda yi+(1-\lambda)yj$$



الشكل (5): عينة جديدة باستخدام تقنية MixUp

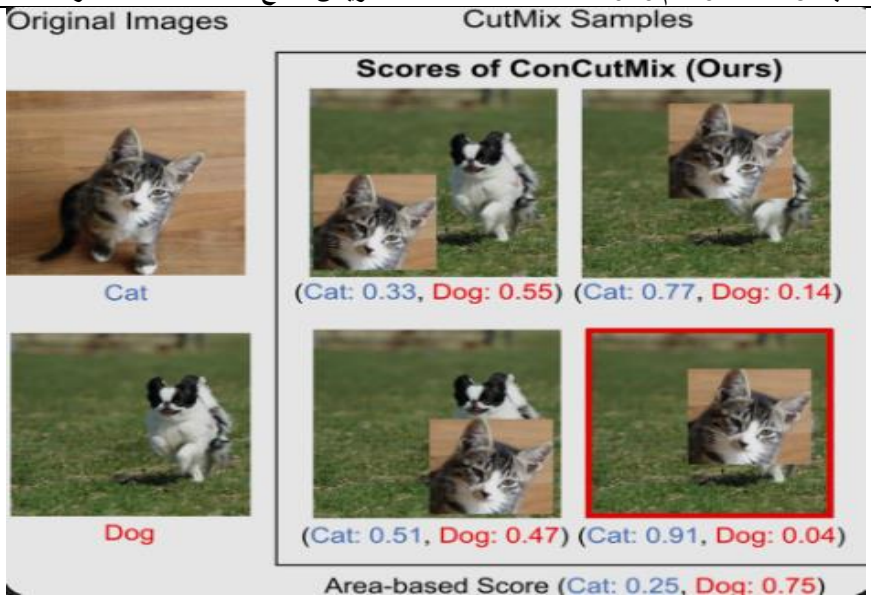
أما **تقنية CutMix** فتعتمد على الدمج الموضعي، حيث يتم استبدال جزء من صورة تدريبية بجزء من صورة أخرى مع تعديل التسمية وفق نسبة المساحة المستبدلة، الأمر الذي يدفع النموذج للتركيز على السمات الجزئية المهمة ويزيد من متانته.

الصياغة الرياضية : لتكن: (xi,yi),(xj,yj) صورتان تدريبيتان

يتم قص منطقة مستطيلة من الصورة xi واستبدالها بمنطقة من الصورة xj كما هو موضح في الشكل (5)، حيث يتم حساب معامل المزج λ وفق نسبة المساحة غير المستبدلة كما يلي:

$$\lambda = 1 - \frac{\text{مساحة الجزء المستبدل}}{\text{المساحة الكلية للصورة}}$$

ثم يتم تعريف التسمية الجديدة وفق المعادلة : $y\sim=\lambda yi+(1-\lambda)yj$



الشكل (6): توليد عينات جديدة للبيانات باستخدام تقنية CutMix

في نهاية مرحلة المعالجة المسبقة، تصبح الصور مكتشفة الوجه، مواءَمة هندسياً، موحدة الأبعاد، ومطبعة القيم، وجاهزة بالكامل للإدخال إلى نموذج Xception وقد أسهم الدمج بين الزيادة التقليدية والمتقدمة في تحسين استقرار عملية التعلم وتعزيز قدرة النموذج على استخلاص السمات الوجهية المرتبطة باضطراب طيف التوحد، مع الحفاظ على دقة ووثوقية التقييم.

3.6 بنية النموذج العميق المقترح:

تم اختيار نموذج Xception كأساس للنظام المقترح للكشف المبكر عن اضطراب طيف التوحد نظراً لقدرته العالية على استخلاص السمات الوجهية الدقيقة اعتماداً على الالتفافات القابلة للفصل عمقياً، مع كفاءة حسابية مناسبة.

يتكون النموذج من ثلاث مراحل رئيسية هي : مرحلة الإدخال(Entry Flow) ، المرحلة الوسطى (Middle Flow)، ومرحلة الإخراج (Exit Flow) .

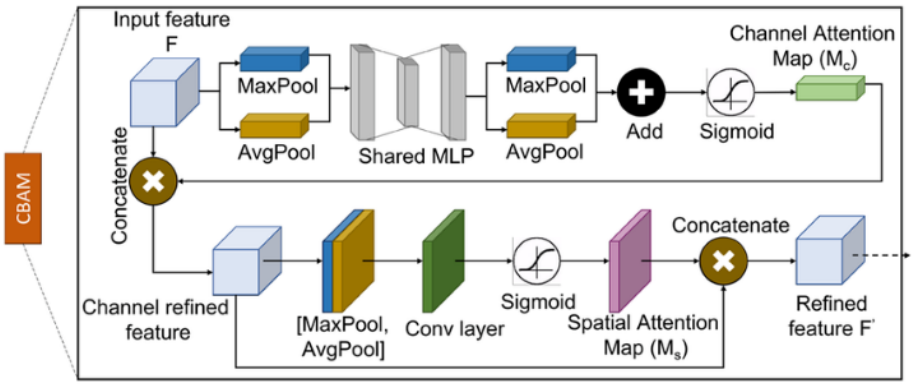
في مرحلة الإدخال، يتم استخلاص السمات الأولية للصورة، مثل الحواف والأنماط البسيطة، والتي تمثل الأساس لفهم البنية العامة للوجه. أما المرحلة الوسطى، فتتكرر فيها وحدات التلافيف القابلة للفصل عدة مرات، مما يسمح بتعلم سمات أكثر تعقيداً، مثل العلاقات المكانية وتوزيع الملامح الوجهية. في حين تركز مرحلة الإخراج على دمج السمات المستخرجة وتحويلها إلى تمثيل عالي المستوى غني بالمعلومات، يُستخدم لاحقاً في عملية التصنيف.

دمج وحدة الانتباه CBAM:

لتعزيز قدرة النموذج على التركيز على المناطق الوجهية ذات الأهمية التشخيصية، تم دمج وحدة الانتباه CBAM (Convolutional Block Attention Module) بعد مخرجات نموذج Xception . تعتمد آليات الانتباه على توجيه تركيز النموذج نحو الأجزاء الأكثر أهمية في البيانات المدخلة بدل معالجة جميع المعلومات بنفس الدرجة[17] . وبالتالي تسهم في إبراز المناطق المرتبطة بالفروقات الدقيقة، مثل العينين والفم وتعابير الوجه، مما يعزز القدرة التمييزية للنموذج.

تتكون وحدة CBAM من مرحلتين متتاليتين تعملان بشكل تكاملي:

- الانتباه القنوي (Channel Attention) : يهدف إلى تحديد القنوات الأكثر أهمية في خريطة السمات، وذلك باستخدام تقنيتي Global Average Pooling و Global Max Pooling، يليهما شبكة عصبية صغيرة مشتركة. تم استخدام معامل تقليل الأبعاد (ratio = 8) لتحقيق توازن بين تقليل التعقيد الحسابي والحفاظ على القدرة التمثيلية للشبكة.
- الانتباه المكاني (Spatial Attention): يركز على تحديد المناطق المكانية الأكثر أهمية داخل الصورة، من خلال دمج خرائط المتوسط والقيمة العظمى عبر القنوات، ثم تطبيق طبقة التفاف بنواة 7×7 متبوعة بدالة Sigmoid لإنتاج خريطة انتباه مكانية وفقاً للشكل (6) :



الشكل (7):مكونات وحدة الانتباه CBAM

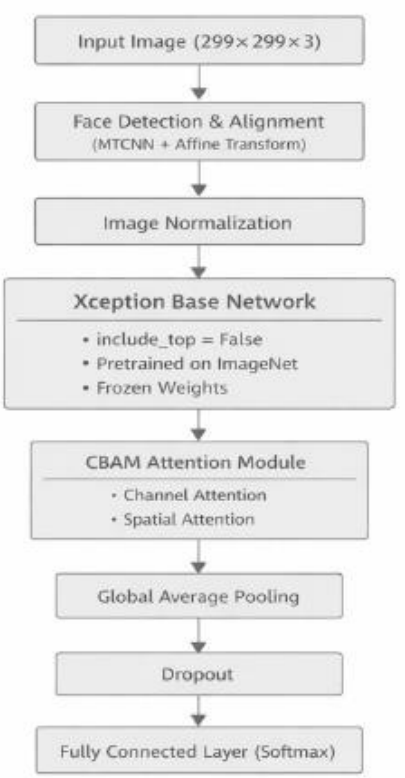
يسهم هذا التسلسل في تحسين جودة السمات المستخرجة، وزيادة حساسية النموذج للتفاصيل الدقيقة، وتقليل تأثير الاختلافات في الإضاءة والخلفيات، مما يدعم عملية التمييز بين الحالات المصابة باضطراب طيف التوحد وغير المصابة.

الطبقات النهائية للتصنيف(Classification Head)

بعد مرحلة استخلاص السمات وتعزيزها باستخدام وحدةCBAM ، يتم تمرير المخرجات إلى مجموعة من الطبقات النهائية المصممة خصيصاً لأغراض التصنيف. تبدأ هذه المرحلة باستخدام طبقة Global Average Poolingلتقليل الأبعاد المكانية لخريطة السمات وتحويلها إلى متجه سمات مضغوط يمثل المعلومات الجوهرية للصورة[18] .

بعد ذلك، تم تطبيق طبقة Dropout بهدف تقليل الاعتماد الزائد على السمات الفردية أثناء التدريب وتحسين قدرة النموذج على التعميم. وفي النهاية، استُخدمت طبقة كثيفة واحدة بعدد عقد يساوي عدد الفئات (عقدتين 2 Neurons)، مع دالة التفعيل Softmax لإجراء التصنيف الثنائي بين فئتي ASDو Non-ASD وفقا للشكل (7).

توفر دالة Softmax مخرجات احتمالية واضحة تسهل تفسير النتائج، وهو أمر بالغ الأهمية في التطبيقات الطبية، كما تتوافق مع دالة الخسارة Categorical Cross-Entropy المستخدمة أثناء التدريب.



الشكل(8) : البنية العامة للنموذج المقترح المعتمد على Xception مع آلية الانتباه CBAM

4.6 آلية تدريب النموذج وضبط المعاملات:

تم اعتماد تقنية **التعلم بالنقل** (Transfer Learning) في تدريب النموذج المقترح [19] من خلال استخدام نموذج Xception مُدرّب مسبقاً على مجموعة بيانات **ImageNet**، وذلك بهدف الاستفادة من المعرفة البصرية العامة التي تعلمها النموذج، مثل الحواف، والأنماط، والعلاقات المكانية، وهي سمات مناسبة بشكل خاص في ظل محدودية البيانات الطبية.

اعتمدت هذه الاستراتيجية على استبعاد الطبقات النهائية الأصلية الخاصة بالتصنيف ، وتحميل الأوزان المدربة مسبقاً، مع تجميد جميع طبقات النموذج الأساسية خلال المرحلة الأولى من التدريب. وقد أسهم

هذا الإجراء في تسريع عملية التدريب، وتحسين الاستقرار في المراحل الأولية، وتقليل خطر فرط التعلم، مع توجيه عملية التعلم نحو الطبقات المضافة الخاصة بالمهمة التشخيصية.

1.4.6 إعداد بيانات التدريب وخطوط البيانات:

تم تنظيم بيانات التدريب والتحقق باستخدام آليات تحميل ديناميكية تتيح قراءة الصور على دفعات أثناء التدريب. في المرحلة الأولى، تم تحميل الصور من مجلدات منظمة حسب الفئات مع تطبيق المعالجة المسبقة وزيادة البيانات الأساسية.

في المراحل المتقدمة، تم الاعتماد على واجهة tf.data.Dataset لإنشاء خطوط بيانات أكثر كفاءة، حيث أُتيح تطبيق الخلط العشوائي والتحميل المسبق، إضافة إلى تقنيات زيادة البيانات المتقدمة (Cutmix & Mixup). ساهم هذا الإجراء في تحسين كفاءة التنفيذ، تقليل استهلاك الذاكرة، وتعزيز قدرة النموذج على التعميم.

2.4.6 دالة الخسارة وخوارزمية التحسين :

تم استخدام دالة الخسارة Categorical Cross-Entropy بالتزامن مع طبقة Softmax في مخرج النموذج. كما تم تطبيق تقنية Label Smoothing بقيمة= ε (0.05).

ولتحديث أوزان النموذج أثناء التدريب تم اعتماد خوارزمية AdamW، وهي نسخة محسنة من Adam تدعم تطبيق Weight Decay بشكل صريح.

يسهم هذا الأسلوب التنظيمي في الحد من تضخم الأوزان، وتقليل فرط التعلم، وتحسين قدرة النموذج على التعميم، خاصة في البيئات الطبية الحساسة.

3.4.6 معدل التعلم :

نظرا لأهمية معدل التعلم في استقرار التدريب وجودة التقارب، تم اعتماد استراتيجية ديناميكية لضبطه. في مرحلة التدريب الأولي، استُخدمت جدولة Cosine Decay with Restarts مع بدء بمعدل تعلم منخفض نسبياً لضمان تقارب مستقر.

أما في مرحلة Fine-Tuning، فقد تم استخدام معدل تعلم صغير جداً (FT_LR = 8e-6) لتجنب إتلاف الأوزان المدربة مسبقاً وتحقيق تحسين دقيق في الأداء.

4.4.6 حجم الدفعة وعدد الحقب :

تم اعتماد حجم دفعة ثابت (Batch Size = 16) باستخدام واجهةtf.data ، بما يحقق توازنا بين كفاءة استخدام الذاكرة واستقرار التدريب، أما عدد الحقب(Epochs) ، فقد تم تحديده بشكل ديناميكي اعتماداً على أداء النموذج، مع استخدام آلية Early Stopping لمراقبة مقياس AUC على مجموعة التحقق واستعادة أفضل الأوزان المحققة تلقائياً.

5.4.6 آليات التحكم في التدريب:

تم استخدام مجموعة من آليات التحكم لضمان استقرار التدريب، شملت:

- Early Stopping
- Model Checkpoint
- Reduce Learning Rate on Plateau

وقد ساعدت هذه الآليات في تقليل زمن التدريب، منع التذبذب، والاحتفاظ بأفضل نسخة من النموذج بناءً على أعلى قيمة AUC .

6.4.6 تدريب رأس النموذج والضبط الدقيق(Fine-Tuning) :

خلال المرحلة الأولى، تم تدريب الطبقات المضافة فقط(CBAM ، Global Average Pooling، Dropout، وطبقة التصنيف النهائية)، بينما بقيت طبقات Xception الأساسية مجمدة.

بعد الوصول إلى استقرار مقبول، تم الانتقال إلى مرحلة Fine-Tuning من خلال فك تجميد عدد محدود من الطبقات العليا، مع تثبيت طبقات Batch Normalizationللحفاظ على الاستقرار الإحصائي وتقليل التذبذب في الأداء.

بعد الانتهاء من التدريب، تم اختبار النموذج على مجموعة بيانات مستقلة لم تُستخدم أثناء التدريب أو التحقق، دون تطبيق أي زيادة بيانات، وذلك لتقييم القدرة الحقيقية على التعميم والتأكد من أن إجراءات التحسين لم تؤدِّ إلى تحيز زائد.

7. مقاييس التقييم :

في التطبيقات الطبية، قد يؤدي أي خطأ تشخيصي إلى تبعات نفسية وسريرية كبيرة، لا سيما في الاضطرابات النمائية العصبية. لذا، لا يُعد الاعتماد على مقياس الدقة Accuracy وحده كافياً، بل يتطلب الأمر استخدام مقاييس متعددة تُبرز قدرة النموذج على اكتشاف الحالات المصابة فعلياً (الحساسية)، واستبعاد الحالات السليمة (النوعية)، وتحقيق توازن مناسب بين الاكتشاف الصحيح وتقليل الإنذارات الكاذبة.

- **إعداد التنبؤات باستخدام(Test–Time Augmentation (TTA**

يوضح الجدول (1) تأثير استخدام تقنية Test–Time Augmentation (TTA) على أداء النموذج. لوحظ أن تطبيق TTA باستخدام الانعكاس الأفقي أدى إلى تحسن ملحوظ في كل من AUC و F1–score مقارنة بالحالة بدون TTA، مما يشير إلى دور TTA في تعزيز استقرار التنبؤات وتحسين التعميم.

Setting	AUC	Best	Best Threshold
No TTA	0.921	0.853	0.50
TTA (Horizontal Flip)	0.9436	0.8787	0.44

الجدول (1) : تأثير استخدام TTA على اداء النمودج

- **مصفوفة الارتباك(Confusion Matrix)**

تُستخدم مصفوفة الارتباك لتحليل توزيع التنبؤات الصحيحة والخاطئة، وتحديد قيم (TN, FP) (TP FN, وتكتسب هذه الأداة أهمية سريرية خاصة، إذ تُعد الأخطاء من النوع الثاني (FN) أكثر خطورة في سياق التوحد، لأنها تعني عدم اكتشاف حالات مصابة وتأخير التدخل المبكر.

- **المقاييس المعتمدة لتقييم الأداء:**

- الدقة (Accuracy): تمثل النسبة الإجمالية للتنبؤات الصحيحة، إلا أنها قد تكون مضللة في حال عدم توازن البيانات.

$$\textbf{Accuracy= (TP+TN) / (FP+FN+TP+TN)}$$

- الحساسية (Sensitivity / Recall): تقيس قدرة النموذج على اكتشاف الحالات المصابة فعلياً، وتُعد من أهم المقاييس في هذه الدراسة نظراً لأولوية تقليل الحالات غير المكتشفة.

$$\textbf{Sensitivity} = \textbf{TP} / \textbf{FN+TP}$$

- النوعية (Specificity): تقيس قدرة النموذج على التعرف الصحيح على الحالات غير المصابة، وتسهم في تقليل القلق الناتج عن التشخيص الخاطئ.

$$\textbf{Specificity} = \textbf{TN} / \textbf{FP+TN}$$

- الدقة الإيجابية (Precision): تعكس مدى صحة التنبؤات الإيجابية، وتساعد في الحد من الإنذارات الكاذبة.

$$\textbf{Precision} = \textbf{TP} / \textbf{FP+TP}$$

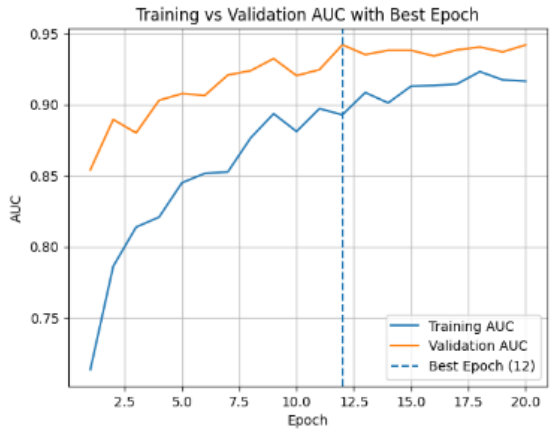
- مقياس F1−score : يجمع بين الحساسية والدقة الإيجابية في قيمة واحدة، ويُعد مناسباً بشكل خاص للبيانات الطبية غير المتوازنة.
- منحنى ROC والمساحة تحت المنحنى(AUC): يُستخدم منحنى ROC لتقييم العلاقة بين معدل الإيجابيات الحقيقية ومعدل الإيجابيات الكاذبة عبر مختلف عتبات القرار، بينما يعبر AUCعن القدرة التمييزية العامة للنموذج.

ويُعد مقياس AUC مناسباً بشكل خاص في تشخيص اضطراب طيف التوحد، لكونه غير متأثر بعدم توازن البيانات، ويقدم تقييماً شاملاً لأداء النموذج بغض النظر عن العتبة المختارة، مما يدعم الاعتماد عليه كنظام مساعد لاتخاذ القرار في التطبيقات الطبية الحساسة.

8. النتائج :

❖ نتائج التدريب والتحقق:

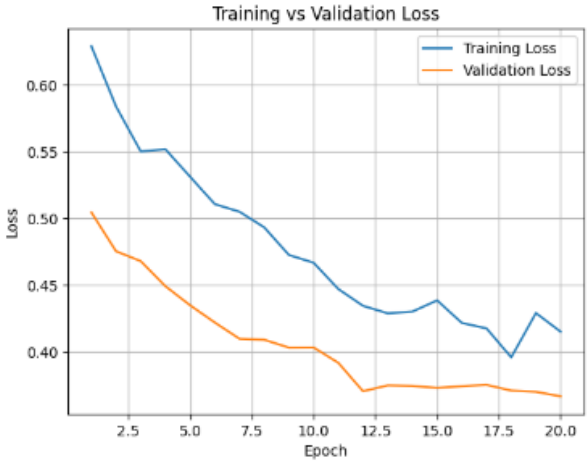
أظهرت نتائج التقييم أن النموذج المقترح يمتلك قدرة عالية على التمييز بين حالات اضطراب طيف التوحد (ASD) والحالات غير المصابة، مع تحقيق توازن مناسب بين الحساسية والنوعية. ويعكس هذا الأداء فعالية دمج نموذج Xception مع استراتيجيات المعالجة المسبقة المتقدمة وآليات التحسين المعتمدة خلال التدريب.



الشكل (9): منحنيات AUC للتدريب والتحقق

يوضح الشكل (8) تطور المساحة تحت منحنى (AUC) لكل من بيانات التدريب وبيانات التحقق. وقد سُجّل أفضل أداء على مجموعة التحقق عند الحقبة الثانية عشرة، حيث بلغت قيمة AUC 0.942، وذلك بالتزامن مع فك تجميد عدد محدود من الطبقات العليا في نموذج Xception بهدف تحسين تمثيل السمات عالية المستوى. ويشير هذا الارتفاع في قيمة AUC إلى قدرة النموذج على التمييز الجيد بين الفئتين، إضافة إلى فعالية آليتي ModelCheckpoint و ReduceLROnPlateau في تحسين الأداء.

بعد هذه الحقبة، لوحظ استقرار واضح في الأداء مع تقارب ملحوظ بين منحنيي التدريب والتحقق، مما يدل على استقرار عملية التعلم وقدرة النموذج على التعميم، ويؤكد فعالية استراتيجيات Fine-Tuning و Early Stopping المعتمدة في هذه الدراسة .



الشكل (10): منحنيات الخسارة loss للتدريب والتحقق

يوضح الشكل (9) منحنيات الخسارة لكل من بيانات التدريب والتحقق عبر 20 حقبة تدريبية. نلاحظ انخفاضاً تدريجياً في خسارة التدريب، يقابله انخفاض مستقر في خسارة التحقق، مما يشير إلى تحسن أداء النموذج وقدرته على التعميم. كما يُلاحظ استقرار خسارة التحقق بعد الحقبة الثانية عشرة، وهي الحقبة التي حقق عندها النموذج أفضل قيمة لمقياسAUC ، مما يبرر اعتماد آلية الإيقاف المبكر (Early Stopping) لتفادي فرط التعلم.

يعرض الجدول (2) أفضل نموذج تم اختياره اعتماداً على أعلى قيمة لمقياس AUC على مجموعة التحقق، وذلك باستخدام آلية الإيقاف المبكر وتخفيض معدل التعلم عند ثبات الأداء.

البند	القيمة	البند	القيمة
أفضل حقبة)	12	Training Loss	0.4344
Training AUC	0.8927	Validation Loss	0.3704
Validation AUC	**0.9420**	Learning Rate	8×10^{-6}
Training Accuracy	0.8860	Early Stopping	نعم
Validation Accuracy	0.8800	Reduce LR on	نعم

الجدول (2): ملخص أفضل أداء للنموذج

❖ **نتائج تقييم النموذج على مجموعة الاختبار:**

بعد تحديد أفضل نموذج بناءً على أداء مجموعة التحقق، تم تحميل الأوزان المحفوظة واختبار النموذج على مجموعة بيانات اختبار مستقلة تتكون من 300صورة موزعة بالتساوي بين الفئتين.

أظهرت نتائج التقييم أن النموذج حقق أداءً متميزاً على مختلف مقاييس التقييم الإحصائية، ويمكن تلخيصها كما يلي:

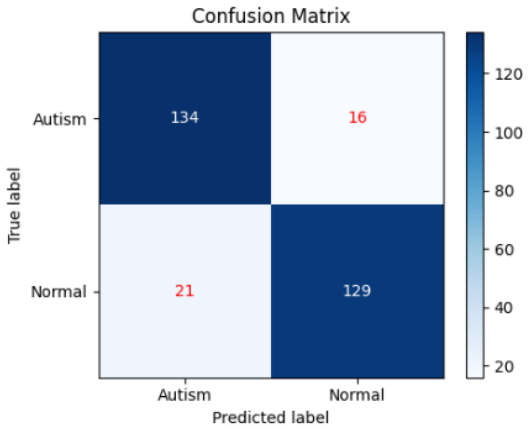
- **الدقة** : بلغت دقة التصنيف الإجمالية 87.67%، مما يدل على قدرة جيدة للنموذج على التنبؤ الصحيح عبر الفئتين.
- **المساحة تحت منحنى (AUC)**: حقق النموذج قيمة AUC = 0.9436، وهي قيمة مرتفعة تعكس قدرة قوية على التمييز بين الحالات المصابة وغير المصابة عبر مختلف قيم العتبة.
- **مقياس F1−score** : بلغ أفضل مقياس F1 قيمة 0.8787عند عتبة تصنيف مقدارها 0.44، وهو ما يعكس توازناً جيداً بين الحساسية والدقة الإيجابية، وهو أمر بالغ الأهمية في التطبيقات الطبية.

الفئة	Precision	Recall	F1-score
Non-ASD	0.8897	0.8600	0.8746
ASD	0.8645	0.8933	0.8787

الجدول (3) : مقاييس تقييم اداء النمودج المقترح حسب الفئات

تشير هذه النتائج إلى أن النموذج أظهر قدرة أعلى على اكتشاف الحالات المصابة فعلياً، وهو أمر بالغ الأهمية في السياق الطبي.

تحليل مصفوفة الارتباك:



الشكل(11): مصفوفة الارتباك

يوضح الشكل (10) مصفوفة الارتباك الخاصة بأفضل نموذج، حيث أسفرت النتائج عن:

- **(TP)** True Positives عددها **134** حالة، وتمثل الحالات المصابة التي صُنفت بشكل صحيح.
- **(TN)** True Negatives عددها **129** حالة، وتمثل الحالات غير المصابة التي صُنفت بشكل صحيح.
- **(FP)** False Positives عددها **21** حالة، وهي حالات غير مصابة صُنفت خطأ كمصابة.
- **(FN)** False Negatives عددها **16** حالة فقط لم يكتشفها النموذج وهو معدل منخفض نسبياً، ويُعد مؤشراً إيجابياً في سياق التشخيص الطبي حيث يُفضل تقليل الحالات غير المكتشفة.

مناقشة النتائج:

تعكس النتائج المحققة فعالية الدمج بين:

- القوة التمثيلية العميقة لنموذج Xception
- دقة آلية الانتباه CBAM في إبراز السمات التشخيصية المهمة
- تقنيات تحسين الأداء مثل Fine-Tuning ، ضبط العتبة، وTTA

كما أن تحقيق قيمة AUC مرتفعة (0.9436) يشير إلى أن النموذج لا يعتمد فقط على نمط واحد من السمات، بل يمتلك قدرة قوية على التعميم والتمييز.

علاوة على ذلك، فإن الأداء المتوازن بين الفئتين، وعدم انحياز النموذج نحو فئة دون أخرى، يعززان من موثوقيته في التطبيقات الطبية الواقعية.

9. التوصيات :

بناءً على نتائج هذه الدراسة، يُوصى بما يلي:

- اعتماد مقاييس مثل AUC و Recall كمؤشرات أساسية في تشخيص ASD
- عدم الاكتفاء بالدقة (Accuracy) وحدها
- استخدام TTA في أنظمة التشخيص المعتمدة على الصور
- التركيز على تقليل False Negatives في التطبيقات الطبية
- تحسين جودة وتوازن مجموعات البيانات الطبية

13. المراجع :

[1]. Autism spectrum disorder (2023). Retrieved from World Health Organization: https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders

[2]. Autism Spectrum Disorder (ASD) (2023). Retrieved from Centers for Disease Control and Prevention: https://www.cdc.gov/ncbddd/autism/facts.html

[3]. Autism and Developmental Disabilities Monitoring (ADDM) Network (2023).Retrieved from Centers for Disease Control and Prevention: https://www.cdc.gov/autism/about/?CDC_AAref_Val

[4] .A. F. Rashid and S. H. Shaker ،
"Autism Spectrum Disorder Detection Using Face Features Based on Deep Neural Network"،Wasit Journal of Computer and Mathematical Science، vol. 2, no. 1, pp. 112–120, 2023. [Online].
Available: https://doi.org/10.31185/wjcm.100

[5]. T. M. Ghazal, S. Munir, S. Abbas, A. Athar, H. Alrababah, and M. A. Khan,"Early Detection of Autism in Children Using Transfer Learning,"Intelligent Automation & Soft Computing, 2023.
[Online].Available: https://doi.org/10.32604/iasc.2023.030125

[6]. M. A. Mareeswaran and K. Selvarajan,
"A Computational Intelligent Analysis of Autism Spectrum Disorder Using Machine Learning Techniques," IAES International Journal of Artificial Intelligence (IJ-AI),
vol. 13, no. 1, pp. 807–816, 2024. [Online].
Available: https://doi.org/10.11591/ijai.v13.i1.pp807-816

[7]. M. Ikermane and A. El Mouatasim,
"Web-Based Autism Screening Using Facial Images and Convolutional Neural Network,"
Indonesian Journal of Electrical Engineering and Computer Science,
vol. 29, no. 2, pp. 1140–1147, 2023. [Online].
Available: https://doi.org/10.11591/ijeecs.v29.i2.pp1140-1147

[8]. Goodfellow, I., Bengio, Y., & Courville, A.
Deep Learning, 1st ed.Cambridge, MA: MIT Press, 2016.
[Online]. Available: https://www.deeplearningbook.org

[9]. LeCun, Y., Bengio, Y., & Hinton, G.
Deep learning. *Nature*, vol. 521, pp. 436–444, 2015.
[Online]. Available: https://doi.org/10.1038/nature14539

[10]. O'Shea, K., & Nash, R.
An introduction to convolutional neural networks.
arXiv preprint arXiv:1511.08458, 2015.
[Online]. Available: https://arxiv.org/abs/1511.08458

[11]. Deeb, M., et al. 2024- Enriching NeuMF System with Full Data Features Using Attention-Based Deep Learning Models to Improve Ratings Prediction, Journal of Homs University vol. 46, no. 9 , pp. 85-128 [In Arabic]. https://journal.homs-univ.edu.sy/index.php/Engineering/article/view/6012

[12]. Dumoulin, V., & Visin, F.
A guide to convolution arithmetic for deep learning.
arXiv preprint arXiv:1603.07285, 2016.
[Online]. Available: https://arxiv.org/abs/1603.07285

[13]. Chollet, F.
Xception: Deep learning with depthwise separable convolutions.
Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1251–1258, 2017.
[Online]. Available: https://doi.org/10.1109/CVPR.2017.195

[14]. Shorten, C., & Khoshgoftaar, T. M.
A survey on image data augmentation for deep learning.
Journal of Big Data, vol. 6, no. 1, pp. 1–48, 2019.
[Online]. Available: https://doi.org/10.1186/s40537-019-0197-0

[15]. Zhang, K., Zhang, Z., Li, Z., & Qiao, Y.
Joint face detection and alignment using multitask cascaded convolutional networks.
IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1499–1503, 2016.
[Online]. Available: https://doi.org/10.1109/LSP.2016.2603342

[16]. Yun, S., Han, D., Oh, S. J., Chun, S., Choe, J., & Yoo, Y.
CutMix and MixUp: Data augmentation techniques for robust image classification. *arXiv preprint arXiv:1905.04899*, 2019
[Online]. Available: https://arxiv.org/abs/1905.04899

[17] .S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, “CBAM: Convolutional Block Attention Module,” *Proceedings of the European Conference on*

Computer Vision (ECCV), pp. 3–19, 2018. [Online]. Available: https://doi.org/10.1007/978-3-030-01234-2_1

[18]. Lin, M., Chen, Q., & Yan, S.
Network in network.
arXiv preprint arXiv:1312.4400, 2013.
[Online]. Available: https://arxiv.org/abs/1312.4400

[19]. Tan, M., & Le, Q.
EfficientNet: Rethinking model scaling for convolutional neural networks.
Proceedings of the International Conference on Machine Learning (ICML), pp. 6105–6114, 2019.
[Online]. Available: https://arxiv.org/abs/1905.11946

[20]. Kessis, N., et al. 2025-Real-Time Arabic Sign Language Recognition Using Deep Learning, Journal of Homs University vol. 47, no. 3 , pp. 53-86 [In Arabic]. https://journal.homs-univ.edu.sy/index.php/Engineering/article/view/7397