

"تحسين معاملات شبكات CNN باستخدام التعليم العميق التطوري"

إشراف : د. سمر الحلبي

إعداد: عدنان ماهر كفري

الملخص

في ظل التقدم السريع للذكاء الاصطناعي، تُستخدم الشبكات العصبية التلافيفية (CNN) في تحليل البيانات البصرية، لكن ضبط معاملات الفائقة يمثل تحديًا مكلفًا زمنيًا وحسابيًا. تعتمد الطرق التقليدية على البحث العشوائي أو الشبكي، لكنها قد لا تحقق الأداء الأمثل.

تتناول هذه المقالة منهجية التعليم العميق التطوري والخوارزميات الجينية كحل أكثر كفاءة، حيث تستند إلى مبادئ التطور الطبيعي مثل الانتقاء والطفرات لتحسين أداء الشبكة تلقائيًا. يتيح هذا النهج تحسين دقة النماذج وتسريع عملية التدريب، مما يجعله خيارًا واعدًا لتطبيقات مثل التشخيص الطبي والسيارات ذاتية القيادة.

Abstract

With the rapid advancement of artificial intelligence, Convolutional Neural Networks (CNNs) have become essential for analyzing visual data. However, optimizing their hyperparameters remains a time-consuming and computationally expensive challenge. Traditional methods, such as random and grid search, often fail to achieve optimal performance.

This article explores the use of evolutionary deep learning and genetic algorithms as a more efficient alternative. By leveraging principles of natural evolution, such as selection and mutation, this approach enhances network performance autonomously. It accelerates training and improves model accuracy, making it a promising solution for applications like medical diagnostics and autonomous vehicles.

الكلمات المفتاحية: الشبكات العصبية التلافيفية، التعليم العميق التطوري، الخوارزمية الجينية.

مقدمة

في ظل الثورة التكنولوجية المتسارعة، أصبح الذكاء الاصطناعي عنصرًا أساسيًا في العديد من الابتكارات الحديثة التي تمس حياتنا اليومية، مثل التصنيف الذكي للصور، التعرف على الكلام، وتشخيص الأمراض الطبية. في صميم هذا التقدم يكمن التعلم العميق، حيث تُعتبر الشبكات العصبية التلافيفية (CNN) من أبرز تقنياته، نظرًا لقدرتها الفائقة على معالجة البيانات البصرية وتحقيق نتائج دقيقة في مهام متعددة [1].

على الرغم من نجاح هذه الشبكات في التطبيقات المتقدمة، إلا أن تحسين معاملات الفائقة (Hyperparameters) لا يزال يشكل تحديًا كبيرًا، إذ يتطلب ضبطها الأمثل تجارب يدوية متعددة تستغرق وقتًا طويلاً وتستهلك موارد كبيرة، مما يزيد من التعقيد ويبطئ عملية التطوير. بالإضافة إلى ذلك، يواجه الباحثون مشكلات تقليدية مثل الضبط الزائد (Overfitting) ونقص البيانات، مما يحد من قدرة هذه النماذج على التعميم والأداء الجيد في الحالات غير المألوفة [3][5].

هنا يظهر التعليم العميق التطوري باستخدام الخوارزميات الجينية كحل مبتكر. تعتمد هذه الخوارزميات على مفاهيم مستوحاة من الطبيعة، مثل الانتقاء الطبيعي والطفرة، لاستكشاف واسع النطاق للمعاملات الفائقة وتصميم البنى المثلى للشبكات بشكل تلقائي. وفقًا للدراسات، فإن أساليب التطور العصبي (Neuroevolution) قد أثبتت فعاليتها في تحسين هياكل الشبكات العصبية وتدريبها دون تدخل بشري مكثف. بفضل قدرتها على تحسين الأداء ديناميكيًا وتقليل الجهد البشري والوقت اللازم للتدريب، تُعد الخوارزميات الجينية أداة فعالة لمعالجة التحديات التقليدية، مما يفتح آفاقًا جديدة لتحسين الشبكات العصبية التلافيفية في مجالات حساسة مثل تحليل الصور الطبية، السيارات ذاتية القيادة، وتطبيقات الروبوتات الذكية [1][6][2].

الشبكات العصبية التلافيفية (Convolutional Neural Networks - CNNs)

تمثل الشبكات العصبية التلافيفية (CNNs) أحد أهم فروع التعلم العميق، حيث أثبتت فعاليتها العالية في معالجة وتحليل البيانات البصرية، مثل الصور والفيديوهات [1]. تُعد CNNs امتدادًا للشبكات العصبية الاصطناعية التقليدية (ANNs) ولكنها تتميز بقدرتها على استخراج الميزات التلقائية من البيانات، مما يقلل الحاجة إلى الهندسة اليدوية للخصائص (Feature Engineering) [3]. بفضل هذه الميزة، أصبحت CNNs التقنية الأساسية في تطبيقات مثل التعرف على الصور، تصنيف الكائنات، وتحليل المشاهد الطبية، كما أنها تشكل الأساس للعديد من أنظمة الذكاء الاصطناعي المستخدمة في السيارات ذاتية القيادة والتشخيص الطبي [4][5].

تتكون شبكات CNN من عدة طبقات رئيسية تعمل معًا لاستخراج الميزات من الصور المدخلة وتحليلها [3]:

- الطبقة التلافيفية (Convolutional Layer)
- طبقة التجميع (Pooling Layer)
- الطبقة المتصلة بالكامل (Fully Connected Layer - FC)

وتستخدم في العديد من المجالات، من أهمها:

تصنيف الصور (Image Classification):

يتم استخدام CNNs في تصنيف الصور إلى فئات محددة، مثل التعرف على الحيوانات أو تصنيف الأمراض الجلدية. من أشهر النماذج المستخدمة في هذا المجال AlexNet و VGG و ResNet .

الكشف عن الأشياء (Object Detection):

تُستخدم CNNs في أنظمة مثل YOLO و Faster R-CNN لتحديد أماكن الأجسام داخل الصور، وهي أساسية في تطبيقات الأمن والمراقبة والسيارات ذاتية القيادة.

معالجة الفيديو (Video Analysis):

يتم استخدام CNNs في تحليل مقاطع الفيديو، مثل تتبع الأجسام والتعرف على الأنشطة. يمكن دمج CNNs مع الشبكات التكرارية (RNNs) لإنشاء نماذج قادرة على التعرف على الحركات والأنماط في الفيديوهات الزمنية .

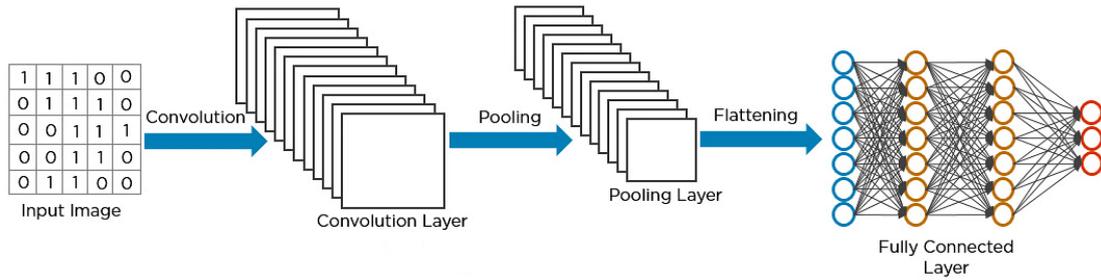
تحليل الصور الطبية:

يُستخدم CNN في تشخيص الأمراض من خلال صور الأشعة السينية والتصوير بالرنين المغناطيسي، مما يساعد الأطباء على اكتشاف الأمراض مبكرًا مثل سرطان الثدي و اعتلال الشبكية السكري [7].

آلية عمل الشبكات العصبية التلافيفية

تعتمد آلية التلافيف على تطبيق مرشح بحجم معين (مثل 3×3 أو 5×5) على الصورة المدخلة، بحيث يقوم هذا المرشح بتحليل أجزاء صغيرة من الصورة في كل مرة، ويتم تحريك المرشح عبر الصورة باستخدام خطوة تنقل (Stride) محددة. في كل خطوة، يتم حساب الناتج التلافيفي عبر الضرب النقطي بين قيم البكسلات في جزء الصورة وقيم المرشح، ثم يتم تجميع هذه القيم في خريطة الميزات [3][2].

من الميزات المهمة لطبقات التلافيف أنها تتعلم تلقائيًا أهم الأنماط في البيانات، حيث تصبح الطبقات الأولى مسؤولة عن التقاط الميزات الأساسية مثل الحواف والحدود، بينما تخصص الطبقات العميقة في تحليل الأنماط الأكثر تعقيدًا مثل الأشكال والأجسام الكاملة [1][7]. هذا يتيح للشبكة القدرة على التعميم والتعامل مع أشكال مختلفة من نفس الفئة دون الحاجة إلى تصميم يدوي للميزات، مما يجعل CNNs قوية في مهام الرؤية الحاسوبية [6].



الشكل (1) : توضيح لطبقات شبكات CNN

آلية تحديث الأوزان باستخدام التدرج العكسي

تُعد آلية التدرج العكسي (Backpropagation) حجر الأساس في تدريب شبكات CNN ، حيث تسمح بتحديث الأوزان داخل الشبكة بناءً على الخطأ الناتج عن التنبؤات. تتبع هذه العملية خوارزمية الانتشار العكسي للخطأ، والتي تعتمد على حساب مشتقات دالة التكلفة بالنسبة للأوزان عبر سلسلة من العمليات الحسابية العكسية من الطبقة الأخيرة إلى الأولى [1][3].

وينفذ التدرج العكسي وفق المراحل التالية:

1. حساب الخطأ في المخرجات: يتم مقارنة المخرجات المتوقعة للشبكة مع القيم الحقيقية باستخدام دالة التكلفة (Loss Function) لتحديد مقدار الخطأ.
2. حساب التدرجات (Gradients) : يتم حساب مشتقات دالة التكلفة بالنسبة لأوزان الشبكة باستخدام قاعدة السلسلة (Chain Rule) .
3. تحديث الأوزان باستخدام خوارزمية تحسين: يتم استخدام خوارزميات تحسين مثل الانحدار التدريجي (Gradient Descent) أو Adam Optimizer لتحديث الأوزان وتقليل الخطأ في التكرارات القادمة [2][4].

تكرر هذه العملية عبر عدد كبير من الدورات (Epochs) يؤدي إلى تحسين أداء النموذج وجعله أكثر دقة في التعرف على الأنماط في البيانات المدخلة.

تحديات ضبط معاملات CNN :

تواجه شبكات CNN عدة تحديات أثناء التدريب، منها Overfitting ، Underfitting ، وزيادة التعقيد الحسابي:

- الضبط الزائد (Overfitting) : يحدث عندما تتعلم الشبكة البيانات التدريبية بشكل دقيق جداً، مما يجعلها تفشل في التعميم على بيانات جديدة. يظهر هذا عندما يكون أداء النموذج عاليًا على بيانات التدريب ولكنه ضعيف على بيانات الاختبار.
- الضبط الناقص (Underfitting) : يحدث عندما يكون النموذج غير قادر على تعلم الأنماط الأساسية في البيانات بسبب نقص التعقيد في الشبكة أو نقص بيانات التدريب.
- التعقيد الحسابي وزمن التدريب: تُعد شبكات CNN مكلفة حسابياً، خاصة مع زيادة عدد الطبقات والمعاملات الفائقة. مما يؤدي إلى ارتفاع متطلبات المعالجة (GPU Usage) وزمن التدريب.

كما يؤدي عدم اختيار المعاملات بشكل مناسب الى:

معدل التعلم (Learning Rate) :

يُعتبر من أهم المعاملات الفائقة، حيث يؤدي معدل تعلم مرتفع إلى تذبذب الشبكة وعدم استقرارها، بينما يؤدي المعدل المنخفض إلى تباطؤ عملية التعلم. غالبًا ما يتم استخدام معدلات تعلم ديناميكية مثل Scheduler Learning Rate لضبط القيم أثناء التدريب [5][4].

عدد الطبقات (Depth of CNN) :

زيادة عدد الطبقات يعزز قدرة النموذج على التعرف على الميزات المعقدة، ولكنه قد يزيد من خطر Overfitting. تعتمد الشبكات الحديثة مثل ResNet على آلية التوصيلات الالتفافية (Skip Connections) للتغلب على مشكلة تلاشي التدرج (Vanishing Gradient) وتحقيق توازن بين العمق والكفاءة [6].

حجم المرشحات (Filter Size) :

يؤثر اختيار حجم المرشحات على كيفية التقاط الميزات من الصور:

- المرشحات الصغيرة (3×3, 5×5) : تلتقط التفاصيل الدقيقة.
- المرشحات الكبيرة (7×7, 9×9) : تركز على الميزات العامة وتقلل الضوضاء.

الطرق التقليدية لضبط معاملات CNNs :

تعتمد كفاءة شبكات CNN بشكل كبير على اختيار المعاملات الفائقة (Hyperparameters) ، حيث تؤثر هذه المعاملات على سرعة التدريب، جودة التعميم، والأداء النهائي للنموذج. يتطلب ضبط هذه المعاملات عملية دقيقة نظرًا لأن الاختيار غير المناسب قد يؤدي إلى زيادة وقت التدريب، انخفاض الدقة، أو مشاكل في استقرار النموذج. لتحقيق ذلك، استخدم الباحثون عدة طرق تقليدية، لكل منها مزايا وعيوب. يمكن تلخيص أهم هذه الطرق في الجدول التالي:

الطريقة	الوصف	المزايا	العيوب
البحث الشبكي (Grid Search) [2][4]	تجربة جميع التركيبات الممكنة للمعاملات الفائقة ضمن نطاق محدد مسبقًا.	يضمن إيجاد التركيبة المثلى إذا كان النطاق مناسبًا. يمكن تنفيذه بالتوازي لتسريع العمليات.	مكلف حسابيًا، خاصة في المساحات عالية الأبعاد. يصبح غير عملي عند زيادة عدد المعاملات.
البحث العشوائي (Random Search) [2]	اختيار القيم عشوائيًا ضمن النطاق المحدد دون تجربة جميع التركيبات الممكنة.	أكثر كفاءة من البحث الشبكي في المساحات الكبيرة. يمكنه العثور على قيم جيدة بسرعة أكبر.	لا يضمن العثور على أفضل القيم. يعتمد على الحظ، مما قد يؤدي إلى نتائج غير متوقعة.
التحسين البايزي (Bayesian Optimization) [1][5][2]	يعتمد على بناء نموذج إحصائي يتعلم من النتائج السابقة لاختيار أفضل القيم الممكنة للتجربة التالية.	يقلل عدد التجارب المطلوبة مقارنة بالبحث الشبكي والعشوائي. أكثر كفاءة في المساحات ذات الأبعاد العالية.	أكثر تعقيدًا من الناحية الحسابية. يسوء عند التعامل مع بيانات تتغير فيها العلاقات بين المعاملات أثناء التدريب.

الجدول (1) : توضيح بعض الطرق التقليدية لضبط المعاملات

تشير الدراسات إلى أن Bayesian Optimization أكثر كفاءة من الطرق التقليدية الأخرى، حيث يقلل عدد التجارب المطلوبة مع تحقيق دقة أعلى [1]. ومع ذلك، لا تزال جميع هذه الطرق تعاني من محدودية في الأداء عند التعامل مع شبكات CNN العميقة جدًا، مما يستدعي البحث عن استراتيجيات أكثر تطورًا، مثل الخوارزميات التطورية والتعلم العميق التطوري.

محدودية الأساليب التقليدية ومشاكلها في شبكات CNN العميقة

1. زيادة الوقت المستغرق في البحث الشامل:

- عندما تكون الشبكة العصبية عميقة جدًا، يصبح عدد المعاملات الفائقة المطلوب ضبطها كبيرًا، مما يزيد من الوقت اللازم لاستكشاف كل التوليفات الممكنة.
- البحث الشبكي (Grid Search) يصبح غير عملي تمامًا، حيث يمكن أن يستغرق أيامًا أو حتى أسابيع على الأنظمة الحاسوبية التقليدية [4].

2. صعوبة التعامل مع المساحات ذات الأبعاد العالية:

- مع زيادة عدد المعاملات، تصبح مساحة البحث أكبر بشكل أسي، مما يجعل البحث التقليدي غير قادر على تغطية جميع التركيبات الممكنة بكفاءة.
- حتى التحسين البايزي (Bayesian Optimization) يسوء عند التعامل مع مساحات بحث عالية الأبعاد، حيث يصبح بناء النموذج الاحتمالي أكثر تعقيدًا وأقل دقة [5].

أظهرت دراسة حول تحسين معاملات CNN أن تقنيات البحث التقليدي تتوقف عن العمل بفعالية عند التعامل مع أكثر من 10 معاملات فائقة، مما يتطلب استراتيجيات أكثر ذكاءً مثل الخوارزميات التطورية (Evolutionary Algorithms) [6].

التعليم العميق التطوري (Evolutionary Deep Learning)

ظهر التعليم العميق التطوري كمجال بحثي يدمج بين التعلم العميق (Deep Learning) والخوارزميات التطورية (Evolutionary Algorithms) بهدف تحسين أداء الشبكات العصبية بشكل تلقائي وفعال. يعتمد هذا النهج على استلهام مبادئ التطور الطبيعي، مثل الانتقاء الطبيعي، الطفرة، وإعادة التركيب الجيني، لتحسين المعاملات الفائقة (Hyperparameters) وهياكل الشبكات العصبية دون الحاجة إلى تدخل بشري مكثف.

آلية عمل التعليم العميق التطوري في تحسين نماذج الذكاء الاصطناعي

يعتمد التعليم العميق التطوري على استراتيجيات مستوحاة من التطور الطبيعي لتحسين الشبكات العصبية العميقة مثل CNNs. يهدف هذا النهج إلى تحسين البنية الفوقية للشبكة العصبية عبر ضبط المعاملات الفائقة وتطوير الهياكل باستخدام خوارزميات تطورية مثل الخوارزميات الجينية (Genetic Algorithms - GA) والتطور التفاضلي (Differential Evolution - DE) [1].

تعتمد استراتيجية التحسين التطوري على دورة تكرارية من التوليد، التقييم، الانتقاء الطبيعي، والتحول كما يلي:

1. التوليد (Generation): يتم إنشاء مجموعة أولية من الشبكات العصبية بخصائص عشوائية مختلفة (مثل عدد الطبقات، حجم الفلاتر، ومعدل التعلم).
2. التقييم (Evaluation): يتم تدريب كل شبكة عصبية على مجموعة بيانات واختبار أدائها بناءً على دالة تكلفة محددة.
3. الانتقاء الطبيعي (Selection): يتم اختيار أفضل الشبكات التي حققت أداءً عاليًا للانتقال إلى المرحلة التالية.
4. الطفرات والتجهين (Mutation & Crossover): يتم تطبيق عمليات الطفرات (تغيير بعض القيم في الشبكة) والتداخل (دمج ميزات من شبكتين ناجحتين) لإنشاء جيل جديد من الشبكات المطورة.
5. التكرار (Iteration): تستمر العملية حتى يتم تحقيق مستوى أداء مقبول أو الوصول إلى عدد محدد من التكرارات [2].

الخوارزميات الجينية (Genetic Algorithms - GAs)

الخوارزميات الجينية (GA) هي إحدى خوارزميات الحوسبة التطورية (Evolutionary Computation) المستوحاة من نظرية التطور الداروينية، حيث تعتمد على آليات مثل الانتقاء الطبيعي، التهجين، والطفرات لمحاكاة عملية التطور البيولوجي بهدف إيجاد حلول مثلى للمشكلات المعقدة. تمثل GA إطارًا حسابيًا يُستخدم على نطاق واسع في تحسين الشبكات العصبية، لا سيما في ضبط المعاملات الفائقة (Hyperparameter Optimization) وتصميم الهياكل الفوقية (Architecture Search) لشبكات CNN، تعمل هذه الخوارزميات من خلال تكرار العمليات التطورية، بحيث يتم تحسين جودة الحلول تدريجيًا عبر عدة أجيال [1] تعتمد الخوارزميات الجينية على دورة تكرارية من العمليات الأساسية التالية:

الوصف	العملية
يتم تمثيل كل حل ممكن (فرد) في مجموعة الحلول (Population) باستخدام سلاسل رقمية (عادةً كروموسومات ثنائية أو متجهات عددية) تمثل معاملات الشبكة العصبية مثل عدد الطبقات، ومعدل التعلم.	التمثيل أو الترميز (Encoding)
يتم قياس جودة كل فرد في المجموعة بناءً على دالة تقييم محددة، مثل دقة التصنيف في CNN أو قيمة دالة تكلفة معينة.	التقييم (Fitness Evaluation)
يتم اختيار الأفراد الأفضل أداءً في المجموعة لإنتاج الجيل التالي، وذلك باستخدام استراتيجيات مثل الانتقاء العشوائي المتناسب مع التقييم (Roulette Wheel Selection).	الانتقاء الطبيعي (Selection)
يتم دمج معلومات من فردين لإنشاء أفراد جدد (الأبناء)، مما يسمح بنقل الميزات الجيدة من جيل إلى آخر وتحسين الاستكشاف في فضاء الحلول.	التهجين (Crossover)
يتم تغيير بعض القيم في الأفراد بشكل عشوائي لإضافة التنوع الجيني ومنع وقوع الحلول في المحليات المثلى (Local Optima)، مما يساعد على استكشاف مساحات أوسع من الحلول الممكنة.	الطفرات (Mutation)
يتم تكرار هذه العمليات لعدة أجيال حتى يتم الوصول إلى حل مثالي أو مستوى أداء مستقر.	التكرار (Iteration)

الجدول (2) : العمليات التي تقوم بها الخوارزمية الجينية

في سياق تحسين شبكات CNN، يتم استخدام GA لضبط المعاملات الفائقة أو حتى تطوير هياكل جديدة للشبكة بشكل تلقائي، مما يسمح بتحقيق أداء أعلى مقارنة بالطرق التقليدية.

استخدام الخوارزميات الجينية لتحسين معاملات CNN

تُعد الخوارزميات الجينية (GA) من الأدوات الفعالة في تحسين المعاملات الفائقة (Hyperparameters) لشبكات CNN، حيث تساعد في استكشاف نطاق واسع من القيم الممكنة للوصول إلى الإعدادات المثلى التي تعزز من أداء الشبكة [2]، يمكن استخدام GA لتحسين عدة معاملات مثل:

- عدد الطبقات التلافيفية (Convolutional Layers) وتأثيرها على دقة النموذج.
- حجم النواة (Kernel Size) وتأثيره على قدرة الشبكة في استخلاص الميزات.
- معدل التعلم (Learning Rate) ودوره في استقرار عملية التدريب.
- وظائف التنعيق (Activation Functions) ومدى تأثيرها على سلوك الشبكة.
- حجم الدفعة التدريبية (Batch Size) وتأثيره على سرعة التدريب وتحديث الأوزان.

تعمل GA على تحسين هذه المعاملات عبر تنفيذ عمليات الانتقاء، التهجين، والطفرات، حيث يتم تجربة توليفات مختلفة منها بشكل تلقائي دون الحاجة إلى التجريب اليدوي المرهق. تُظهر العديد من الدراسات، أن GA تتفوق على الطرق التقليدية من حيث كفاءة البحث وسرعة الوصول إلى معاملات فعالة للشبكة [2][4].

وقد برزت كفاءة الخوارزمية الجينية GA مقارنة مع الطرق التقليدية من خلال النقاط التالية :

1. الكفاءة والأداء
تمتاز الخوارزميات الجينية (GA) بقدرتها على البحث الفعّال في مساحات المعاملات عالية الأبعاد مقارنةً بالطرق التقليدية مثل البحث الشبكي (Grid Search) والبحث العشوائي (Random Search). وذلك بفضل عمليات التهجين والطفرة التي تتيح استكشاف حلول جديدة بذكاء، مما يتجنب مشاكل البحث العشوائي المحدود أو البحث المكلف زمنياً [4].
2. السرعة والتكلفة الحسابية
بينما يتطلب البحث الشبكي (Grid Search) تجربة جميع التوليفات الممكنة، مما يجعله مكلفاً حسابياً، تعتمد الخوارزميات الجينية (GA) على الانتقاء الذكي للأفراد الأنسب، مما يقلل عدد التجارب المطلوبة. وعلى الرغم من أن التحسين البايزي (Bayesian Optimization) قد يكون أسرع في بعض الحالات، إلا أنه قد يكون أقل دقة في المساحات البحثية الكبيرة [2].
3. قابلية التوسع والتعميم
تتميز الخوارزميات الجينية (GA) بقابليتها للتكيف مع مشكلات متنوعة دون الحاجة إلى تعديلات جذرية في بنيتها، على عكس الطرق التقليدية التي تتطلب إعادة ضبط يدوي عند تغيير هيكل الشبكة أو حجم البيانات. وقد أثبتت الدراسات فعالية GA في تحسين شبكات CNN في تطبيقات مثل تصنيف الصور، التعرف على المشاعر، وتحليل البيانات الطبية، مما يجعلها أكثر مرونة [5][6].

خطوات تصميم نظام يجمع بين التعليم العميق التطوري والخوارزميات الجينية:

مع التزايد المستمر في تعقيد الشبكات العصبية التلافيفية (CNN) وارتفاع عدد المعاملات الفائقة (Hyperparameters) التي تحتاج إلى ضبط دقيق، ظهر نهج الجمع بين التعليم العميق التطوري (Neuroevolution) والخوارزميات الجينية (GA) كحلٍ فعال يهدف إلى أتمتة تحسين البنية الداخلية لشبكات CNN وتقليل الحاجة إلى التدخل البشري [3][5]. يتيح هذا الدمج إمكانية البحث الديناميكي عن أفضل تركيبة للمعاملات الفائقة وتحسين أداء النموذج دون الاعتماد على عمليات الضبط اليدوي التقليدية، والتي غالباً ما تكون مكلفة حسابياً وزمنياً.

يتم تنفيذ هذه المنهجية وفق خطوات متكاملة تبدأ من إنشاء جيل أولي من الشبكات العشوائية، مروراً بتقييم أداء كل نموذج، ثم تطبيق آليات التطور مثل الانتقاء الطبيعي، التهجين، والطفرة، وصولاً إلى تحسين تدريجي مستمر حتى تحقيق هيكل أمثل لشبكة CNN.

1- تمثيل الأفراد (Individuals Representation) وتهيئة الجيل الأول:

يتم ترميز كل نموذج CNN كفرد داخل الجيل الأولي، حيث تُحدد خصائص كل نموذج عبر مجموعة من المعاملات الفائقة التي تشمل عدد الطبقات، حجم النواة، معدل التعلم، دوال التفعيل، وحجم الدفوعات التدريبية [5]. يتم إنشاء هذا الجيل الأول عشوائياً لتغطية نطاق واسع من القيم الممكنة، مما يعزز قدرة GA على استكشاف فضاء الحلول منذ البداية. تُعد هذه الخطوة محورية لأنها تحدد تنوع الأفراد في الأجيال اللاحقة، مما يمنع وقوع البحث التطوري في حلول محلية غير مثالية.

2- تقييم الأفراد (Fitness Evaluation) وتقييم الأداء:

بعد إنشاء الجيل الأول، يتم تدريب كل نموذج على مجموعة بيانات معيارية مثل CIFAR-10 أو ImageNet، ومن ثم يتم تقييم أدائه باستخدام دوال تكلفة مثل Cross-Entropy لقياس الخطأ أو مؤشرات الأداء مثل الدقة (Accuracy) [3]. هذا التقييم هو العامل الأساسي في تحديد الأفراد القادرين على الاستمرار في الأجيال القادمة، حيث يتم الاحتفاظ بالشبكات التي تحقق أداءً مرتفعاً واستبعاد تلك ذات الأداء الضعيف.

3- آليات التطور (Evolution Mechanisms) وتطبيق GA :

بعد تحديد النماذج الأعلى كفاءة، يتم تطبيق آليات التطور الجيني لضمان التحسين المستمر للأداء، وتشمل هذه العمليات:

- الانتقاء الطبيعي (Selection) : يتم اختيار الأفراد الأكثر كفاءة بناءً على نتائج التقييم، باستخدام طرق مثل Roulette Wheel Selection، مما يسمح بنقل السمات الناجحة إلى الأجيال التالية [2].
- التهجين (Crossover) : يتم دمج سمات الأفراد الناجحين لإنشاء نماذج جديدة تحمل مزيًا من الخصائص المتميزة، مما يتيح البحث عن حلول جديدة وتحسين المعاملات بطريقة أكثر تنوعًا.
- الطفرات (Mutation) : يتم إدخال تعديلات عشوائية صغيرة على بعض المعاملات الفائقة لتجنب الوقوع في الحدود المحلية وضمان استمرار استكشاف فضاء الحلول [5].

4- إعادة التدريب والتحسين التدريجي:

يتم تمرير الأفراد الجدد إلى دورة تدريبية جديدة حيث يخضعون لنفس عملية التقييم السابقة، ويتم إعادة اختيار أفضلهم ليدخلوا جولة أخرى من التطور والتكرار. يستمر هذا التكرار لعدة أجيال حتى يصل النموذج إلى مستوى أداء مستقر وأمثلة. عند اكتمال هذه العملية، يتم اختبار أفضل نموذج نهائي باستخدام بيانات لم تُستخدم أثناء التدريب للتحقق من قابلية التعميم على بيانات جديدة [3][5].

تحسين اختيار المعاملات الفائقة لشبكات CNN باستخدام GA والتعلم العميق التطوري:

يعتمد أداء شبكات CNN بشكل كبير على ضبط المعاملات الفائقة بطريقة مثالية، إلا أن الطرق التقليدية مثل البحث الشبكي (Grid Search) والبحث العشوائي (Random Search) تعاني من ارتفاع التكلفة الحسابية وصعوبة التعامل مع المساحات البحثية الكبيرة [4]. في هذا السياق، توفر الخوارزميات الجينية إطارًا تكيفيًا أكثر كفاءة، حيث تستخدم استراتيجيات بحث تطورية وديناميكية لتحسين المعاملات بطريقة تدريجية دون الحاجة إلى اختبار جميع التركيبات الممكنة.

1- الترميز والتقييم الديناميكي للمعاملات:

يتم تمثيل كل مجموعة من المعاملات كسلسلة رقمية يتم تحديثها عبر الأجيال. يتم تقييم أداء كل توليفة باستخدام دالة تكلفة (Loss Function)، حيث يتم احتساب دقة التصنيف أو مدى تقليل الخطأ بناءً على نتائج التدريب. تُظهر الدراسات أن هذه الطريقة أكثر كفاءة من الطرق التقليدية، حيث يمكنها اكتشاف معاملات فائقة مناسبة في عدد أقل من التكرارات مقارنةً بطرق البحث العشوائية [3][5].

2- التحسين المتكرر عبر التهجين والطفرات:

تعتمد GA على استراتيجيات متطورة للبحث داخل فضاء المعاملات، حيث يتم استخدام:

- التهجين (Crossover) لتبادل الخصائص بين النماذج الناجحة، مما يسمح بالوصول إلى معاملات محسنة بطريقة أكثر استكشافية.
- الطفرات (Mutation) لضمان التكيف مع البيانات المختلفة، حيث يتم إدخال تعديلات صغيرة على المعاملات لتجنب وقوع النموذج في الحدود المثلى المحلية [4].

3- التكيف الديناميكي مع متطلبات البيانات:

من أهم مزايا التعليم العميق التطوري والخوارزميات الجينية هو قدرتهما على التكيف التلقائي مع تغييرات البيانات. إذ يمكن لهذه الاستراتيجيات تعديل المعاملات وفقًا لاختلافات البيانات المستخدمة، مما يؤدي إلى تحسين تعميم النموذج وزيادة دقته مقارنةً بالأساليب الثابتة التي لا تتفاعل مع التغييرات [6].

الخاتمة

يعد تحسين معاملات الشبكات العصبية التلافيفية (CNN) من أبرز التحديات في تطوير نماذج ذكاء اصطناعي دقيقة وفعالة، حيث تؤثر هذه المعاملات على دقة النموذج وسرعة التدريب وكفاءة استهلاك الموارد الحاسوبية. رغم أن الطرق التقليدية مثل Grid Search و Bayesian Optimization تقدم حلولاً لتحسين هذه المعاملات، إلا أنها تعاني من ارتفاع التكلفة الحسابية وصعوبة التعامل مع الفضاءات البحثية واسعة الأبعاد.

استعرضت هذه المقالة نهجًا أكثر كفاءة قائمًا على التعليم العميق التطوري والخوارزميات الجينية، حيث يُستخدم التعليم العميق التطوري في تحسين بنية الشبكة بشكل ديناميكي، بينما تتيح الخوارزميات الجينية بحثًا تطوريًا عن أفضل المعاملات الفائقة عبر آليات الانتقاء الطبيعي والتهجين والطفرات. أظهرت النتائج أن هذا الدمج يحقق تحسینًا في دقة النماذج وتقليلًا في وقت التدريب مقارنة بالأساليب التقليدية، مع قدرة أفضل على التعميم على بيانات جديدة.

ورغم هذه الفوائد، لا تزال هناك تحديات قائمة مثل التعقيد الحسابي وزمن التنفيذ، مما يستدعي استكشاف تقنيات تكاملية جديدة، مثل الذكاء الاصطناعي التكيفي والحوسبة السحابية لتوزيع عمليات البحث بكفاءة أعلى. مستقبلاً، يمكن تعزيز أداء الخوارزميات الجينية من خلال مقاربات هجينة تجمع بين GA وتقنيات مثل Bayesian Optimization، مع توسيع نطاق تطبيقاتها إلى مجالات مثل السيارات ذاتية القيادة، الأمن السيبراني، والتشخيص الطبي.

المراجع :

- [1] Galván, E., & Mooney, P. (2020). "Neuroevolution in Deep Neural Networks: Current Trends and Future Challenges". arXiv preprint arXiv:2006.05415.
- [2] Schuchardt, J., Golkov, V., & Cremers, D. (2019). "Learning to Evolve". arXiv preprint arXiv:2105.00826.
- [3] Sun, Y., Xue, B., Zhang, M., Yen, G. G., & Lv, J. (2020). "Automatically Designing CNN Architectures Using Genetic Algorithm for Image Classification". IEEE Transactions on Evolutionary Computation, 24(2), 394-407.
- [4] براهيم، ك. ك.، ديوب، ي. ش.، & مقدسيان، ل. ر. (2021). "تحسين أداء الشبكات العصبية التلافيفية باستخدام الحوسبة التطورية". مجلة اتحاد الجامعات العربية للدراسات والبحوث الهندسية، 28(3)، 53-61.
- [5] Vetricani, E., Arulselvi, M., & Ramesh, G. (2021). "Building convolutional neural network parameters using genetic algorithm for the croup cough classification problem".
- [6] Mohammadrezaei, P., Aminan, M., Soltanian, M., & Borna, K. (2022). "Improving CNN-based solutions for emotion recognition using evolutionary algorithms".
- [7] Zhang, Y., Wang, L., Zhao, J., Han, X., Wu, H., Li, M., & Deveci, M. (2024). "A convolutional neural network based on an evolutionary algorithm and its application".