

## دراسة وتحليل خوارزميات التفرع في التعرف على الوجوه

الدكتورة: ربي خيجه - قسم الفيزياء - كلية العلوم - جامعة اللاذقية

الملخص:

اتجهت التقنيات الحديثة في الوقت الحالي إلى رفع أداء النظم الحاسوبية والمهام التي تقوم بها، كما تسببت زيادة استخدامات الشبكات ومواقع التواصل ومخدمات تزويد المكتبات والمواقع الحكومية إلى زيادة كم البيانات التي تتواجد على المواقع البعيدة إلى زيادة حالات الضغط على تلك المواقع مما يسبب المزيد من التأخير الزمني للوصول إلى البيانات المطلوبة.

اتجهت العديد من الدراسات إلى استخدام تقنيات الذكاء الصناعي والشبكات العصبية بهدف تصنيف البيانات على المخدمات حسب النوع أو المحتوى بهدف تسريع عمليات التنقيب على البيانات وزيادة الدقة في عمليات البحث، ومن ثم اتجهت الدراسات إلى اعتماد واستخدام نظم البرمجة التفرعية التي تهدف إلى توزيع عمل البيانات على معالجات أو بطاقات إظهار مختلفة.

أثبتت الشبكات العصبية قدرتها على حل العديد من القضايا، ومع ازدياد حجم القضايا باتت تحتاج إلى شبكات ذات حجم أكبر وتتطلب عمليات معالجة أكبر بالإضافة إلى زمن تدريب أطول، وبالتالي أصبحت البرمجة التفرعية واستخدامها من أجل عمليات التعلم مهمة رئيسية للحصول على تطبيقات فعالة. إلا أن اعتمادها الرئيسي على سمات خوارزميات التعلم والبنى الصلبة المتاحة بجوارها تطلب استخدام تقنيات مثل شبكات RNN(Recurrent Neural Networks) اعتماداً على التوازي في عمليات التعلم، تم الاعتماد على تقنيتين رئيسيتين وهما قطاعات المعالجة Computing Clusters اعتماداً على مكتبات MPI, OpenMP والأخرى على Graphics processing باستخدام مكتبات CUDA حيث حقق كلا النموذجين الأداء المطلوب من الخوارزميات التفرعية.

كلمات مفتاحية: شبكات RNN، معايير الأداء، البرمجة التفرعية.

## **Study and Analysis of Branch Prediction Algorithms in Face Recognition**

Abstract:

Modern technologies have recently shifted towards enhancing the performance of computing systems and the tasks they carry out. This is primarily due to the increasing use of networks, social media, and government websites, leading to a surge in data volume. As a result, the load on these websites has increased, causing delays in data access, which in turn necessitates the development of new approaches to manage this growing demand.

Several studies have focused on utilizing artificial intelligence techniques and deep neural networks to classify data across different content types, aiming to accelerate processing and improve accuracy in data search operations. These studies have also examined the development of branch prediction systems, which aim to distribute tasks across different data streams more efficiently.

Deep neural networks have proven effective in solving various problems. As the size and complexity of data grow, larger networks are required, which in turn demand more processing time and memory. Consequently, branch prediction systems have become a crucial tool to optimize processing, and machine learning algorithms have become key to achieving high-performance applications. However, these systems often require the use of techniques such as Recurrent Neural Networks (RNN), which specialize in sequential data prediction, and rely on advanced techniques for processing, such as MPI and OpenMP.

The research also discusses how computing clusters and Graphics Processing Units (GPUs), particularly CUDA technology, can be used to enhance the performance of these branch prediction algorithms.

Keywords: RNN Networks, Performance Metrics, Branch Prediction Algorithms.

## ١. مقدمة:

اتجهت صناعة الحواسيب التجارية في الفترة الأخيرة إلى المعالجة التفرعية والموزعة. وقد سببت محدوديات صناعة أنصاف النواقل إلى المزيد من المحدوديات في تطور تلك الصناعة، مع التطور الحاصل في تطوير انصاف النواقل وبنيتها وحجمها ضمن الشبكات المختلفة تم الوصول إلى المزيد من القدرات والأداء بالإضافة على تطوير ملحوظ في أداء الخوارزميات مع التوفير الإضافي للوقت والمال.

في الوقت الحالي، تعتبر القطاعات clusters المؤلفة من الحواسيب المتباعدة بنية تقليدية منتشرة بقوة للاستخدام في عمليات البرمجة التفرعية حي تنتم العمليات المختلفة على عقد حوسبة ضمن القطاع وتتصل مع بعضها البعض من خلال الرسائل. تعتبر واجهات تمرير الرسائل (MPI) message passing interface الاستخدام المعياري المنتقل لهذه القيمة حيث تؤمن عدة بنى تجريدية والتي تبسط من استخدام الحواسيب التفرعية مع الذاكرة الموزعة [1].

في الوقت الحالي، أدى تطوير وحدات معالجة بطاقات الإظهار (GPUs) graphics processing units إلى رفع أداء نظم البرمجة التفرعية والتي يمكن ان تطبق في معظم نماذج بطاقات الإظهار [2]، يوجد العديد من التقنيات التي يمكن اعتمادها في هذا المجال ولعل أشهرها المرفقات المقدمة من شركة Nvidia والمعروفة باسم compute unified device architecture (CUDA) حيث تتضمن تطويرات برمجية باستخدام C/C++ ومكتبات للتوابع وميكانيكية صلبة مجردة والتي تقوم بإخفاء وحدات بطاقات الإظهار GPU عن المستخدمين [3].

تحصل بذلك الخوارزميات المختلفة والتي تعتمد على البرمجة التفرعية على العديد من المزايا بالاعتماد على تلك البنى التفرعية وذلك بسبب الخصائص التي تتميز بها تلك الشبكات التفرعية [4][5][6].

يعتبر تدريب الشبكات العصبية عملية بالغة التعقيد والصعوبة بالإضافة لحاجتها إلى الزمن وخصوصاً في الحالات التي تكون فيها الشبكات العصبية كبيرة الحجم أو او تحتاج إلى بيانات تدريب كبيرة كدخل خرج. تمتلك كل شبكة عصبية خصائصها التي تتفرد بها عن الشبكات الأخرى. كما ان الشبكات العصبية، حتى في حالة التشابه، وفي حال امتلاكها محددات مختلفة ومجموعات تدريب مختلفة قد تتمتع بسلوك مختلف أيضاً ضمن نفس الكيان الصلب التفرعي. وبالتالي فإن الحصول على حلول وخصائص عامة يعتبر مستحيلاً لكي يكون هنالك خصائص أداء ذات معايير ثابتة.

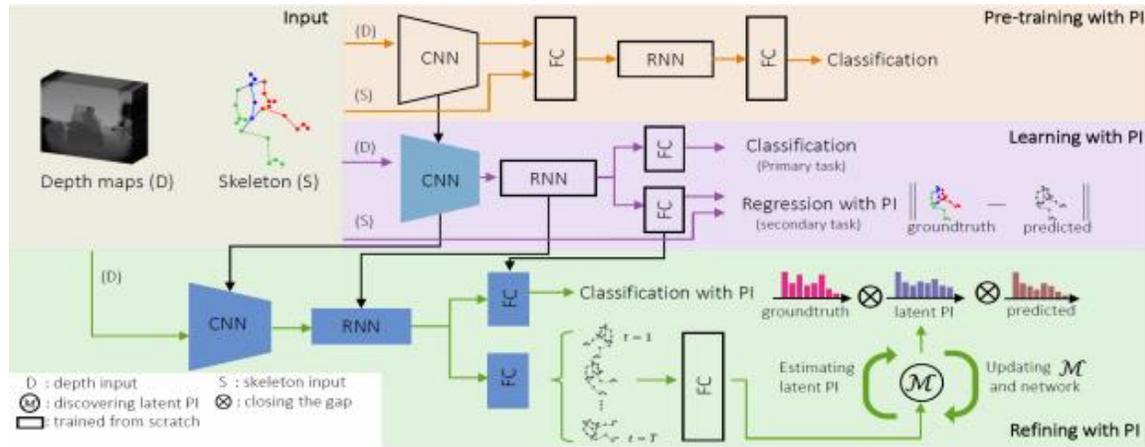
قدم Jiang Wang, Yi Yang, Junhua Mao, Zhiheng Huang, Chang Huang, Wei Xu دراسة تهدف إلى من أجل عنونة المحتوى الخاص بالصور الرقمية باستخدام الشبكات العصبية CNN-RNN وقد بينت هذه الدراسة ان أداء الشبكة يختلف من ناحية دقة الخرج وفقاً لنوع المصنفات المستخدمة وتوابع التعتيب المعتمدة حيث استخدم الباحثون توابع Soft Max و Knn و WARP، كما وضح الباحثون ارتفاع دقة النظام الكلية مع ازدياد حجم الصور في قاعدة البيانات [7]. اعتمدت التقنيات

## دراسة وتحليل خوارزميات التفرع في التعرف على الوجوه

الآخري من أجل عنونة الصور باستخدام الشبكات العصبية على [8] support vector machines و neural networks [9] حيث تتم العمليات على البكسلات اللونية بصورة رئيسية. في دراسات أخرى تم إدخال عدد من البكسلات المجاورة على عمليات التحليل بهدف رفع الدقة وزيادة النعومة المكانية اخرى التصنيف الناتجة [10]. تم أيضاً اعتماد سمات التصميم اليدوية [11].

قام [12] Chen et al. بتصميم شبكة RNN والتي تعتمد على مرشح التحويل المكاني [13] وهو مرشح يقوم بزيادة حدة الحواف في خرائط التصنيف الناتجة

وضح كل من Pichao Wang, Wanqing Lia, Philip Ogunbonaa, Jun Wang, Sergio Escalera دراسة تعتمد على استخدام الشبكات من نمط RNN بهدف التنبؤ بحركة الانسان اعتماداً على تسلسل من الصور ضمن الفيديو واستخدام البرمجة التفرعية، اعتمدت هذه الخوارزمية على بيانات من نمط RGB-D والتي تعتبر أكثر فعالية وأقل كلفة في هذه المهام وهي صور لونية تعطي قدرة على التنبؤ بالعمق، حيث تم اعتماد شبكة نموذجية بالشكل (1) الموضح:

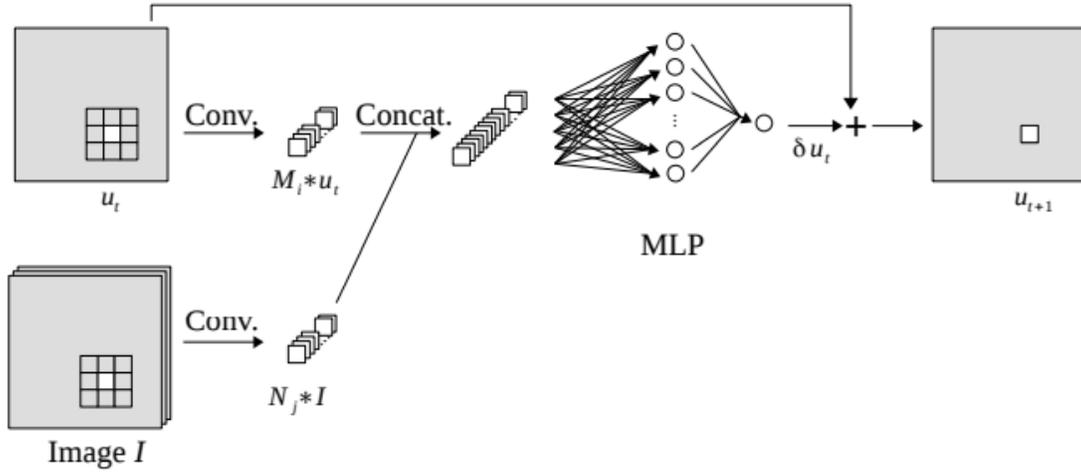


الشكل (1) بنية شبكة RNN المستخدمة في تصنيف الصور من نمط RGB-D

اوضحت النتائج قدرة النظام على التنبؤ بالحركة بدقة ما بين 60-80 بالمئة وهو مايسمح بتطبيقه في العديد من التطبيقات في الحياة العملية. [8]

استخدم Emmanuel Maggiori, Guillaume Charpiat, Yuliya Tarabalka, Pierre Alliez الشبكات من نمط RNN من أجل تصحيح صور الاقمار الصناعية وتصحيحها، ونحتاج إلى قيادة وتحكم بعدد قليل من البارامترات على كجزء من محددات الشبكة.

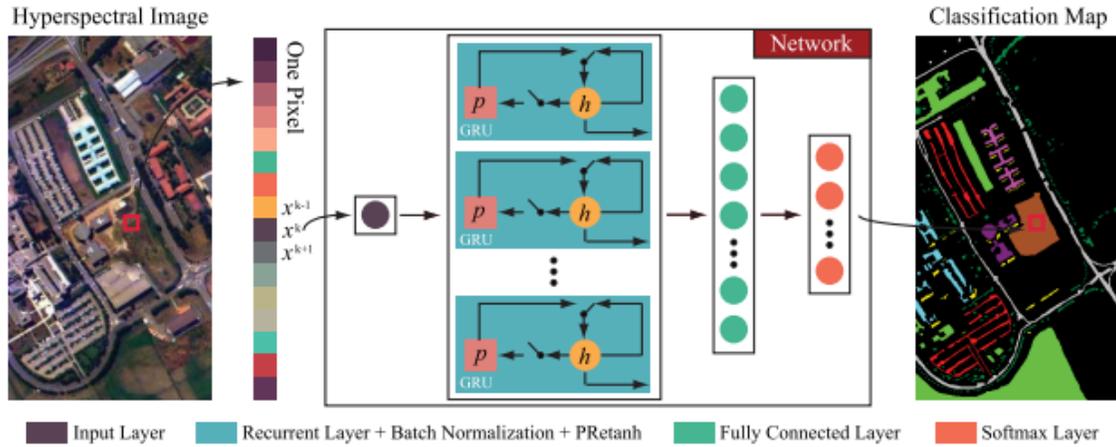
يوضح الشكل (2) التالي المخطط العملي المعتمد من قبل الخوارزمية:



الشكل (٢) بنية شبكة RNN بهدف تحسين صور الأقمار الصناعية

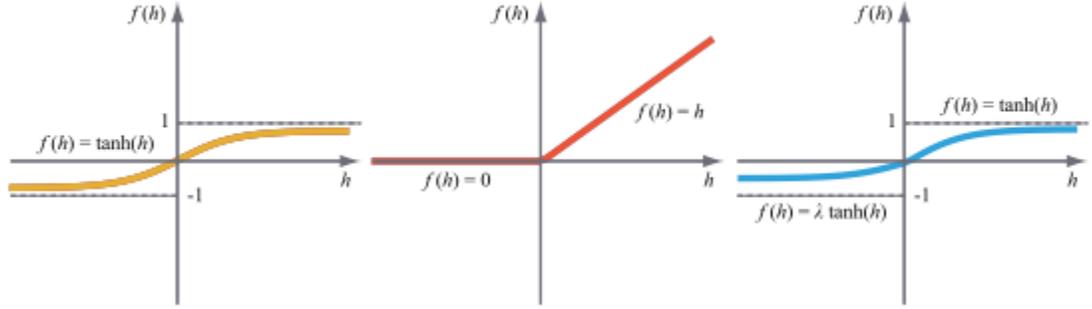
يتم استخلاص السمات من كلا الصورتين، الدخل  $I$  والخريطة الحرارية  $heat map$  التابعة للتكرار السابق ومن ثم يتم دمجها وإدخالها إلى MLP والتي تقوم بحساب تعديل الوزن حيث يتم تعديل صورة الخريطة الحرارية بالقيم الجديدة. أوضحت النتائج ارتفاعاً ملحوظاً بدقة التصنيف مقارنة مع الخوارزميات الأخرى، كما أن رفع الدقة يزداد باستخدام عمل مشترك بين شبكات CNN, RNN.

قدم Xiao Xiang Zhu, Lichao Mou, Pedram Ghamisi دراسة تعتمد على الشبكات من نمط RNN من أجل كشف تغيرات التضاريس الأرضية، تم توزيع المهام على مجموعة من وحدات Graphics كما في الشكل (٣) التالي:



الشكل (٣) بنية شبكة RNN والمجزئة إلى وحدات غرافية GRU بهدف التعرف وتصنيف المواقع

تم الاعتماد على توابع تفعيل مختلفة وملاحظة الفارق بالاداء عند استخدام توابع التفعيل تلك، تعتبر توابع  $\tanh, \text{ReLU}$ ,  $\text{PRetanh}$  الاكثر استخداما واعتمادية في شبكات RNN.



الشكل (٤) توابع التفعيل المعتمدة في شبكات RNN

أوضحت النتائج ارتفاع نسبة الخطأ عند استخدام تابع Relu مقارنة مع التوابع الأخرى.

## ٢. أهمية البحث وأهدافه:

تتنوع النماذج التي تستخدم فيها الشبكات العصبية ويعطي هذا الاختلاف العديد من الفوارق بالأداء، ولعل أهم معايير الأداء التي تؤخذ بعين الاعتبار هي معايير الدقة وسرعة التنفيذ. تتراوح الشبكات العصبية المستخدمة في مجال تصنيف وتمييز الصور الرقمية من الشبكات ذات التمرير الأمامي أو ما يعرف بشبكات البرسبترون مروراً بالشبكات التكرارية RNN والشبكات الالتفافية CNN.

تهدف الدراسة إلى تسريع أداء الشبكات التكرارية RNN والشبكات الالتفافية CNN في تصنيف الوجوه في الصور الرقمية بغية بناء نظام آمن للتعرف على المطلوبين وتقييم أداء هذا النظام من جهة معايير الدقة والسرعة وذلك لانتقاء الخوارزمية القابلة للاستخدام في مجال البرمجة التفرعية واستخدام المصادر الحاسوبية المختلفة الخاصة على أقصى مجال لتحقيق الدقة المناسبة في زمن مثالي.

تختلف الطرق التي يتم فيها إجراء عمليات البرمجة التفرعية في نظم المراقبة والتعرف والهدف الرئيسي هو التعرف على العناصر الموجودة في الصورة، يعتمد العمل على تقنيتين رئيسيتين وهما تعدد الخيوط Multi threading وتعدد العمليات multiprocessing.

يهدف البحث إلى دمج تقنيات البرمجة التفرعية مع خوارزميات الشبكات العصبية للحصول على نظام مسرع من أجل التعرف على الأشخاص ضمن الفيديو، سيتم قراءة الفيديو على شكل threading في حين سيتم تشغيل الشبكات العصبية بصورة تفرعية للوصول إلى الاداء الامثل للنظام.

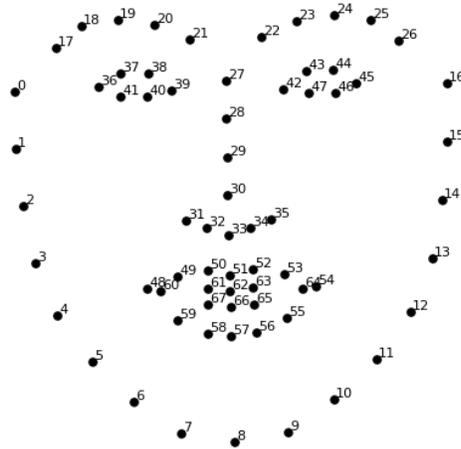
## ٣. طرائق البحث ومواده:

يصف هذا البحث المنهجية المعتمدة لزيادة سرعة أداء نظام أمني للتعرف على وجوه المطلوبين باستخدام الشبكات العصبية والبرمجة تفرعية بنماذجها المختلفة لاختيار الخوارزمية الأفضل التي تحقق السرعة الأعلى ضمن مجموعة الموارد المتاحة.

#### ٤. حالة الدراسة:

##### الحالة المدروسة Case Study: التعرف على الوجوه:

يعتمد بشكل عام في مجال التعرف على الوجوه على تقنيات كشف الوجوه والتعرف على الوجوه، تعمل هذه الخوارزمية على إيجاد ٦٨ نقطة مهمة في الوجه، تسمى هذه النقاط بالنقاط الفاصلة Landmarks وتتواجد أعلى الدقن - الحافة الخارجية للعين - الحافة الداخلية للعين - الحاجب - حول الفم - حول الأنف . يتم إيجاد هذه النقاط من خلال تدريب خوارزمية تعلم آلة. بعد تحديد سمات الوجه وتخزينها ضمن قاعدة بيانات بحيث تتبع كل مجموعة من السمات لشخص محدد يتم في المرحلة التالية تمرير تلك البيانات إلى الشبكة العصبية والقيام بعملية التدريب وربط كل مجموعة من السمات مع شخص محدد.



بعد ذلك يتم إجراء الاختبار للبيانات من خلال صور جديدة يتم إدخالها للشبكة وملاحظة الدقة في التعرف على الوجه بالإضافة إلى سرعة التعرف الناتجة عن النظام المصمم. يتم ذلك من خلال بناء وتدريب شبكة الطّي العصبونية على الكثير من صور الوجوه مهمتها استخراج شعاع مكون من ١٢٨ بُعد لكل وجه ويسمى بشعاع التضمين Embedding، الفائدة منه اختزال الصورة ببياناتها الكبيرة إلى بعض الأرقام القليلة بالتالي بدلاً من مقارنة الصورة بشكل كامل مع الصور الموجودة في قاعدة البيانات، يتم مقارنة شعاع التضمين لصورة الوجه المُدخلة مع أشعة التضمين الخاصة بكل صورة من صور الوجوه الموجودة في قاعدة البيانات.

## دراسة وتحليل خوارزميات التفرع في التعرف على الوجوه

يتم معرفة اسم الشخص من خلال تدريب مصنف آلة متجه الدعم Support Vector Machine مهمته التعرف على الشخص من خلال إيجاد شعاع التضمين الأقرب للصورة المُدخلة من بين أشعة التضمين الخاصة بصور الوجوه الموجودة في قاعدة البيانات، تحدث هذه العملية بسرعة كبيرة بأقل من 1 ثانية.

### ٤,١ . مكتبة TensorFlow:

طورت شركة غوغل في ٢٠١٥، Tensorflow وهو مكتبة مفتوحة المصدر اعتمد في البداية الباحثون كلغة برمجة لبناء العديد من النماذج model والتي كانت مبنية بالأساس على خوارزميات التعلم الآلي . يستخدم حالياً في العديد من الأبحاث ومختلف المجالات كالعلوم الاجتماعية والطب والهندسة والاقتصاد، كذلك يستخدم في مختلف أنواع مهام الفهم الإدراكي واللغوي كالاعتراف على الكلام والنص والصور والوجوه ، أيضاً عمليات التصنيف والتمييز .

يجعل Tensorflow من الأسهل على المطورين تصميم وبناء وتدريب نماذج التعلم العميق. هناك العديد من الأطر الأخرى المتاحة للمطورين ، ولكن Tensor Flow هي أكثر إطار استخداماً على نطاق واسع فيما بينها ، وقد تم اعتماده من قبل العديد من العملاقة مثل Airbus و Twitter و IBM وغيرها - ويرجع ذلك أساساً إلى هيكل النظام المرنة للغاية.

إن الفائدة الأكبر التي توفرها TensorFlow لتنمية تعلم الآلة هي التجريد. وبدلاً من التعامل مع التفاصيل الدقيقة لتنفيذ الخوارزميات ، أو اكتشاف الطرق المناسبة لمخرجات دالة واحدة إلى مدخلات أخرى ، يمكن للمطور التركيز على المنطق العام للتطبيق .

يوفر TensorFlow وسائل راحة إضافية للمطورين الذين يحتاجون إلى التصحيح واكتساب الاستبطان في تطبيقات يتيح لنا TensorFlow وضع التنفيذ المتدرج لتقييم كل عملية رسم بياني وتعديلها بشكل منفصل وشفاف ، بدلاً من إنشاء الرسم البياني بأكمله ككائن معتم واحد وتقييمه دفعة واحدة. نتيح لنا مجموعة التصور في TensorBoard إمكانية فحص وطريقة عرض الرسوم البيانية من خلال لوحة تحكم تفاعلية تستند إلى الويب.

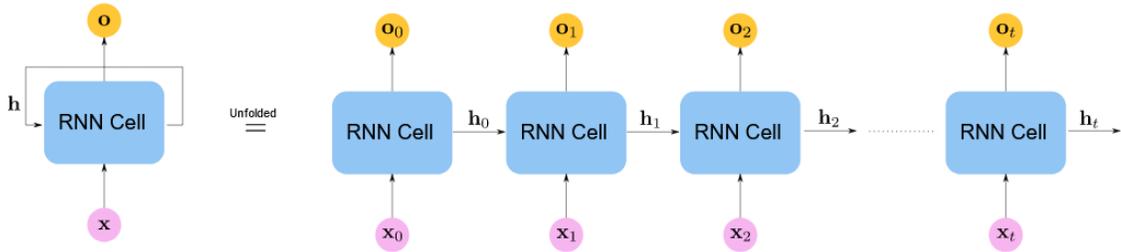
على الرغم من أن TensorFlow ليس مشروعاً رسمياً من نوع Apache ، إلا أنه كان مفتوح المصدر منذ بضع سنوات من قبل مطور البرامج Google LLC والذي يستمر في تحقيق استثمارات عميقة في إطار العمل.

### ٤,٢ . الشبكات العصبية الملتفة Recurrent Neural Networks:

تعتبر شبكات RNN إحدى أهم شبكات التعلم العميق من أجل معالجة البيانات ذات النمط التتابعي، اعتبر هذا النمط من الشبكات الأفضل في معالجة البيانات التتابعية ضمن فترة من الزمن حيث يحتاج النموذج العميق ذو التغذية الأمامية لمجموعة من

البارامترات لكل معامل في التتابع. بالإضافة إلى ذلك، قد لا يكون من الممكن التعميم لتحقيق التتابع على طول المتحولات المدروسة.

تقوم هذه الشبكات بتطبيق الوزن نفسه على جميع المعاملات في التتابع كما تقوم بتقليل عدد المعاملات وتسمح للنموذج بالتعميم على كامل طول المتحولات ضمن التتابع. تسمح RNN بتعميم النموذج إلى بنى المعطيات أكثر من البيانات التتابعية كما في البيانات المكانية أو الجغرافية.



الشكل (٥) الشبكات من نمط Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN)

وتعتبر الشبكات من نمط Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN) إحدى أقوى أدوات التصنيف الديناميكية المعتمدة، يركز نموذج تعلم الآلة على تطوير الخوارزميات التي تسمح بتطوير الأداء بشكل ذاتي ديناميكي اعتماداً على التدريب. أي كلما زادت عمليات التدريب للخوارزمية سيكون أداء هذه الخوارزمية أكثر فعالية، تتم تلك العملية من خلال خلق تابع تصنيف من بيانات التدريب المتاحة. يتم بعد ذلك قياس الأداء للمصنف المصمم عن طريق تطبيق البيانات غير الملاحظة ومطابقة الخرج.

تم استنباط الشبكات العصبية الصناعية (ANN) Artificial Neural Networks من نظم التعلم البيولوجية والتي تتمثل بشبكة معقدة من العصبونات، تعتبر العصبونات الوحدات الأبسط في الشبكة والتي تستقبل شعاعاً من قيم الدخل وتقوم بتهيئة قيمة خرج وحيدة. تعتبر الشبكة العصبية ذات التمرير الأمامي النموذج الأكثر انتشاراً حيث يتم ترتيب العصبونات على شكل طبقات وهي طبقة الدخل، طبقة الخرج، بالإضافة إلى طبقة خفية واحدة على الأقل.

تعتبر شبكات التمرير الأمامي محدودة الأداء فقط لعمليات التصنيف من النمط Static وبالتالي فهي تعاني من المحدودية على تأمين العنونة static mapping ما بين الدخل والخرج. من أجل نمذجة التنبؤ الزمني فإننا بحاجة إلى المصنفات الديناميكية.

من أجل تحويل نموذج التغذية الأمامية إلى نموذج ديناميكي فإننا بحاجة إلى إعادة تمرير الإشارات من العينات الزمنية السابقة timesteps من جديد إلى الشبكة، وتسمى هذه الشبكات مع وصلات من نمط recurrent بشبكات Recurrent Neural Network (RNN)، هذه الشبكات محدودة في البحث عن العينات الزمنية بالمجال السابق بما لا يتعدى عشر عينات زمنية فقط

والسبب في ذلك كون إشارة التغذية الخلفية تتعرض للتلاشي وهو ماتم الإشارة إليه بالنمط Long Short-term Memory Recurrent Neural Network(LSTM-RNN)

تستطيع الشبكات من هذا النمط بالتعلم حتى ١٠٠٠ عينة زمنية timesteps وذلك وفقاً لدرجة التعقيد الذي قد تحتويه الشبكة. تسمح هذه البنية التسلسلية للشبكات التكرارية بإضافة بعد جديد من التعقيد إلى عمليات البث الخلفي، تمتلك كل حلقة في شبكات RNN على زوجها الخاص من الدخل-خرج بينما تتشارك جميع الحلقات بنفس الأوزان وبالتالي يجب تحديد الفترة التي يجب أن تتعدل فيها الأوزان.

تم عنونة هذه المشكلة باسم (BPTT) back-propagation through time حيث يتم إجراء هذه العملية بعد عملية البث الامامي لتتابع من المداخل ويتم تحديث أوزان شبكة RNN كما يلي:

١- يتم بسط الشبكة كسلسلة من شبكات ANN حيث تمثل كل شبكة منها نسخة من شبكة RNN الاساسية، تتشارك هذه الشبكات بنفس الأوزان وتوابع التفعيل.

٢- تعمل BPTT بشكل خلفي عبر التسلسل المدروس ويتم حساب تدرج الضياع والخطأ عبر كل ANN في السلسلة.

٣- يتم إعادة بسط الشبكة للوضع الأصلي وتعديل الأوزان.

٤,٣ . محدوديات شبكات RNN:

تكافح ذاكرة العمل working memory في الشبكات التكرارية التقليدية للحفاظ على اعتماديات الاصطلاحات بشكل أطول والسبب الاساسي لهذه المشكلة هي التلاشي المتدرج، ويمكن أن تسبب المشاكل عندما تتضمن تسلسلات الدخل المبكرة معلومات هامة. تعرف هذه المشكلة بشكل واسع في عمليات البث الخلفي والتدرج المتناقص ويمكن ملاحظتها عند تدريب شبكات RNN وشبكات التغذية الامامية ANN.

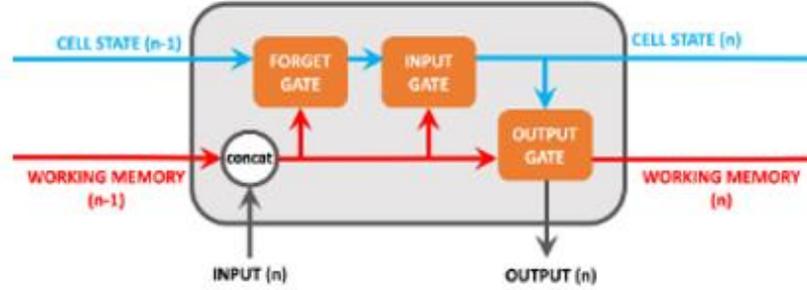
- تمثل العنصر الرئيسي الذي يسبب الاخطاء في شبكات ANN ، يسبب البث الخلفي عبر مجموعة من الطبقات المخفية مجموعة من الأخطاء التدرجية مما يسبب صعوبة في تقدير التدرج المتناقص وتحقيق القيم المثالية للأوزان.
- في شبكات RNN وعند القيام بعمليات البسط تتعرض BPTT عبر مجموعة من الخطوات الزمنية على نفس أثر الخطأ التدرجي مما يسبب نفس المشكلة السابقة من ناحية الأوزان والتدرج.

٤,٤ . حل مشكلة تلاشي التدرج مع شبكات LSTM:

يوجد العديد من الشبكات التكرارية والتي تعمل من اجل حل مشكلة التلاشي حيث تعتبر طريقة Long Short Term Memory Networks (LSTMs) الأكثر استخداماً.

كما في شبكات RNN التقليدية تقوم LSTM بالدوران عبر تسلسل من البيانات لتحافظ على استمرار وتجميع ذاكرة العمل عبر مجموعة من الدورات كما تقوم بمشاركة الأوزان وتوابع التفعيل عبر الدورات باستخدام الأوزان التي تم تحقيقها من خلال BPTT.

تقوم LSTM بإضافة المزيد من العناصر كما هو موضح بالشكل (٦) التالي:



الشكل (٦) بنية العصبون في شبكات RNN

تعمل ذاكرة العمل في هذا المخطط بنفس فعالية الاداء وبفلس الطريقة كما هو في شبكات RNN التقليدية، تتضمن المعلومات الواردة إلى العقدة:

- ١- حالة الخلية cell state: من الدورة السابقة.
- ٢- ذاكرة العمل working memory: من الدورة السابقة كما في شبكات RNN التقليدية.
- ٣- قيمة الدخل Input Value: من الدورة الحالية كما في شبكات RNN التقليدية.

في حين تتضمن العناصر الحديثة:

- حالة الخلية cell state: وتمثل حجر الاساس في LSTM حيث تمثل مهمتها بالعمل كذاكرة ذات مدى بعيد long-term memory وتضمن استمرار المعلومات في حال الضرورة عبر كل الدورات في العقدة. كما يمكن تعديل عملها لحذف البيانات غير الضرورية أو الحفاظ على البيانات السياقية الضرورية. كما تضمن عدم ضياع البيانات من الدورات المبكرة عبر التسلسلات الطويلة.
- بوابة النسيان Forget Gate: ومهمتها الوحيدة تحديد البيانات التي يجب ان يتم حذفها من حالة الخلية. ويتم ذلك من خلال إجراء الحسابات على ذواكر العمل المتصلة وقيم الدخل ومن ثم تطبيق ذلك على حالة الخلية.
- بوابة الدخل Input gate: وهي تحدد أي المعلومات التي يجب إضافتها إلى حالة الخلية اعتماداً على ذواكر العمل المتصلة وقيم الدخل.

## دراسة وتحليل خوارزميات التفرع في التعرف على الوجوه

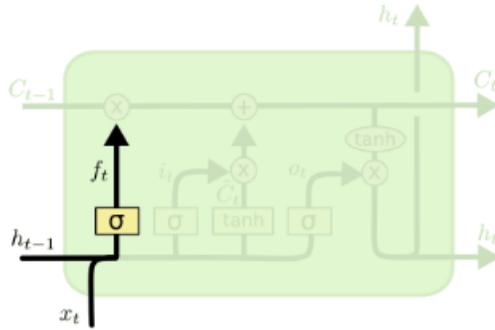
- بوابة الخرج output Gate: وهي تحدد أي ذاكرة عمل ستقوم العقدة بتمريرها للخروج اعتماداً على مجموعة من الحسابات على حالة الخلية الحالية وقيم ذاكرة العمل والدخل المتصلة.

يتضمن الخرج من هذه العقدة:

- 1- حالة الخلية cell state: وهي تمثل long-term memory للتكرار الحالي.
  - 2- ذاكرة العمل Working Memory: للتكرار الحالي.
  - 3- قيمة الخرج output value: والذي سيستخدم للتنبؤ بالتكرارات وهو يعمل بنفس طريقة ذاكرة العمل.
- ٤,٥ . كيفية عمل LSTM:

تمثل الخطوة الأولى في LSTM تحديد أي البيانات التي سنقوم بحذفها من حالة الخلية، يتم ذلك من خلال طبقة السيغمويد Sigmoid layer والتي تسمى بطبقة بوابة النسيان forget gate layer حيث تنظر إلى الخرج  $h_{t-1}$  والدخل  $x_t$  وتقوم بجبراج رقم قيمته بين الصفر والواحد لكل عدد ضمن حالة الخلية  $C_{t-1}$  تمثل القيمة 1 عبارة "احتفظ بهذه القيمة" في حين تمثل القيمة 0 "تخلي عن هذه القيمة".

باعتبار نموذجنا من اجل التنبؤ بالبيكسل التالي اعتماداً على البيكسل السابق ضمن الصورة الرقمية، قد تتضمن حالة الخلية على سبيل المثال القيمة اللونية للبيكسل، عند الوصول إلى لون جديد يجب أن نقوم بحذف اللون السابق للوصول إلى حالة الارتباط اللوني ضمن مجال محدد من البيكسلات ضمن الصورة.

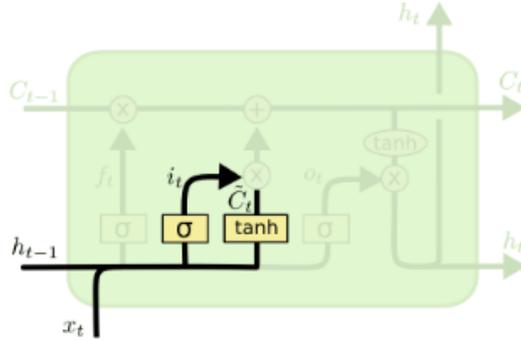


$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

الشكل (٧) الخطوة الأولى في الشبكات التكرارية، تحديد البيانات المحذوفة

تمثل المرحلة التالية نمط البيانات التي سنقرر تخزينها في حالة الخلية. وتتضمن هذه العملية خطوتين. الأولى، تحدد طبقة سيغمويد تسمى طبقة بوابة الدخل أي قيمة سوف نقوم بتحديثها وفي المرحلة الثانية تقوم طبقة tanh بتوليد شعاع من القيم المرشحة  $\tilde{C}_t$  والتي سيتم إضافتها إلى الحالة. في الخطوة التالية سوف يتم دمج كلا هاتين العمليتين لتوليد تحديث للحالة.

في حالتنا سنقوم بإضافة القيمة اللونية الجديدة لاستبدال القيمة اللونية القديمة.



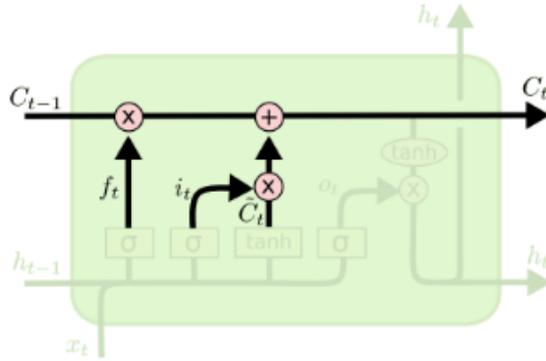
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

الشكل (٨) بناء مجموعة القيم المرشحة لعمليات الاستبدال

وهنا سيتم تحديث حالة الخلية السابقة  $C_{t-1}$  إلى الحالة الجديدة  $C_t$  وتتم هذه العملية عن طريق جداء الحالة القديمة بالتابع  $f_t$  حيث تم تحديد القيم المرغوب نسيانها مسبقاً ومن ثم يتم إضافة  $i_t * \tilde{C}_t$  وهي قيم المرشحين الجديدة مضبوطة بمقدار الرغبة حول قرار تحديث كل قيمة حالة.

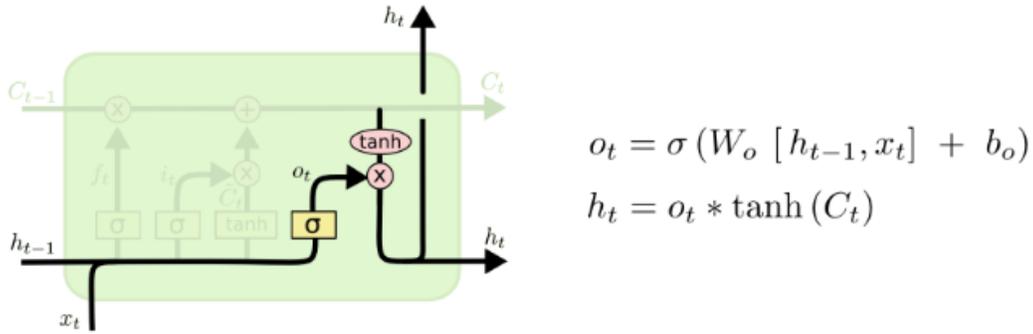
في حالتنا تم اتخاذ القرار بحذف المعلومات حول قيمة البيكسل السابق وإضافة البيانات الجديدة.



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

الشكل (٩) إجراء عملية الاستبدال

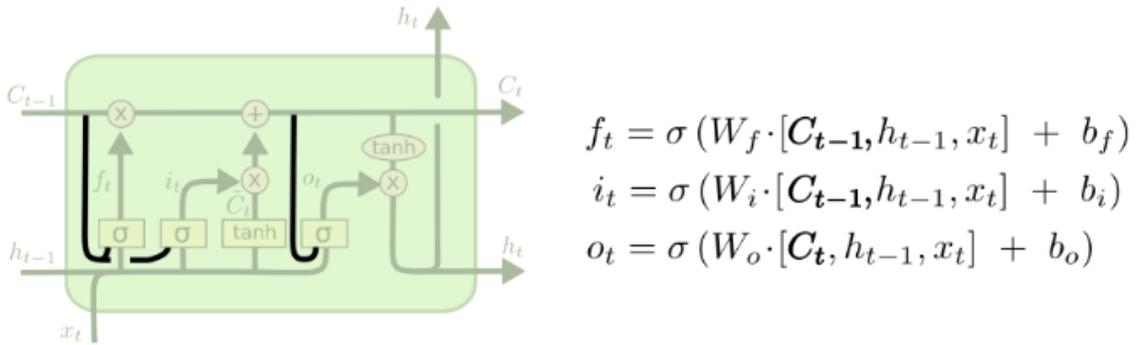
تمثل المرحلة التالية اتخاذ القرار بما سيتم تمريره للخروج وهو يعتمد على حالة الخلية، إلا أنه سيكون نسخة مرشحة، في البداية سنقوم بتشغيل طبقة السيغمويد والتي ستقرر أي الحالات من حالة الخلية سيتم اخراجها ومن ثم يتم تمرير الحالة إلى  $\tanh$  (لدفع القيم لتكون بين ١ و -١) وضربها مع خرج بوابة السيغمويد وبالتالي يتم فقط تمرير القيم التي نحتاجها. الهدف من هذه العملية في مثالنا التفريق بين القيم اللونية.



الشكل (١٠) تمرير القيمة للخروج

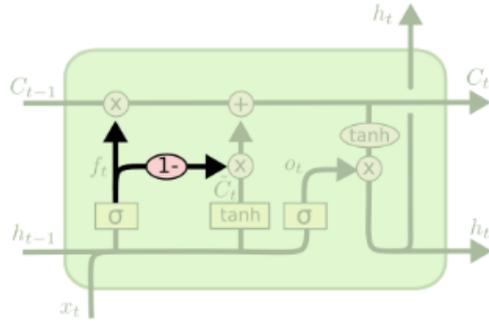
#### ٤,٦ المتحولات في LSTM:

ضمن نسخة LSTM التقليدية، والموضحة من قبل (2000) Gers & Schmidhuber تتمثل بإضافة peephole connections وهذا يعني أننا نسمح لطبقات البوابات بالنظر إلى حالة الخلية.



الشكل (١١) إضافة peepholes إلى كل البوابات

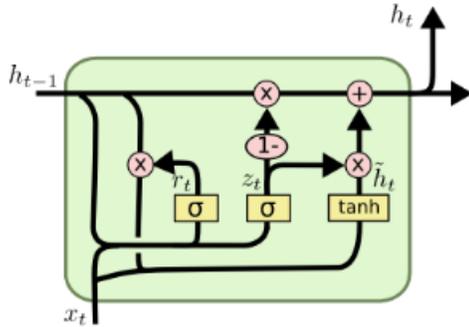
يوضح الشكل (١١) السابق إضافة peepholes إلى كل البوابات ، من النماذج الأخرى المعتمدة استخدام زوج من بوابات النسيان والدخل. وبدلاً من اتخاذ القرارات بصورة منفصلة حول ما يتم نسيانه وما يتم إضافته من المعلومات، سيتم اتخاذ هذه القرارات معاً. أي نقوم بنسيان البيانات فقط عندما نريد إدخال شيء ما بدلاً منها. نقوم بإدخال القيم الجديدة إلى الحالة عندما نقوم بنسيان شيء أقدم.



$$C_t = f_t * C_{t-1} + (1 - f_t) * \tilde{C}_t$$

الشكل (١٢) شبكة LSTM مع زوج من بوابات Forget

يوجد نسخة معدلة من LSTM وهي Gated Recurrent Unit والمقدمة من قبل Cho, et al. (2014) حيث تقوم بدمج بوابات الدخل والنسيان ببوابة واحدة تسمى بوابة التحديث update gate كما تقوم بدمج حالة الخلية مع الحالة الخفية وتجري عدد من التغييرات الأخرى. يعتبر النموذج الناتج أكثر بساطة من نموذج LSTM المعياري، مما رفع من حالة استخدام هذا النموذج.



$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

الشكل (13) شبكة LSTM مع وحدة Update Gate

من النماذج الأخرى المعتمدة نموذج Depth Gated RNNs by Yao, et al. (2015) ونموذج مختلف كلياً يسمى Clockwork RNNs by Koutnik, et al. (2014).

#### ٤,٧ . شبكات RNN الديناميكية:

يسمح استخدام خوارزمية Tensorflow في شبكات RNN بالقيام بالعمليات الحسابية ضمن أطوال متغيرة، في حين تؤمن شبكات RNN التقليدية إجراء التدريب على شبكة RNN بقيم أطوال ثابتة للبيانات، أي إذا كان لدينا سلسلة زمنية مؤلفة من ٢٠٠ خطوة وبالتالي فإننا نقوم ببناء مخطط شبكة ثابت ب ٢٠٠ خطوة RNN، يعتبر توليد النموذج الأول بطيئاً كما لا يمكننا إدخال قيم تسلسل أكبر من ٢٠٠ والذي قمنا بتحديد مسبقاً.

في الشبكات الديناميكية يتم حل هذه المشكلة باستخدام حلقة while والتي تقوم بضبط حجم المخطط عند التنفيذ، يكون بناء النموذج أسرع بالإضافة للقدرة على إضافة مجالات ذات أحجام متغيرة. تختلف أيضاً شبكات RNN الديناميكية بطريقة إعداد المداخل والمخارج والتي أيضاً تكون متغيرة وفقاً لحجم batch.

#### ٤,٨ . شبكات RNN ثنائية الاتجاه:

وتتألف من شبكتي RNN متصلتين معاً، يتم تغذية الدخل بصورة طبيعية ضمن التسلسل الزمني المحدد إلى الشبكة الأولى، في حين يتم تغذيته بصورة عكسية إلى الشبكة الثانية ويتم ربط مخارج الشبكتين معاً ضمن كل خطوة زمنية، يسمح هذا التركيب بامتلاك الشبكة معلومات أمامية وخلفية حول التسلسل في كل لحظة زمنية.

تعاني هذه الشبكات من التعقيد حيث يجب تحديد طريقة تمرير البيانات من المخارج إلى طبقة العصبونات dense neural network.

في الشبكة التقليدية تتم عملية التمرير هذه في اللحظة الزمنية الأخيرة، وهنا سوف ترى الشبكة العكسية فقط آخر قيمة زمنية مما يفقد الشبكة القدرة على التنبؤ.

#### ٥ . النتائج العملية:

من أجل تطوير نظام البرمجة التفرعية يجب تقسيم العمل إلى عدة مراحل تتمثل بما يلي:

- المرحلة الأولى: بناء قاعدة الصور: يتم إنشاء قاعدة الصور إما من صور مفردة أو من مجموعة من الصور الملتقطة من الفيديو، تم إعداد نظامين لالتقاط صور المستخدمين من ملفات الفيديو الأول يعتمد على البرمجة التسلسلية التقليدية والثاني اعتماداً على توزيع مهام التقاط اطارات الصور بصورة تفرعية بهدف تسريع أداء النظام لتجميع البيانات. يسمح النظام المقترح بالتقاط الفيديو من كاميرا واحدة وتوزيع مهمة معالجة الإطارات أو النقاط مجموعة من الصور من الكاميرات المختلفة ومعالجة كل كاميرا على حدى وفقاً لنواة مستقلة.
- بإجراء الاختبار على النقاط الفيديو المعد مسبقاً والمسجل من كاميرا ما تم الاختبار على ملف فيديو بطول دقيقتين من أجل استخلاص الاطارات بصورة تسلسلية أو صورة تفرعية.
- في حالة النقاط وتخزين الإطارات بصورة تسلسلية استغرقت العملية زمن 61.16 ثانية وفي حال معالجة الإطارات بصورة تفرعية تم الحصول على زمن ٥٩ ثانية. سمحت عملية تخزين الإطارات بصورة تفرعية بتوفير الزمن المطلوب لتمرير الإطارات إلى ملف التخزين.
- المرحلة الثانية: اختيار الشبكة العصبية للاختبار: تم إجراء الاختبارات على مجموعة من الشبكات العصبية وهي RNN بدون استخدام TensorFlow، ومع استخدام TensorFlow تم اختبار شبكات RNN, CNN, Dynamic-SoftMax، وتمت عمليات RNN, Bidirectional-RNN وتقييم معايير السرعة والدقة. تم اعتماد تابع التفعيل من نمط SoftMax وتمت عمليات

الاختبار على معالج core i5 ثنائي النواة رباعي الخيوط (2cores-4threads) بذاكرة 6GB ومن دون الاعتماد على وحدة معالجة بطاقة الإظهار GPU. تم اعتماد عدد من الدورات الموحد لجميع الخوارزميات بمقدار ١٠٠٠ دورة وتم حساب الزمن بدءاً من عملية تحميل البيانات إلى الشبكة وحتى الوصول للخروج. يوضح الجدول (١) التالي النتائج التي تم الوصول إليها.

	Accuracy	time
RNN – no tensorflow	0.28	0:18:2
RNN-with tensorflow	0.45	5:04:00
CNN-with tensorflow	0.98	10:41:10
RNN Dynamic-with tensorflow	0.59	2:49:20
RNN-bidirectional-with tensorflow	0.49	4:47:30

من النتائج الموضحة فقد تبين أن شبكات CNN تعطي الدقة الأعلى والأفضل ولكن خلال زمن أطول، في حين تتمتع الشبكات الديناميكية RNN بأفضل ثنائية سرعة - دقة حيث سيتم رفع الدقة بشكل أكبر مع زيادة عدد الدورات. توضح النتائج أيضاً أهمية استخدام Tensorflow للحصول على الدقة الأفضل حيث تعاني شبكات RNN بدون استخدامه من انخفاض كبير للدقة على الرغم من السرعة العالية لانتهاء عملية التصنيف. المرحلة الثالثة: تم إجراء عملية التدريب للشبكة العصبية للتعرف على الوجوه بصورة تسلسلية وصورة تفرعية، بعد إجراء الاختبار والتعرف اعتماداً على أربعة خيوط كان لدينا النتائج التالية:

النظام التسلسلي التقليدي	النظام التفرعي	استخدام المعالج
13%	22%	Thread1
94%	23%	Thread2
12%	92%	Thread3
8%	99%	Thread4
31%	63%	الاستخدام الكلي
1.4sec	0.3sec	زمن التعرف

## ٦. النتائج والتوصيات:

تقدم هذه الورقة خوارزمية من أجل اختيار الشبكة العصبية المناسبة الأفضل في عمليات تصنيف الصور الرقمية من ناحية الدقة- السرعة وذلك بهدف استخدامها في النظم الموزعة.

بعد حساب الزمن الكلي لأداء المراحل المختلفة للتعرف على الوجوه تم الوصول إلى أن سرعة النظام المتفرع المعتمد على شبكات CNN أفضل بالأداء بنحو ٣٠ بالمئة من النظام المستخدم بصورة تسلسلية بالإضافة إلى قدرة النظام على القيام بعمليات القراءة المتعددة بنفس الوقت بدلاً من القيام بمهمة وحيدة.

تتمتع هذه الشبكات أيضاً بإمكانية استخدامها في النظم الموزعة حيث يتم توزيع المهام على الذواكر أو المعالجات مما يسمح برفع فعالية الاداء بشكل أكبر واستخدام المصادر الحاسوبية المتاحة بشكل أكبر فعالية في المجالات العلمية و التطبيقية.

## المراجع:

- [1]Quinn, M.: Parallel Programming in C with MPI and OpenMP. McGraw Hill, Boston (2003)Google Scholar
- [2]Halfhill, T.R.: Parallel Processing With CUDA. Microprocessor report (2008), <http://www.MPRonline.com>
- [3]Nvidia: Nvidia CUDA Compute Unified Device Architecture, Programming Guide, Version 1.1 (2007), <http://nvidia.com/cuda>
- [4]Seiffert, U.: Artificial neural networks on massively parallel computer hardware. In: ESANN 2002 proceedings, Bruges, Belgium, pp. 319–330 (2002)Google Scholar
- [5]Lotrič, U., Dobnikar, A.: Parallel implementations of feed–forward neural network using MPI and C# on.NET platform. In: Ribeiro, B., et al. (eds.) Adaptive and natural computing algorithms: proceedings of the International Conference in Coimbra, Portugal, pp. 534–537 (2005)Google Scholar
- [6]Catanzaro, B., Sundaram, N., Keutzer, K.: Fast support vector machine training and classification on graphics processors. In: McCallum, A., Roweis, S. (eds.) Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland, pp. 104–111 (2008)Google Scholar
- [7] CNN–RNN: A Unified Framework for Multi–label Image Classification, Jiang Wang, Yi Yang, Junhua Mao, Zhiheng Huang, Chang Huang, Wei Xu, University of California, 2018.
- [8] Jean Mas and Juan Flores, “The application of artificial neural networks to the analysis of remotely sensed data,” International Journal of Remote Sensing, vol. 29, no. 3, pp. 617–663, 2008.
- [9] Mathieu Fauvel, Yuliya Tarabalka, Jon Atli Benediktsson, Jocelyn Chanussot, and James C Tilton, “Advances in spectral–spatial classification of hyperspectral images,” Proceedings of the IEEE, vol. 101, no. 3, pp. 652–675, 2013.
- [10] Christopher Lloyd, Suha Berberoglu, Paul Jand Curran, and Peter Atkinson, “A comparison of texture measures for the per–field classification of mediterranean land cover,” International Journal of Remote Sensing, vol. 25, no. 19, pp. 3943–3965, 2004.

- [11] Liang-Chieh Chen, Jonathan T Barron, George Papandreou, Kevin Murphy, and Alan Yuille, "Semantic image segmentation with task-specific edge detection using CNNs and a discriminatively trained domain transform," arXiv preprint arXiv:1511.03328, 2015.
- [12] Eduardo S. L. Gastal and Manuel M. Oliveira, "Domain transform for edge-aware image and video p
- [13] RGB-D-based Human Motion Recognition with Deep Learning: A Survey, Pichao Wang, Wanqing Lia , Philip Ogunbonaa , Jun Wanc , Sergio Escalera, arXiv:1711.08362v2 [cs.CV] 24 Apr 2018.
- [14] Qiwei Yin, Ruixun Zhang, and XiuLi Shao, CNN and RNN mixed model for image Classification, MATEC Web of Conferences 277, 02001 (2019).
- [15] Agarap, F. A. (2017). An Architecture Combining Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for Image Classification. arXiv preprint arXiv:1712.03541.2. Bhatnagar, S., Ghosal, D., Kolekar, M. (2017).
- [16] Classification of fashion article images using convolutional neural networks. 2017 Fourth International Conference on Image Information Processing (ICIIP), 1-6.
- [17] H. Lyu, H. Lu, L. Mou, W. Li, J. Wright, X. Li, X. L. Li, X. X. Zhu, J. Wang, L. Yu and P. Gong, "Long-Term Annual Mapping of Four Cities on Different Continents by Applying a Deep Information Learning Method to Landsat Data," Remote Sensing, vol. 10, no. 3, p. 471, 2018.
- [18] L. Mou, L. Bruzzone and X. X. Zhu, "Learning Spectral-Spatial-Temporal Features via a Recurrent Convolutional Neural Network for Change Detection in Multispectral Imagery," 2018.