

تطبيق التعلم بالنقل في تصنيف الصور

طالبة الدكتوراه: نهله سعد الدين

كلية الهندسة المعلوماتية - جامعة البعث

إشراف: د. يسر السيد سليمان الأتاسي

الملخص:

يعتبر التصوير الشعاعي تقنية مقبولة للمجتمع الطبي للكشف عن التشوهات. ومع ذلك، فإن تفسير الصور يستغرق وقتاً طويلاً وهو معرض للخطأ من قبل أخصائيي الأشعة الذين يتعرضون للعوامل الخارجية بما في ذلك التعب المحتمل الناتج عن العمل لساعات طويلة أو الإرهاق أو التفكير بأمر الحياة الأخرى.

ولتحسين كفاءة عمل أخصائيي الأشعة، قمنا بالعمل على تطوير نموذج تشخيص بمساعدة الكمبيوتر لتصنيف الصور الشعاعية إلى تصنيفين: حالة طبيعية وحالة شذوذ (أو حالة غير طبيعية) وذلك لتسهيل عملية التشخيص الشعاعي، من خلال نقل مجموعة من الشبكات العصبية التلافيفية العميقة المختارة بين مجموعة شبكات متوفرة قمنا بدراستها وتطبيقها على أساس المناطق المحتمل أن تكون غير طبيعية التي يوفرها اختصاصيو الأشعة لحالة الدراسة التي قمنا باختيارها.

نلقي في هذه الدراسة الضوء على دراسة تجريبية لإحدى الشبكات العصبية التلافيفية المتصلة كلياً بتطبيقها على عينة البيانات المدروسة من خلال تقنية التعلم بالنقل. حصلنا على مجموعة من النتائج الجيدة، والتي حققت دقة تشخيص عالية بلغت حوالي 88% في بعض الحالات المدروسة.

الكلمات المفتاحية: التعلم بالنقل - صنع قرار - تعلم آلة - تصنيف صور - MURA -
- صور شعاعية - DenseNet - شبكة عصبية تلافيفية متصلة.

Applying Transfer Learning in image classification

Abstract:

Radiography is an accepted technique for the medical community to detect abnormalities. However, the interpretation of the images is time-consuming and is subject to error by radiologists who are exposed to external factors including possible fatigue from working long hours, overwork, or thinking about other life matters.

To improve the efficiency of the radiologists' work, we developed a computer-assisted diagnostic model to classify the radiographs into two classifications: a normal condition and an anomaly (or an abnormal state) in order to facilitate the radiological diagnostic process, by transferring a set of selected deep convolutional neural networks between a set of available networks. We have studied and applied them on the basis of potentially abnormal areas that radiologists provide for the study case we have selected.

In this study, we shed light on an experimental study of a fully connected convolutional neural network (DenseNet) by applying it to the studied data sample through a learning transfer technique.

We obtained a set of good results, which achieved high diagnostic accuracy of about 88% in some of the studied cases.

Keywords: Transfer Learning – Decision Making– Machine Learning- Image Classification – MURA – Radiologist Images – DenseNet – Fully Connected Convolutional Neural Network

1- مقدمة:

تعد صور الصدر الشعاعية من أكثر الصور التي يتم الحصول عليها بشكل متكرر في مجال الأشعة وغالباً ما تكون موضوعاً لأبحاث الرؤية الحاسوبية. معظم النماذج المستخدمة لتصنيف الصور الشعاعية مستمدة من الشبكات العصبية العميقة المتاحة والمدرية على مجموعات بيانات الصور الكبيرة وذلك لمساعدة أطباء الأشعة على تحسين كفاءتهم في العمل بالإضافة إلى دقة التشخيص، تم إنجاز أعمال كبيرة لتصميم أنظمة التشخيص بمساعدة الحاسب (Computer Aided CAD (Diagnosing) مساعدة في هذا المجال [1]. وهناك فئتان رئيسيتان لأنظمة CAD حسب الوظيفة: أولاً الكشف وثانياً التشخيص.

عادةً ما يكون تحديد المناطق غير الطبيعية في الصور (أو ما يسمى عادة حالات الشذوذ) هو الخطوة الأساسية والأهم في جميع أنظمة CAD تقريباً. كما يمكن أن تكون هذه الأنظمة مفيدة للغاية عندما يصعب على أخصائي الأشعة اكتشاف أو تحديد المناطق غير الطبيعية نظراً للمحتويات المعقدة في الصور. بالإضافة إلى ذلك، يتم تطبيق أنظمة التشخيص بشكل عام لمساعدة أخصائي الأشعة على اتخاذ قرارات بشأن تشخيص الصور عند الضرورة [2].

2- هدف البحث:

تم العمل في هذا البحث على تطبيق دراسة عملية تجريبية لشبكة DenseNet201 العميقة والكبيرة من ناحية عدد الطبقات. وبالاستفادة من خلال تقنية التعلم بالنقل / Transfer Learning قمنا بتطبيق دراستنا على عينة بيانات رسمية تم نشرها من قبل فريق تعلم الآلة في جامعة ستانفورد (تفاصيل أكثر عن عينة البيانات نجدها لاحقاً ضمن البحث)، ودراسة مدى تأثير تقنية التعلم بالنقل على النتائج.

كما سنتناول الدراسة تفاصيل بنية الشبكة المدروسة، وآلية تطبيق التعلم بالنقل بالاستفادة من التقنيات المتوفرة للأبحاث، إضافةً إلى النتائج والملاحظات التي حصلنا عليها بعد تنفيذ هذه الدراسة.

3- الدراسات السابقة:

صمم نيراج وآخرون نظاماً للكشف مع سلسلة من التعلم العميق ومصنفات الغابات العشوائية المقدمة [3]. كما تم استخدام شبكة المعتقدات العميقة كمصنف لتحديد المناطق الخادعة متعددة المقاييس في المرحلة الأولى يليها مصنف الغابات العشوائية لتحديد المناطق التي نتجت من المرحلة الأولى لتكون جماعية أم لا. أظهرت التجارب على مجموعتين مختلفتين من البيانات نتائج واعدة. تم اقتراح عمل مشابه لاكتشاف التكلس باستخدام المصنفات المتتالية.

هناك أيضاً عدد كبير من الأعمال على أنظمة التشخيص المتعلقة بالتكلس المجهري والكتلة [5] [4] [6].

بحيث تم في [4] دراسة للعلاقات بين علم الأمراض والبنية البلورية في تكلسات الثدي وتصنيفها إلى مجموعتين "حميدة" و "خبيثة". في الدراسة [5] تم قياس التركيب الأولي للتكلسات الموجودة في الأقسام النسيجية لخزعات أنسجة الثدي، وربط ذلك بكل من البارامترات البلورية المقاسة سابقاً في نفس العينات، وتقرير التشريح المرضي والاستفادة منها في عملية التشخيص. في الدراسة [6] تم تقديم فحص الخلايا بالإبرة الدقيقة للثدي للفئات غير الخبيثة بحيث يعتمد التشخيص الدقيق الحالي لآفات الثدي على نهج تقييم ثلاثي يشمل الفحوصات السريرية والتصويرية والمرضية. في هذه الدراسة تم اعتماد فحص الخلايا بالإبرة الدقيقة (FNAC) على نطاق واسع للتقييم المرضي بسبب دقته وسهولة استخدامه.

تفوقت شبكة CNN المدربة المكونة من تسع طبقات في نظام الكشف عن سرطان الثدي، والتي صممها بان، على ستة أساليب أخرى من أحدث الأساليب [7]. وصلت دقة التشخيص إلى 94% لكنها لا تزال قابلة للتحسين. وفي نظام آخر، تم تقديم شبكة عصبية تلافيفية جديدة (CNN) تسمى GlimpseNet [27]. يتم استخراج مناطق الاهتمام المتعددة (ROI) في وقت واحد ثم يتم تصنيفها. ثم يتم إعطاء نتيجة تشخيصية للصورة الكاملة عن طريق تجميعها لإعطاء نتيجة مشوهة للصورة الكاملة. ومع ذلك، كما ذكر المؤلفون، لم يتم إصلاح مشكلة زيادة التخصيص بينما اكتسب أداء الدقة 4.1% مقارنة بالخوارزميات الموجودة.

الشبكات العصبية التلافيفية:

على الرغم من أن شبكات CNN العميقة قد استخدمت على نطاق واسع في المجالات التي لا تشمل فقط التحليل والتشخيص الطبي [9]، فإنه لا تزال هناك بعض المشكلات التي يتعين حلها في الاستخدام العملي لكل مشكلة. ومن قيود المشاكل العملية نذكر مثلاً الحجم المحدود لبيانات التدريب. أيضاً، كما أشار جيسون [10]، تتعلم الطبقات السفلية في شبكات CNN العميقة وخاصة الطبقة الأولى بعض الميزات العامة المشابهة لمرشحات Gabor أو نقاط الألوان، لكن الطبقات العليا تتعلم المزيد من الميزات المحددة فيما يتعلق بمجموعة البيانات. التعلم الانتقالي، الذي تم تقديمه لتبني المصنفات المدربة لفئات أخرى لتصنيف فئات معينة، تبين أنه طريقة فعالة لحل مثل هذه الحالات [12] [13]. ولهذا قمنا بالعمل على تطبيقه ودراسة دوره في عينة البيانات المستخدمة في بحثنا هذا. هناك طريقتان لتحويل الشبكات القائمة إلى شبكات موجهة للمشاكل. الإجراء الأول لطريقتي للتعلم بالنقل هو نفسه وهو نسخ أول n من الطبقات السفلية من الشبكات الأساسية إلى الشبكات المستهدفة بينما تتم تهيئة البارامترات (Parameters) في الطبقات المتبقية من الشبكات المستهدفة بشكل عشوائي. الفرق بين الطريقتين هو في

الإجراء الثاني وهو أن إحداهما تترك البارامترات في الطبقات المنسوخة مجمدة بينما تقوم الأخرى بضبط البارامترات عند تدريب الشبكات المستهدفة. يُقترح تقنية التجميد عندما يكون حجم مجموعة البيانات المستهدفة صغيراً بينما تكون أعداد البارامترات كبيرة وإلا يجب اعتماد الضبط الدقيق عندما تكون مجموعة البيانات الكبيرة مصحوبة بعدد صغير من البارامترات [14].

في التعلم بالنقل، يتم تطبيق معرفة نموذج التعلم الآلي الذي تم تدريبه بالفعل على مشكلة مختلفة ولكنها ذات صلة. على سبيل المثال، إذا قمنا بتدريب المصنف البسيط للتنبؤ فيما إذا كانت الصورة تحتوي على حقيبة ظهر، فيمكننا استخدام المعرفة التي اكتسبها النموذج أثناء تدريبه للتعرف على أشياء أخرى مثل النظارات، أقلام،

أي أننا نحاول بشكل أساسي استغلال ما تم تعلمه في مهمة A لتحسين التعميم في مهمة أخرى B بحيث نقوم بنقل الأوزان التي تعلمتها الشبكة في "المهمة A" إلى "المهمة B" الجديدة.

أصبح هذا النوع من النموذج ذو المرحلتين شائعاً للغاية في العديد من التطبيقات العملية، وخاصة في التصوير الطبي. بحيث تم ضبط البنى القياسية المصممة لـ ImageNet مع الأوزان المقابلة مسبقاً للتدريب بدقة على المهام الطبية التي تتراوح من تفسير الأشعة السينية للصدر وتحديد أمراض العيون ، إلى الاكتشاف المبكر لمرض الزهايمر.

على هذا النحو، هناك العديد من المهام والمسائل المميزة لاعتماد التعلم بالنقل في مجالات مختلفة بما في ذلك الكشف عن المرض وتشخيصه. قام الباحثون في [15] بنقل شبكة CNN العميقة المكونة من تسع طبقات والتي تم تدريبها مسبقاً على مجموعة كبيرة من بيانات تصوير الثدي بالأشعة السينية ولكنها حسنت دقة الكشف عن الكتلة في التحليل المقطعي للثدي الرقمي (DBT) بشكل كبير من 0.80 إلى 0.91.

وفي محاولة أخرى، قام بنيامين وآخرون بتحسين دقة تصنيف الورم الشعاعي للثدي من 0.81 إلى 0.86 بعد تطبيق التعلم بالنقل [16].

طبقتنا في دراستنا أحدث شبكات DenseNet وأضفناها إلى بحثنا كمصنفات ذات نتائج مرضية بشكل جيد، والتي تزيل الإجراءات الأولية مثل التجزئة في أنظمة التشخيص التقليدية. لتحديد ما إذا كان يجب تجميد الشبكات الأساسية أم لا في هذه المشكلة، استخدمنا طريقة للتعلم بالنقل، كما تحققنا من صحة دراستنا من خلال مقارنة دقة النتائج التي حصلنا عليها بأحدث الأعمال وتوصلنا إلى استنتاج مفاده أن طريقتنا تعمل بشكل أفضل بينها جميعاً على عينة البيانات المدروسة.

التعلم بالنقل (Transfer Learning):

قد تستغرق نماذج الشبكة العصبية التلافيفية العميقة أياماً أو حتى أسابيع للتدريب على مجموعات بيانات كبيرة جداً. وتتمثل إحدى طرق اختصار هذه العملية في إعادة استخدام أوزان النموذج من النماذج المدربة مسبقاً التي تم تطويرها لمجموعات البيانات المعيارية، مثل مهام التعرف على الصور ImageNet. يمكن تنزيل النماذج الأفضل أداءً واستخدامها مباشرة، أو دمجها في نموذج جديد لحل مشكلات الرؤية الحاسوبية.

يشير مصطلح التعلم بالنقل إلى تأثير المعرفة المكتسبة من قبل الشبكة العصبية المدربة مسبقاً على مجموعة بيانات معينة (كبيرة عادة) متاحة لحل المشكلات الجديدة التي تتوفر لها أمثلة تدريب قليلة، ودمج المعرفة الحالية مع المعرفة الجديدة المستفادة من الأمثلة القليلة من مجموعة البيانات الخاصة بالمسألة. وبالتالي، يتم استخدام التعلم بالنقل بشكل كبير، غالباً مع تقنيات أخرى مثل زيادة عينة البيانات المدروسة، من أجل معالجة مشكلة نقص بيانات التدريب.

يعمل التعلم الانتقالي فقط في التعلم العميق إذا كانت ميزات النموذج التي تم تعلمها من المهمة الأولى عامة.

في التعلم بالنقل، نقوم أولاً بتدريب شبكة أساسية على مجموعة بيانات أساسية ومهمة، ثم نقوم بإعادة توظيف الميزات المكتسبة، أو نقلها، إلى الشبكة الهدف ليتم

تدريبها على مجموعة بيانات ومهمة هدف. ستتم هذه العملية إذا كانت الميزات عامة، ما يعني أنها مناسبة لكل من المهام الأساسية والهدف، بدلاً من المهمة المحددة للمهمة الأساسية.

طريقة النموذج المدرب مسبقاً:

1- نحدد النموذج المصدر: يتم اختيار نموذج مصدر مدرب مسبقاً من النماذج المتاحة بحيث تطلق العديد من المؤسسات البحثية نماذج على مجموعات بيانات كبيرة وصعبة يمكن تضمينها في مجموعة النماذج المرشحة للاختيار من بينها.

2- إعادة استخدام النموذج: يمكن استخدام النموذج الذي تم تدريبه مسبقاً كنقطة انطلاق لنموذج في مهمة ثانية. قد يشمل ذلك استخدام النموذج بالكامل أو أجزاء منه، اعتماداً على تقنية النمذجة المستخدمة.

3- ضبط النموذج: قد يحتاج النموذج إلى تكيفه أو تعديله على بيانات زوج الإدخال والإخراج المتاحة للمهمة الهدف.

ولكن يمكن أن يتساءل الباحثون هنا ما مقدار المساعدة التي يمكن أن يساعدها التعلم بالنقل فعلياً، قمنا بالإجابة على هذا السؤال فيما يتعلق بتصنيف مجموعة البيانات الخاصة بالصور الشعاعية الطبية التي ذكرناها في بند عينة البيانات المدروسة.

يعتبر التعلم بالنقل تحسين، اختصار لتوفير الوقت أو الحصول على أداء أفضل. وبشكل عام، ليس من الواضح أنه ستكون هناك فائدة من استخدام التعلم بالنقل في المجال إلا بعد تطوير النموذج وتقييمه.

طرق التعلم بالنقل:

تبين أن التعلم عن طريق النقل هو وسيلة فعالة لتقديم الشبكات ذات الأداء الجيد في التصنيف لحل مشاكل التصنيف العملية بشرط محدودية حجم البيانات [17] ومع ذلك، فإن شبكات CNN العميقة الحديثة مصممة بشكل عام للتعرف على الأشياء على نطاق واسع [18].

لذلك، في الخطوة الأولى من التعلم بالنقل، يجب استبدال الطبقة المتصلة بالكامل التي تحتوي على n خلية عصبية بطبقة تحوي عدد محدد من الخلايا العصبية لتلبية احتياجات التصنيف المحددة (الخاصة بدراستنا هنا).
يعتمد ما إذا كان ينبغي تطبيق تقنية التجميد أو الضبط الدقيق على عدد البارامترات التي سيتم تدريبها وعلى حجم البيانات المتاحة للدراسة.
للتمييز بين الأداء المختلف لهاتين التقنيتين، استخدمنا كلتا الطريقتين عند نقل الشبكات الأساسية بسبب غموض حجم البيانات وكذلك عدد البارامترات التي سيتم تدريبها. بالنظر إلى التصنيف الثنائي في حالتنا، فقد استبدلنا جميع الطبقات المتصلة بالكامل في الشبكات الأساسية بخلية عصبونية واحدة. لذلك، تم إعادة تدريب ثلاث طبقات عليا فقط، بما في ذلك الطبقة المتصلة بالكامل، وطبقة softmax، وطبقة التصنيف عندما قمنا بتجميد الطبقات التي تسبق الطبقة المتصلة بالكامل أثناء إعادة تدريب الشبكات بأكملها من خلال عدد صغير من الفترات لضبط البارامترات في حالة الدقة.

4- دراسة الحالة التجريبية:

لعبت مجموعات البيانات الكبيرة والعالية الجودة دوراً مهماً في دفع عجلة التقدم في مجالات التعليم العميق.
تتكون المهمة المختارة لتجربة التعلم بالنقل من تصنيف صور شعاعية طبية لسبعة أنواع مفاصل وعظام في الجسم إلى فئتين. يرجع اختيار هذه المهمة بشكل أساسي إلى توفر هذه المجموعة من البيانات، بالإضافة إلى مجال المشكلة، وهو عام بما يكفي ليكون مناسباً للتطبيق الفعال للتعلم بالنقل مع الشبكات العصبية المدربة مسبقاً على مجموعة بيانات ImageNet الشهيرة.

حسب الاحصائيات الأخيرة المتوفرة في مجال الحالات المتعلقة بمجال بحثنا (صور شعاعية) فإن أمراض العضلات والعظام تؤثر على أكثر من 1.7 مليار شخص في جميع أنحاء العالم، وهي السبب الأكثر شيوعاً للألم والإعاقة الشديدة وطويلة الأجل،

بحيث تزداد 30 مليون زيارة لقسم الطوارئ سنوياً وتزايد بسبب هذا النوع من الحالات [19][20].

مهمة الكشف عن المرض، أو بمعنى آخر تحديد ما إذا كانت دراسة التصوير الشعاعي طبيعية أو غير طبيعية، هي عملية إشعاعية مهمة بحيث تعتبر دراسة تفسر على أنها تستبعد المرض بشكل طبيعي ويمكن أن تلغي حاجة المرضى للخضوع لمزيد من الإجراءات التشخيصية أو التدخلات.

5- عينة البيانات المدروسة:

نعمل من خلال بحثنا على MURA Dataset، وهي مجموعة بيانات كبيرة للكشف عن التشوهات في التصوير الشعاعي العضلي الهيكلي تتألف من 14863 دراسة (وتحتوي كل دراسة على واحد أو أكثر من المشاهدات (الصور)) من 12173 مريضاً، بإجمالي 40,561 صورة إشعاعية متعددة الرؤية.

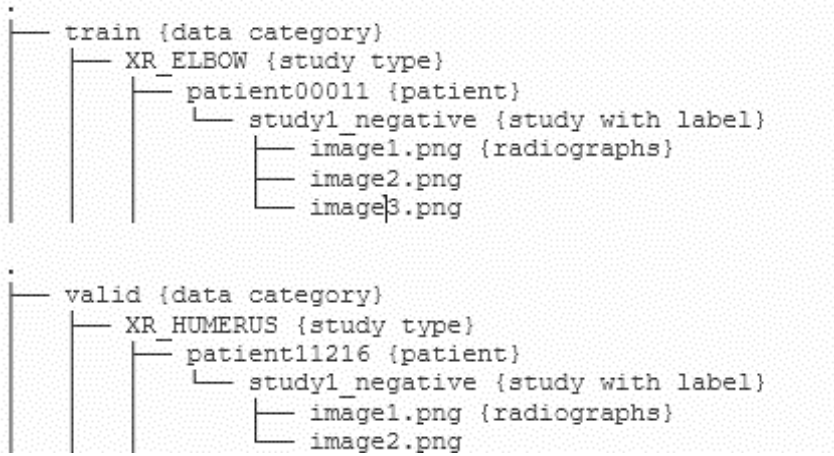
تحتوي مجموعة البيانات على 9,045 دراسة طبيعية و5818 دراسات إشعاعية عضلية هيكلية غير طبيعية [21].

ينتمي كل منها إلى واحد من سبعة أنواع من الدراسات التصويرية للأطراف العلوية: الكوع، الإصبع، الساعد، العضد، الكتف والمعصم واليد. تم تصنيف كل دراسة يدوياً على أنها طبيعية أو غير طبيعية من قبل أطباء الأشعة المعتمدين من مستشفى ستانفورد في وقت التفسير الشعاعي السريري في بيئة الأشعة التشخيصية بين 2001 و2012.

في الدراسة التي قدمتها جامعة ستانفورد تم تقسيم مجموعة البيانات إلى ثلاث مجموعات: مجموعة التدريب (11184 مريض، 13457 دراسة، 36808 صورة)، مجموعة التحقق (783 مريض، 1199 دراسة، 3197 صورة)، ومجموعة الاختبار (206 مريض، 207 دراسة، 556 صورة). كما أنه لا يوجد تداخل في المرضى بين أي من المجموعات.

أما في أغلب الدراسات التي تمت على نفس عينة البيانات فإنها تنقسم الصور إلى 36808 صورة تدريبية و3197 صورة للتحقق منها. بحيث تأتي عينة البيانات

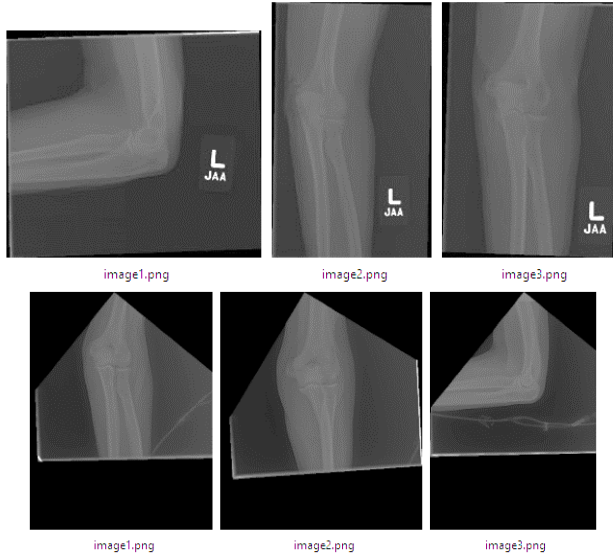
كمجموعة بيانات منقسمة إلى مجموعتين: مجموعة التدريب ومجموعة التحقق في مجلدات منفصلة. كل مجموعة من مجموعات البيانات (التدريب والتحقق) لها مجلدات فرعية (كما في الشكل (1))



الشكل (1): طريقة تقسيم عينة البيانات المدروسة

كما نرى من الشكل أنه يتم تقسيم البيانات ضمن المجموعتين المذكورتين سابقاً (في مجلد منفصل لكل مجموعة منها)، كما أنه يوجد مجلد خاص بكل فئة من الفئات السبعة والمذكورة أيضاً، ضمن كل فئة يوجد مجلد خاص لكل مريض (حالة دراسة مستقلة) ويتم تحديد حالة الدراسة مع وضع الـ label أو التصنيف لها، بحيث أن التصنيفين المحددين هما positive و negative، وضمن المجلد الخاص بحالة الدراسة هناك مجموعة من المشاهدات (الصور).

كما نقدم بعض الأمثلة في الشكل الآتي.



إحدى الدراسات لحالة دراسة مع صور متنوعة (Abnormal - الصورة من جهة المين وحالة Normal - الصورة من جهة اليسار) لعظم الكوع

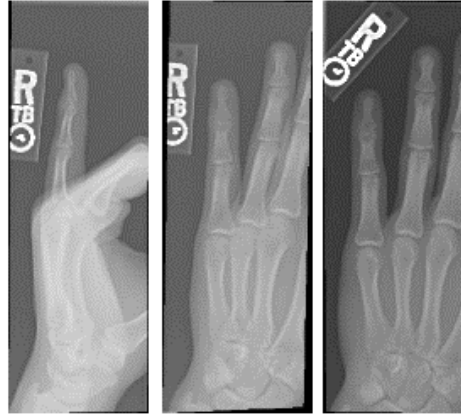
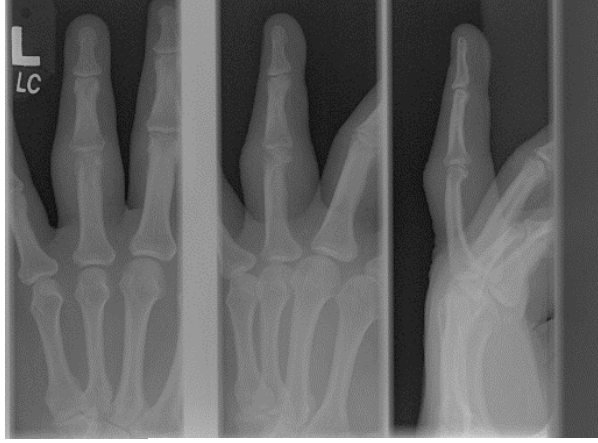


image1.png image2.png image3.png

إحدى الدراسات لحالة دراسة مع صور متنوعة (Abnormal - الصورة من جهة المين وحالة Normal - الصورة من جهة اليسار) لعظم الأصابع

الشكل (2): مجموعة من حالات الدراسة في عينة البيانات المستخدمة في البحث

تقسيم البيانات وزيادتها:

كما ذكرنا سابقاً عن البيانات المدروسة وتصنيفها إلى فئتين وفقاً للتسميات المقدمة من اختصاصي الأشعة. قمنا بتقسيم 80% بشكل عشوائي في مجموعة التدريب (Train Set) من مجموعة البيانات بالكامل، بينما بقي الجزء الباقي 20% في مجموعة الاختبار (Test Set).

تم تقييم أداء المصنف (النموذج) وفقاً لمتوسط دقة التصنيف الناتج في فترة تشغيل تبلغ 100 تكرار.

ملاحظة: السبب الذي جعلنا نحتفظ بـ 80% من مجموعة البيانات كمجموعة تدريب هو أننا اتبعنا قاعدة الاستبعاد، التي قسمت 80% من مجموعة البيانات الكاملة كمجموعة تدريب بينما تم استخدام البيانات المتبقية كمجموعة اختبار.

قاعدة الاستبعاد:

وفقاً لـ Andrew Ng، في Coursera MOOC on Introduction to Machine Learning، تتمثل القاعدة في تقسيم مجموعة البيانات إلى نسبة 3:1:1 (60:20:20) للتدريب والتحقق من الصحة والاختبار على التوالي. فعندما يحتوي المصنف على بارامترات يجب تعيينها، يمكن تقسيم مجموعة البيانات إلى 3 أجزاء (التدريب والتحقق والاختبار) بحيث يمكن ضبط البارامترات مع تجنب Overfitting.

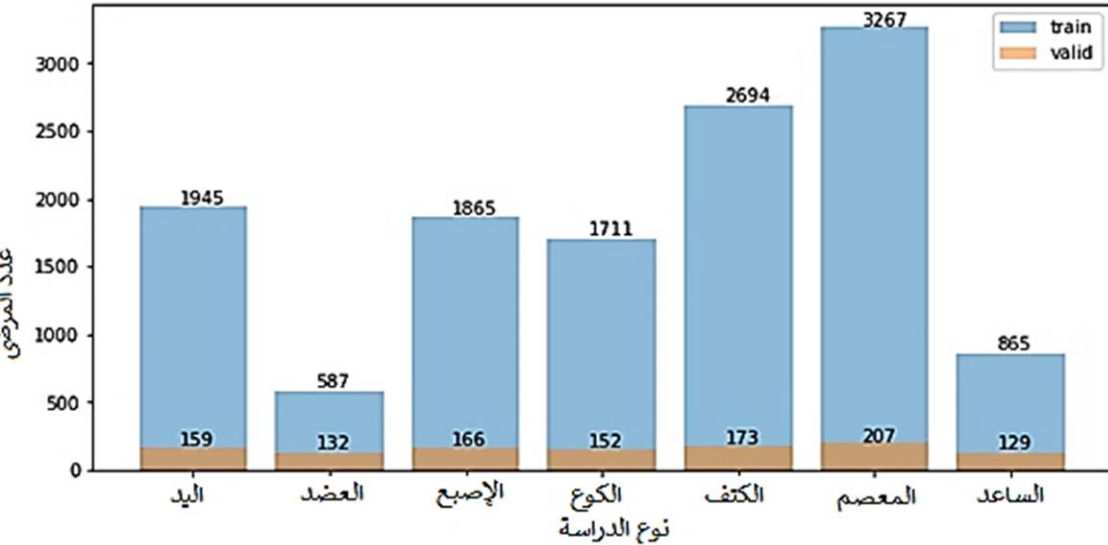
عند تدريب نظام التعلم على عينات البيانات، قد لا نعرف إلى أي مدى يمكنه التنبؤ بالعينات غير المرئية بشكل صحيح، لذلك يتم إجراء مفهوم التحقق (Cross Validation) لتعديل البارامترات المستخدمة للتدريب من أجل تحسين دقتها وإلغاء تأثير التناسب الزائد Overfitting على تلك البيانات. ولا ينبغي أن يتم ذلك في مجموعة الاختبار نفسها، وبالتالي علينا الفصل بين مجموعة الاختبار ومجموعة التحقق من الصحة.

في الحالات التي لا ينطبق فيها التحقق (Cross Validation)، من الشائع فصل البيانات بنسبة 8:2 (80:20) للتدريب والاختبار على التوالي. لتقدير القدرة التنبؤية الفعلية للمصنف (على سبيل المثال، الدقة/accuracy)، عادةً ما نأخذ 20-fold cross-validation. أي سيكون هناك 5 مجموعات (حسابات لتقدير القوة التنبؤية للمصنف، مع كل حساب يتم استخدام قسم مختلف (20%) كمجموعة اختبار والـ 80% المتبقية كمجموعة تدريب. يتم أخذ متوسط نتيجة هذه الحسابات الخمسة كأفضل تقدير للقدرة التنبؤية للمصنف.

في بعض الأحيان، عندما تكون مجموعة البيانات كبيرة بما فيه الكفاية (مثل حالة الدراسة لدينا)، فإن إجراء التحقق من الصحة بمقدار 5 مجموعات سيستغرق وقتاً طويلاً. لذلك فمن المقبول عادةً تقسيم مجموعة البيانات الكبيرة إلى 80% للتدريب و20% للاختبار، ويتم أخذ النتائج في مجموعة الاختبار كتقدير لأداء المصنف عند نشره لتصنيف بيانات جديدة.

تحليل الشذوذ (الحالات غير الطبيعية) Abnormality:

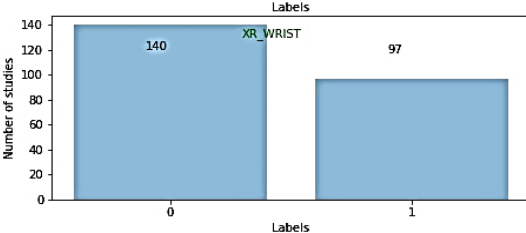
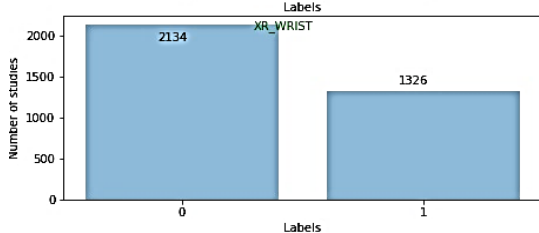
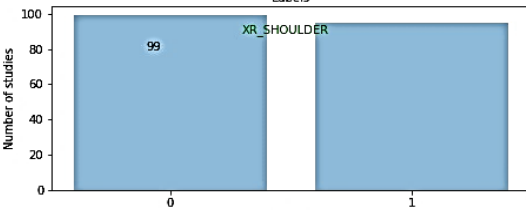
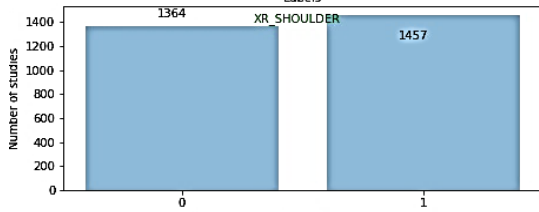
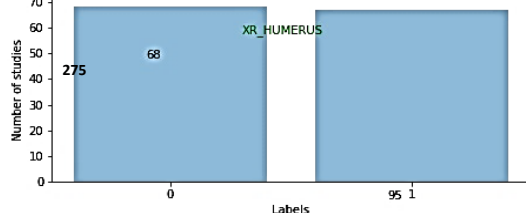
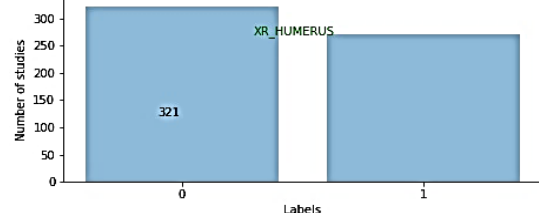
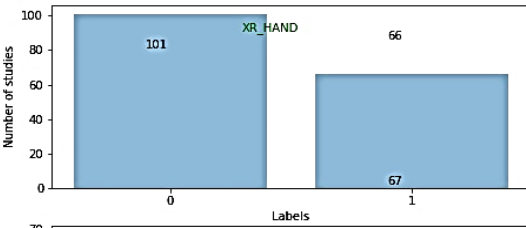
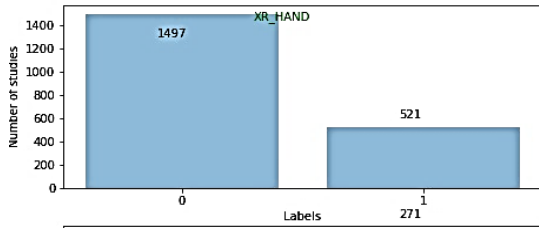
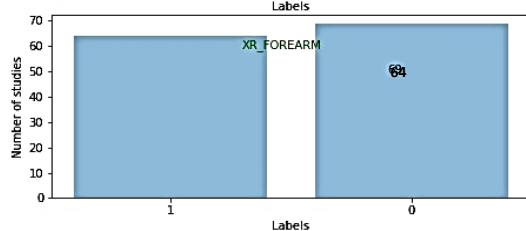
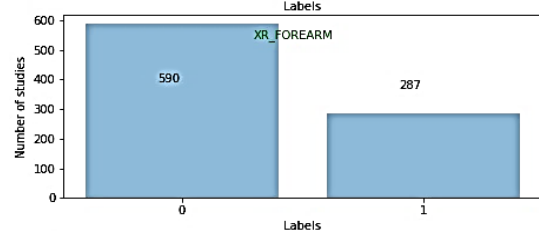
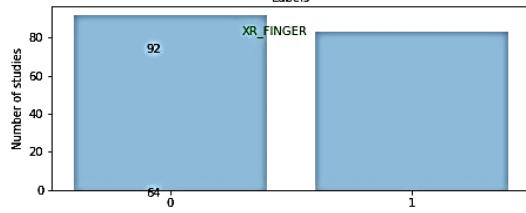
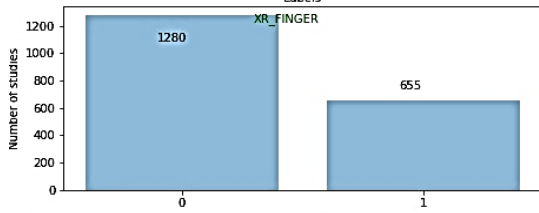
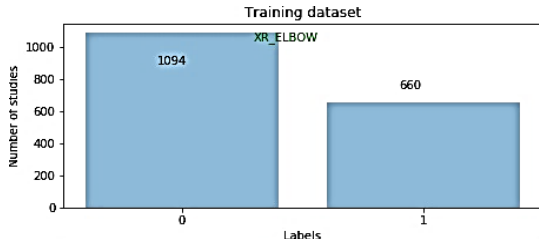
للتحقق من أنواع التشوهات الموجودة في مجموعة البيانات، تمت مراجعة تقارير أخصائي الأشعة لتمييز 100 دراسة غير طبيعية يدوياً مع اكتشاف الشذوذ: بحيث تم تصنيف 53 دراسة بالكسور، و48 دراسة مع الأجهزة، و35 دراسة مع أمراض تنكسية، و29 دراسة مع تشوهات أخرى، بما في ذلك الآفات والخلع الفرعي [21]. تم التعامل مع مجموعة البيانات من خلال إنشاء نوعين من البيانات الوصفية. واحدة تحتوي على معلومات لكل صورة مع التصنيف الخاص بها ونوع الدراسة. والأخرى تحتوي على كل نوع من أنواع الدراسة والتصنيف وعدد الصور كي يتمكن من رسم البيانات لفهم الخصائص المختلفة لها.



الشكل (3): مخطط إحصائي يمثل عدد المرضى في كل فئة من الفئات السبعة المحددة في نوعي عينة البيانات (التدريب والتحقق)

نرى في الشكل بيانات كل أنواع الدراسة ونلاحظ أن نوع المعصم يحتوي على أكبر عدد من البيانات مقارنة بأنواع الدراسة الأخرى ولكن لا يوجد اختلاف كبير في أحجام بيانات التحقق بين أنواع الدراسة.

تطبيق التعلم بالنقل في تصنيف الصور



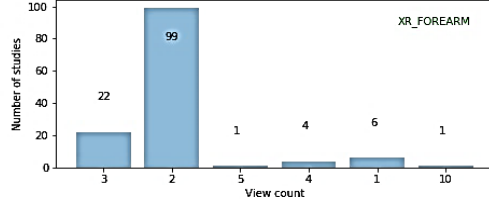
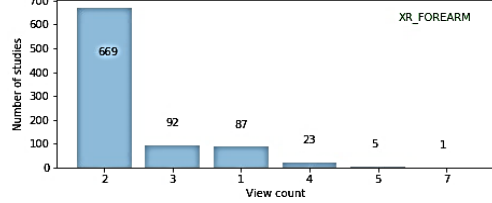
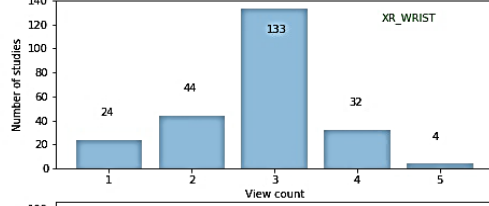
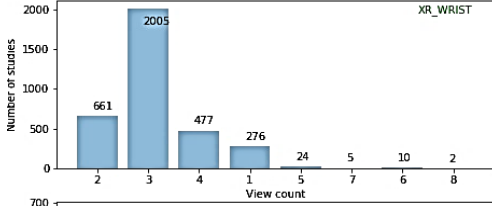
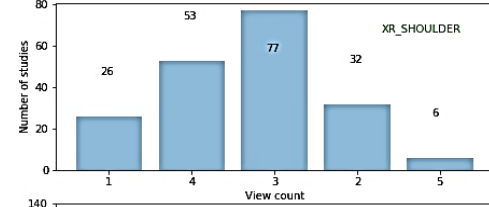
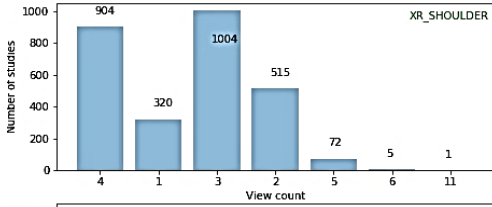
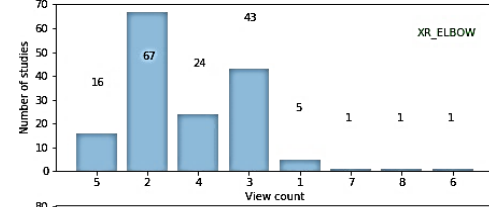
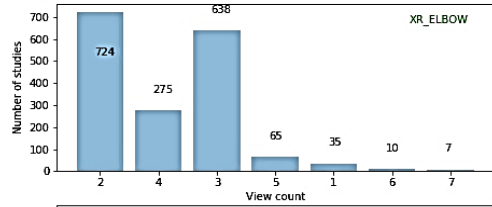
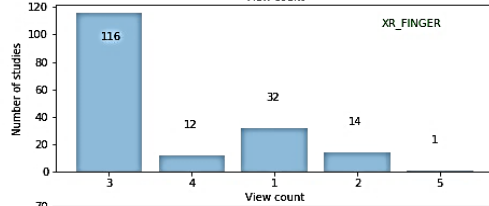
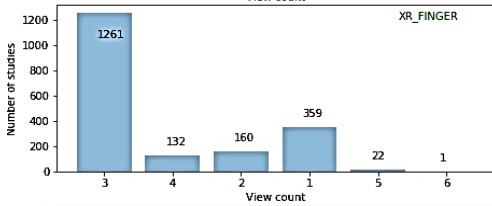
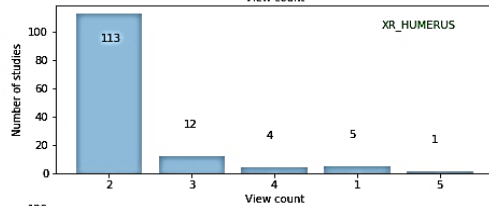
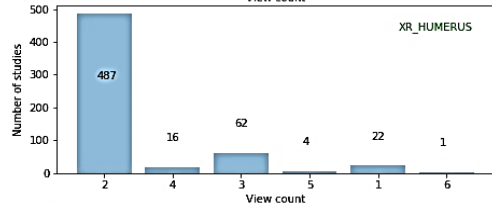
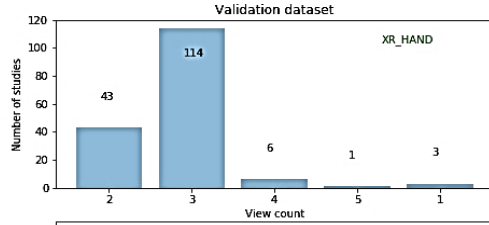
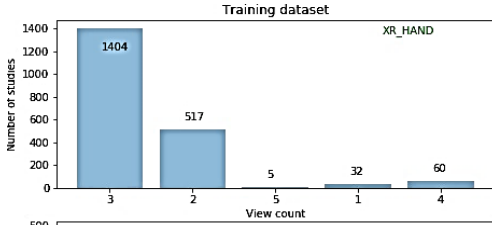
الشكل (4): يوضح توزيع أعداد الدراسات في كل دراسة (على مستوى المريض)

يتطلب التشخيص في بعض الأحيان دراسات متعددة عن الصور الشعاعية المختلفة. تمت ملاحظة أن أكثر من 90% من الحالات يحتاج المريض فيها إلى دراسة واحدة فقط. ومع ذلك، فمن الممكن أيضاً أن تتضمن كل دراسة صوراً شعاعية متعددة. قد يحتاج المريض إلى أخذ صور إشعاعية متعددة خلال كل دراسة اعتماداً على مدى تعقيد المشكلة.

استنتجنا من خلال دراسة عينة البيانات وتحليلها بشكل دقيق أن أنواع الدراسة "الساعد والكتف واليد والمعصم والإصبع" تحتوي على 3 صور لكل دراسة في معظم دراساتنا، كما يحتوي نوع الدراسة "الكوع" على صورتين لكل دراسة في معظم دراساتنا.

يوضح الشكل الآتي عدد الدراسات في كل نوع من الأنواع المدروسة مع عدد الصور (views) في كل دراسة.

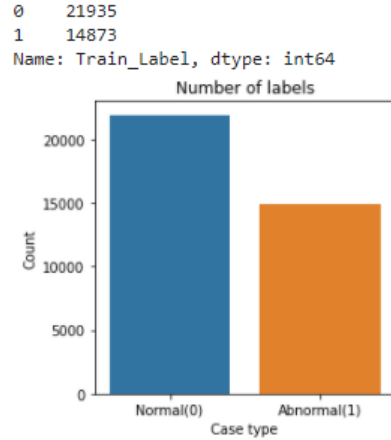
تطبيق التعلم بالنقل في تصنيف الصور



الشكل (5): توزيع التصنيف في البيانات لكل أنواع الدراسة

المعالجة المسبقة للبيانات:

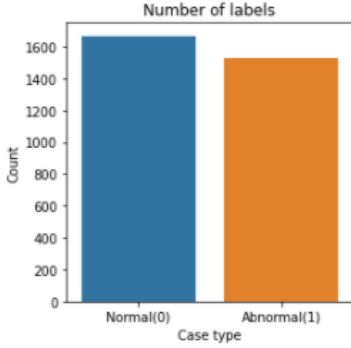
وفقاً للورقة الرسمية المنشورة من جامعة ستانفورد [21] ، فإنه يتم القيام بتطبيع كل صورة بحيث يكون لها نفس المتوسط والانحراف المعياري للصور في مجموعة التدريب ImageNet ، تم استخدام صور متغيرة الحجم وتحويلها لحجم ثابت 320×320 . وخلال دراستنا تم الاعتماد على قياس 224×224 . كما قمنا بزيادة البيانات أثناء التدريب عن طريق تطبيق الانعكاسات الجانبية العشوائية والتدوير حتى 30 درجة باستخدام ImageDataGenerator



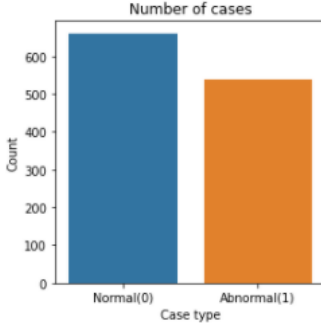
عدد التصنيفات في عينة

العدد الفردي (التفصيلي) لتصنيفات الصور في مجموعة التدريب
بيانات التدريب كاملة (عدد الدراسات)

```
0 1667
1 1530
Name: Valid_Label, dtype: int64
```



```
0 661
1 538
Name: Valid_Label, dtype: int64
```



عدد التصنيفات في

العدد الفردي لتسميات الصور في مجموعة التحقق من الصحة
مجموعة التحقق من الصحة (عدد الدراسات)

الشكل (6): مخططات إحصائية لعينة الدراسة (عدد الدراسات وعدد الصور في كل من بيانات التدريب وبيانات التحقق من الصحة)

مسابقة MURA

نشرت مجموعة التعلم الآلي في جامعة ستانفورد ورقة بحثية [21] تتعلق بمشكلة تصنيف الصور الشعاعية الطبية وقدمت واحدة من أكبر مجموعة بيانات صور شعاعية عامة في العالم تسمى MURA. MURA هي اختصار للتصوير الشعاعي العضلي الهيكلي (Musculoskeletal Radiographs). استخدمت مجموعة ML بجامعة ستانفورد خوارزمية DenseNet169 لتدريب شبكة عصبية عميقة يمكنها الكشف عن التشوهات في الصور الشعاعية بدقة أقرب إلى كبار أطباء الأشعة.

وكان الهدف من مسابقة MURA هو تصنيف كل دراسة إلى طبيعية أو غير طبيعية (تنبؤات ثنائية)، وليس كل صورة. وكان أفضل أداء في جامعة ستانفورد. في مجموعة البيانات هذه، قام الفريق الذي نشرها بتدريب شبكة تلافيفية متصلة بكثافة DenseNet169 مكونة من 169 طبقة للكشف عن التشوهات وتحديد مكانها.

6- مواد وطرق البحث:

قمنا بتطبيق عدد من نماذج التعلم بالنقل المتوفرة من خلال تدريب نموذج لكل منها ثم قمنا بمقارنة أداء النموذج وأخصائي الأشعة (والذي تم أيضاً إعطاؤهم الصور وقاموا بتصنيفها يدوياً، حيث قام 6 من أخصائي الأشعة بالعمل على ذلك وفقاً للورقة البحثية الخاصة بالبيانات المنشورة) [21]. نود أن يحقق نموذجنا أداءً مشابهاً لأداء أطباء الأشعة. بالنسبة للدراسات التي أجريت على اليد والرسغ والمعصم، فإن نتائج النموذج الخاصة بنا أعلى قليلاً من ناحية الدقة من تلك النتائج الخاصة بأطباء الأشعة؛ وفي دراسات الكوع والساعد والعضد والكتف، كانت الدقة الخاصة بنموذجنا أقل قليلاً من درجات أخصائي الأشعة، ولكنها ليست ذات دلالة إحصائية، مما يشير ويؤكد إلى أن مجموعة البيانات تمثل مشكلة تحدي جيدة للبحث في المستقبل.

التجارب:

بيئة التكوين وإعداد البارامترات:

تم تطوير جميع التجارب وتنفيذها على منصة Google Colaboratory السحابية باستخدام Keras مع TensorFlow 2.4.0. **Colaboratory** أو Colab اختصاراً، هو أحد منتجات Google Research، والذي يسمح للمطورين بكتابة وتنفيذ كود Python من خلال متصفحهم. يعد Google Colab أداة ممتازة لمهام التعلم العميق. كما يعتبر فعلياً كإصدار

متخصص من Jupyter Notebook، والذي يعمل على السحابة ويوفر موارد حوسبة مجانية مثل GPUs وTPU.

هناك عدة أسباب لاختيار استخدام Google Colab بدلاً من Jupyter Notebook العادي، وأهمها:

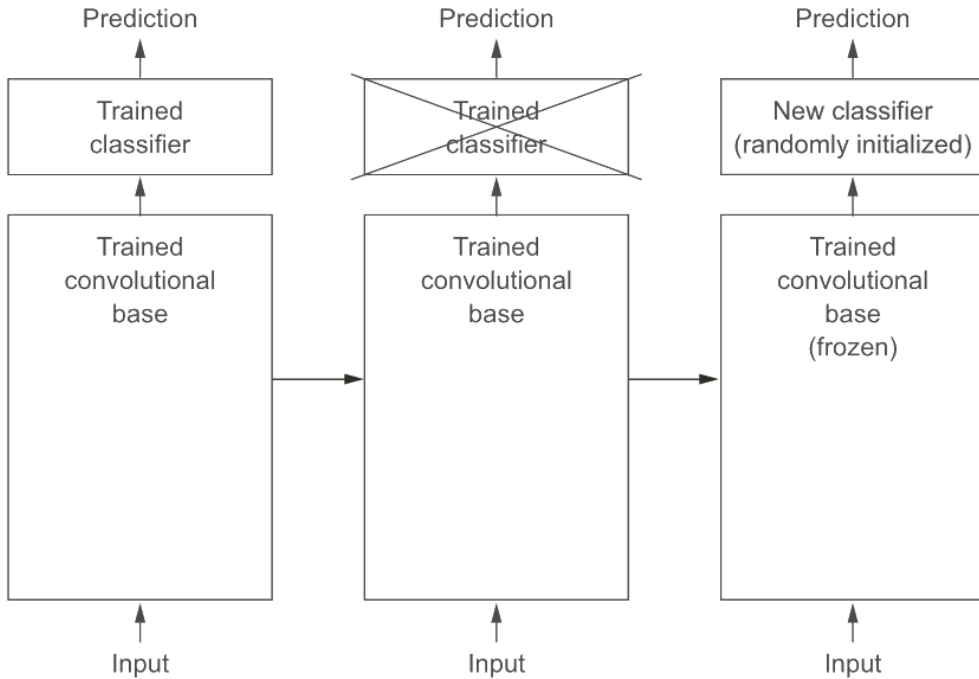
- وجود مكتبات مثبتة مسبقاً (Keras, TensorFlow, and PyTorch مثلاً)
 - يتم الحفظ على السحابة (يتم حفظ جميع دفاتر Google Colab ضمن حساب Google Drive، تماماً مثل ملفات Google Docs و Google Sheets)
 - التعاون (وهي ميزة أخرى يقدمها Colab فإذا كنت تعمل مع مطورين متعددين في مشروع ما، فمن الأفضل استخدام Colab. تماماً مثل التعاون في مستند محرر مستندات Google، كما يمكن البرمجة المشتركة مع مطورين متعددين باستخدام Colab notebooks)
 - استخدام GPU و TPU مجاناً (تتيح Google Research استخدام وحدات معالجة الرسومات (GPU) و (TPU) المخصصة لمشاريع التعلم الآلي كون تسريع GPU و TPU يحدثان فرقاً كبيراً حتى في المشاريع الصغيرة)
- لمقارنة أداء الطرق المختلفة للتعلم بالنقل، قمنا بتعيين نفس البارامترات لكل الطرق المستخدمة بدون مواصفات خاصة، البارامترات هي: معدل التعلم الأولي 0.0001، عدد الفترات القصوى هو 100 فترة.
- نماذج شبكات التعلم التي قمنا بتجريبها هي ثلاثة نماذج، وهي DenseNet169، VGG16، DenseNet201:
- ولكن اكتفينا بوضع نتائج DenseNet201 كونها حققت الأداء الأفضل بينها جميعاً.

الاستراتيجيات المعتمدة في التعلم بالنقل:

تتم معالجة مهمة تصنيف الصور الموضحة سابقاً من خلال تبني الطريقتين الشائعتين اللتين يتم استخدامهما بشكل شائع عند تطبيق التعلم بالنقل باستخدام شبكات CNN المدربة مسبقاً، وهما استخراج الميزات والضبط الدقيق [23].

1- استخراج الميزات:

يتكون استخراج الميزات أساساً من أخذ القاعدة التلافيفية لشبكة مُدرّبة مسبقاً، وتشغيل البيانات المستهدفة من خلالها وتدريب مُصنّف جديد على الناتج (أي الاستفادة من الشبكة أو النموذج المُدرّب مسبقاً والاستفادة من الميزات منها في تدريب شبكة أو نموذج جديد لمسألة تصنيف جديدة)، كما هو موضح في الشكل الآتي:



الشكل (7): توضيح تطبيق استخراج الميزات على شبكة عصبية تلافيفية، بحيث يتم

تبديل المصنفات بينما يتم الاحتفاظ بنفس القاعدة التلافيفية. [23]

ملاحظة: تعني كلمة "مجمد" في كتلة المصنف الجديد أن الأوزان لم يتم تحديثها أثناء التدريب.

يمكن أن يكون المصنف المكس أعلى القاعدة التلافيفية إما مجموعة من الطبقات المتصلة بالكامل أو مجرد طبقة تجميع global واحدة، يتبعها طبقة كثيفة مع تابع تنشيط softmax.

اعتماداً على ما ورد في الدراسات السابقة فإنه لا توجد قاعدة محددة بشأن نوع المصنف الذي يجب اعتماده، ولكن كما وصفه لين وآخرون. آل [2]، يؤدي استخدام طبقة Global Pooling واحدة بشكل عام إلى تقليل التخصيص وذلك نظراً لعدم وجود معلمات للتحسين في هذه الطبقة.

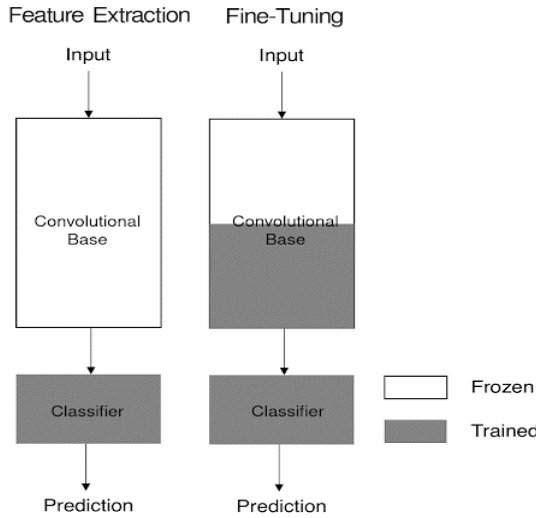
وبالتالي، ولأن مجموعات التدريب المستخدمة في التجارب صغيرة نسبياً، فإن المصنّف المختار يتكون فقط من طبقة تجميع متوسط global واحدة يتم تغذية المخرجات مباشرة في طبقة تنشيط softmax والتي تنتج الاحتمالات لكل فئة من فئات التصنيف المدروسة.

يتم تحديث أوزان المصنفات العليا فقط أثناء التدريب، بينما يتم "تجميد" أوزان القاعدة التلافيفية وبالتالي الاحتفاظ بها دون تغيير. بهذه الطريقة، يتعلم المصنّف الضحل كيفية تصنيف الصور إلى فئتين من التمثيلات الجاهزة التي تعلمها مسبقاً نموذج المصدر لمجالها. وإذا كان المصدر والمجال الهدف متشابهين، فمن المحتمل أن تكون هذه التمثيلات مفيدة للمصنّف ويمكن أن تؤدي المعرفة المنقولة بالتالي إلى تحسين أدائها بمجرد تدريبها.

2- التكوين المثالي / Fine Tuning:

يمكن النظر إلى التكوين المثالي أو الضبط الدقيق على أنه خطوة مختلفة على عكس استخراج الميزات التي تتكون من إعادة تدريب اختياري لبعض الطبقات العليا للقاعدة التلافيفية المستخدمة سابقاً لاستخراج الميزات. وبهذه الطريقة، يتم تعديل التمثيلات الأكثر تجريباً للنموذج المصدر التي تعلمتها طبقاته الأخيرة بشكل طفيف لجعلها أكثر صلة بالمشكلة الهدف (أو المشكلة المدروسة).

يمكن تحقيق ذلك من خلال إلغاء تجميد بعض الطبقات العليا للقاعدة التلافيفية، والحفاظ على تجميد جميع طبقاتها الأخرى وتدريب القاعدة الالتفافية بشكل مشترك باستخدام المصنّف نفسه المستخدم سابقاً لاستخراج الميزات، كما هو موضح في الشكل أدناه.



الشكل (8): استخراج الميزات مقارنةً بالضبط الدقيق [23]

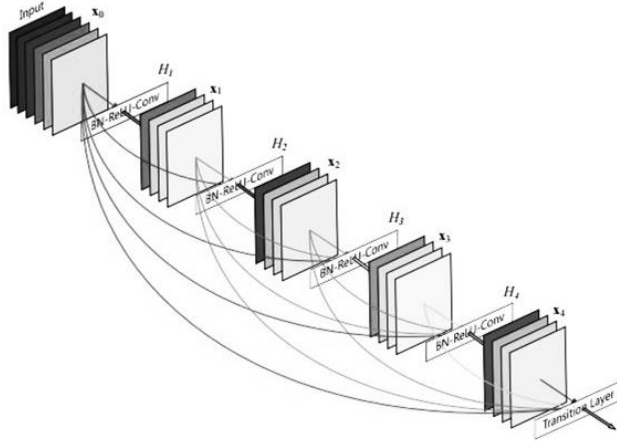
من المهم الإشارة إلى أنه وفقاً لـ F. Chollet [24]، لا يمكن ضبط الطبقات العليا لقاعدة تلافيفية مُدرّبة مسبقاً إلا إذا كان المصنّف الموجود فوقها قد تم تدريبه مسبقاً بالفعل. والسبب هو أنه إذا لم يكن المصنّف مدرّباً بالفعل، فسيتم تهيئة أوزانه بشكل عشوائي. نتيجة لذلك، ستكون إشارة الخطأ التي تنتشر عبر الشبكة أثناء التدريب كبيرة جداً وسيتم تحديث الأوزان غير المجمدة لتعطيل التمثيلات المجردة التي تعلمتها القاعدة التلافيفية سابقاً.

ولأسباب مشابهة، يوصى أيضاً بإجراء الضبط الدقيق باستخدام معدل تعلم أقل من المعدل المستخدم لاستخراج الميزات.

والجدير بالذكر أن السبب وراء إلغاء تجميد الطبقات العليا فقط هو أن الطبقات السفلية تشير إلى ميزات عامة مستقلة عن المشكلة، بينما تشير الطبقات العليا إلى الميزات المعتمدة على المشكلة والتي ترتبط بشكل أكبر بالمجال المحدد والذي تم تدريب الشبكة في الأصل من أجله. وبالتالي، فإن الميزات التي تعلمتها الطبقات الأولى مناسبة بشكل عام لمعالجة مجموعة واسعة من المجالات، في حين أن الميزات التي تعلمتها الطبقات العليا تحتاج إلى تعديل لكل مجال محدد [25].

DenseNet

من بين جميع شبكات CNN العميقة المقترحة حديثاً، كانت DenseNet معروفة بالأداء الرائع في المهام المعيارية للتعرف على الأشياء التنافسية مثل ImageNet و [7] [CIFAR-100] [25]. ويعود ظهور شبكات CNN العميقة إلى عام 2012. في عام 2012، فازت شبكة ذات بنية جديدة تسمى AlexNet، وهي المحاولة الأولى لحل تحدي تصنيف الصور على نطاق واسع بواسطة شبكة CNN العميقة [2]، بالمركز الأول والثاني في مهمة التعريب والتصنيف على التوالي. قد يكون عمق AlexNet، الذي يتكون فقط من خمس طبقات تلافيفية وثلاث طبقات متصلة تماماً، قليلاً جداً مقارنةً بعمق شبكات CNN العميقة الحديثة المصممة في السنوات الأخيرة، ولكنها أشارت إلى مستقبل واعد لشبكات CNN العميقة والذي تم إثباته من خلال النمو الهائل في عدد نماذج لشبكات CNN عميقة جديدة. الشكل الآتي يوضح بنية شبكة DenseNet مكونة من 5 طبقات:



الشكل (9): بنية شبكة DenseNet مكونة من 5 طبقات وبمعدل نمو $k=4$ بحيث تأخذ

كل طبقة جميع خرائط البارامترات السابقة كمدخل لها

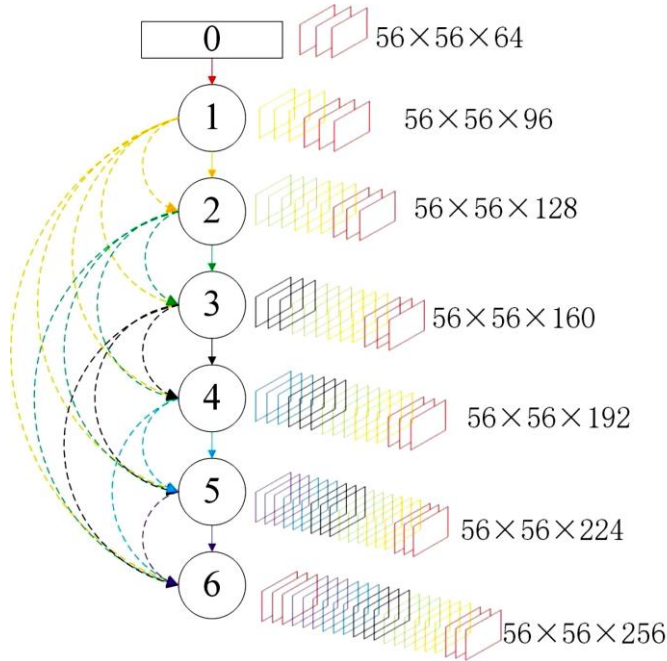
بالنسبة للشبكات المطورة حديثاً، تبين أن العديد من الاتصالات المختصرة هي طرق فعالة للتغلب على المشكلات المحتملة مثل تلاشي التدرج عند تدريب شبكات CNN العميقة. في [26]، حلت وحدة البداية المحسنة المكونة من فرع مختصر وعدد قليل

من الفروع العميقة مشكلة تلاشي / انفجار التدرجات عندما تتعمق الشبكات. لتسهيل عملية تدريب الشبكات العميقة، تم تصميم الشبكات المسماة شبكات الطرق السريعة للسماح بتدفق المعلومات التي تنتجها الطبقات السابقة إلى الطبقات اللاحقة دون فقدان المعلومات. تم تحقيق نفس الهدف ولكن باستخدام أسلوب التعلم المتبقي في [14]. في حين أن الاستفادة من البنى العميقة والواسعة يمكن أن تكون تقنيات مفيدة لتحسين أداء شبكات CNN العميقة، فإن DenseNet تدرك هدف التدريب السهل وكفاءة البارامترات من خلال إعادة استخدام الميزات، والتي تُدخل خرائط الميزات المتسلسلة التي تنتجها جميع الطبقات السابقة في الطبقة اللاحقة. بهذه الطريقة، يُسمح للطبقات العميقة في الشبكات بالوصول إلى جميع خرائط البارامترات التي تنتجها الطبقات السابقة وبالتالي إعادة استخدام الميزات.

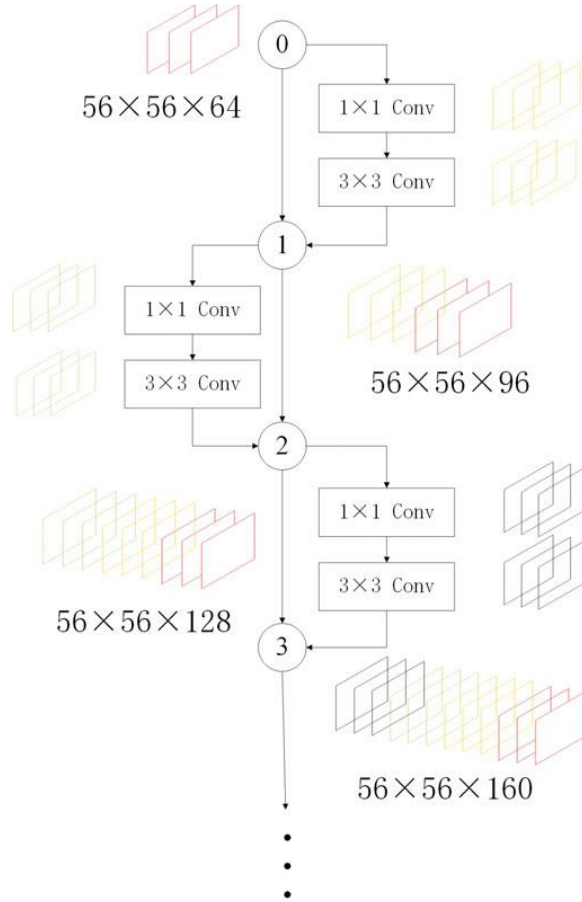
في شبكات CNN التقليدية العميقة، كانت الطبقات التلافيفية متبوعة عموماً بطبقات تصغير الحجم تقلل عرض خرائط البارامترات وارتفاعها إلى نصفها. وبالتالي، فإن تسلسل خرائط البارامترات قبل طبقات العينات السفلية وبعدها سيكون مشكلة بسبب الأحجام المختلفة. لحل هذه المشكلة، تم تصميم الكتل الكثيفة بمهارة متبوعة بطبقات تأخذ العينات السفلية بينما ترتبط الطبقات في الكتل الكثيفة بشكل كامل. ونتيجة لذلك، ظلت أحجام خرائط البارامترات في الكتل الكثيفة دون تغيير بينما تم تخفيضها إلى النصف بعد أخذ العينات للأسفل. لذلك، بالنسبة للكتلة الكثيفة ذات الطبقات L ، يكون العدد الإجمالي للتوصيلات المباشرة بين الطبقات هو $(L + 1) / 2$ في حين أن عدد التوصيلات في الشبكة التلافيفية التقليدية مع عدد طبقات التفاف L هو فقط L طبقات كلما زاد عدد مدخلات خرائط البارامترات المتسلسلة للطبقات التالية.

إذا لم تكن هناك قيود على النمو الخطي في عدد خرائط الميزات، فإن كلفة الحساب الضخمة ستكون كبيرة جداً. لذلك تم تصميم معدل النمو k للتحكم في عدد خرائط البارامترات المنتجة في كل طبقة. ونتيجة لذلك، فإن العدد الإجمالي لخرائط البارامترات في الطبقة l من كتلة كثيفة هو $k * (l - 1) + k_0$ ، حيث k_0 هو عدد القنوات في طبقة الإدخال.

لتقليل تعقيد الحساب، تم إدخال طبقات عنق الزجاجة (Bottleneck Structure) عن طريق استخدام النفاث 1×1 قبل كل النفاث 3×3 . إلى جانب طبقات الزجاجة، تعمل الطبقات المسماة طبقات الانتقال على التحكم في عدد خرائط ميزات الإخراج في عمق معين للشبكات وبالتالي تحسين ضغط الشبكات بأكملها. يوضح الشكل الآتي عملية ربط وإنتاج خرائط ميزات جديدة في أول كتلة كثيفة من DenseNet201. نظراً لأن الطبقة الانتقالية تظهر عموماً بعد كتلة كثيفة بعمق محدد، فلا توجد طبقة انتقالية بعد الكتلة الكثيفة الأولى.



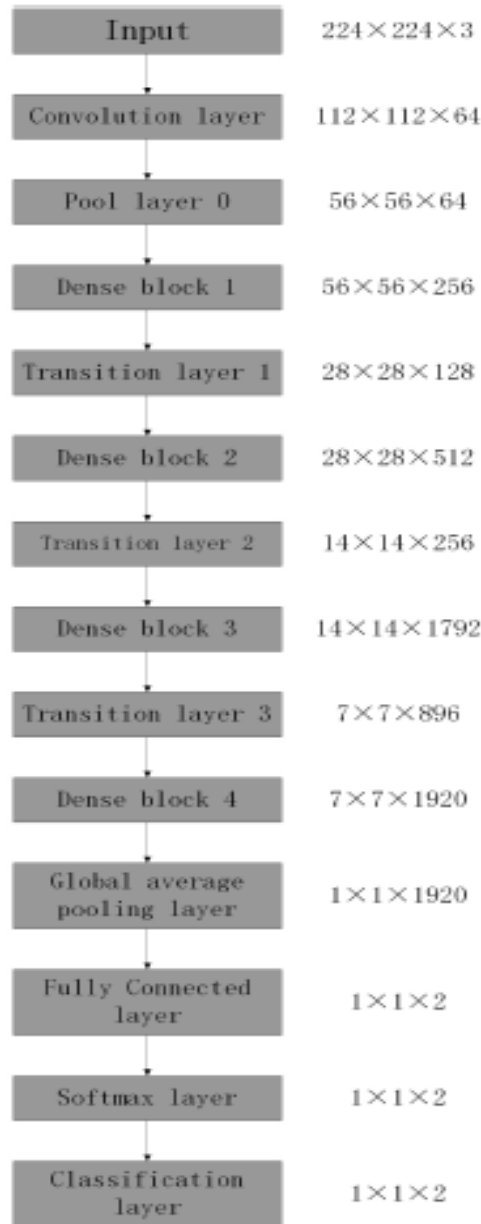
الشكل (10): خرائط البارامترات المتسلسلة في كل طبقة، تشير الخطوط المتصلة إلى تسلسل خرائط البارامترات السابقة وخرائط البارامترات المنتجة في الطبقة التالية حيث يشير السهم إليها كما تقدم الخطوط المتقطعة اتصال طبقات مختلفة



الشكل (11): اتصالات أول كتلة كثيفة في DenseNet201 (معدل النمو $k = 32$)

(هنا

تم إنتاج الخرائط المميزة حديثاً في كل طبقة.



الشكل (12): هيكلية DenseNet201 المنقولة المستخدمة في عملية التدريب والتقييم
Dense Block: (أو ما يسمى بـ الكتلة الكثيفة) هي وحدة مستخدمة في الشبكات العصبية التلافيفية التي تربط جميع الطبقات (بأحجام خرائط ميزات متطابقة) مع بعضها البعض مباشرة. تم اقتراحها في الأصل كجزء من بنية DenseNet.

تتكون الكتلة الكثيفة من كتل التقاف متعددة، كل منها يستخدم نفس عدد قنوات الإخراج. ومع ذلك، في الانتشار الأمامي، نقوم بربط المدخلات والمخرجات لكل كتلة التقاف على بُعد القناة.

للحفاظ على طبيعة التغذية الأمامية، تحصل كل طبقة على مدخلات إضافية من كل الطبقات السابقة وتكرر خرائط الميزات الخاصة بها إلى كل الطبقات اللاحقة. وبالتالي، تحتوي الطبقة l على l مدخل، تتكون من خرائط الميزات لجميع الكتل التلافيفية السابقة. يتم تمرير خرائط الميزات الخاصة بها إلى جميع الطبقات اللاحقة $l - L$. وهذا يقدم $L(L+1)/2$ اتصال في شبكة ذات L طبقة، بدلاً من مجرد L اتصال، كما هو الحال في البنى التقليدية.

Transition Layer (طبقة النقل): نظراً لأن كل كتلة كثيفة Dense Block ستزيد من عدد القنوات، فإن إضافة الكثير منها سيؤدي إلى نموذج معقد للغاية. بالتالي يتم استخدام طبقة نقل للتحكم في مدى تعقيد النموذج ما يقلل من عدد القنوات باستخدام طبقة تلافيفية 1×1 وتقليل ارتفاع وعرض طبقة التجميع المتوسطة بخطوة 2، مما يقلل بشكل أكبر من تعقيد النموذج. بشكل أبسط، يمكن القول إن طبقات الانتقال (التفاف + التجميع) هي طريقة لاختزال التمثيلات المحسوبة بواسطة Blocks Dense ببطء حتى النهاية. وتتكون طبقة النقل من الأجزاء الآتية:

Batch Normalization

1x1 Convolution

Average pooling

7- دراسة النتائج العملية

التدريب:

تمت تهيئة الشبكة بأوزان من نموذج شبكة DenseNet201 تم تدريبه مسبقاً على عينة البيانات الشهيرة imagenet وتم تدريب الشبكة من طرف إلى طرف باستخدام Adam Optimizer مع البارامترات الافتراضية. استخدمنا نفس حجم batch = 8

ومعدل التعلم 0.0001 قمنا بتدريب الشبكات لكل أنواع الدراسة بشكل منفصل بحيث قمنا بتطبيق نموذج التعلم بالنقل على عينة البيانات المدروسة على كل نوع بشكل مستقل (أي التطبيق على فئة المعصم بشكل مستقل، على فئة الأصابع بشكل مستقل... الخ) ويعدد مرات 100 لكل الأنواع، كما قمنا بمتابعة تدريب النموذج على عينة البيانات كاملةً دون فصل الأنواع بشكل مستقل، وقمنا بتسجيل النتائج لكل الحالات المدروسة. ونعرضها في القسم الآتي:

تقييم الأداء (وفقاً لدقة التصنيف):

قمنا بتقييم أداء النموذج على الاختبار الذي تم تعيينه حسب أنواع الدراسة ومقارنته بنموذج المجموعة. تم إنشاء نموذج المجموعة باستخدام النماذج الأفضل أداءً من كل أنواع الدراسة. تمت مقارنة الأداء باستخدام مقياس الدقة. لاحظنا أن أداء النماذج الفردية المدربة على أنواع الدراسة المحددة أفضل بكثير مقارنةً بنموذج المجموعة. وعلى الرغم من أنه يمكن استخدام نموذج المجموعة للتنبؤ بأي نوع من أنواع الدراسة، إلا أن اختلاف الأداء في بعض أنواع الدراسة كان ضعيفاً.

ملاحظات حول النتائج:

نوضح في الجدول أدناه نتائج المقارنة (من ناحية دقة التصنيف)، التي قمنا بإجرائها بين نتائج الورقة البحثية ونتائج دراستنا لتطبيق التعلم بالنقل على تصنيف الصور الشعاعية الطبية [21].

نلاحظ من الجدول أننا حصلنا على نتائج أفضل (دقة) في أغلب حالات الدراسة ضمن عينة البيانات-لاحظنا أن أفضل تحسين في قيمة الدقة كانت بالنسبة لنوع HUMERUS، كما أننا قمنا بإجراء تدريب للنموذج على عينة البيانات كاملةً بغض النظر عن نوع الدراسة (تدريب لكل الصور في عينة البيانات)، بحيث وصلنا إلى دقة تصل 82% تقريباً وهي نسبة مقبولة نوعاً ما مقارنة بالدقة التي حصلنا عليها لكل نوع بشكل مستقل، كون بعض حالات الدراسة أعطت دقة أقل من 82% مثل

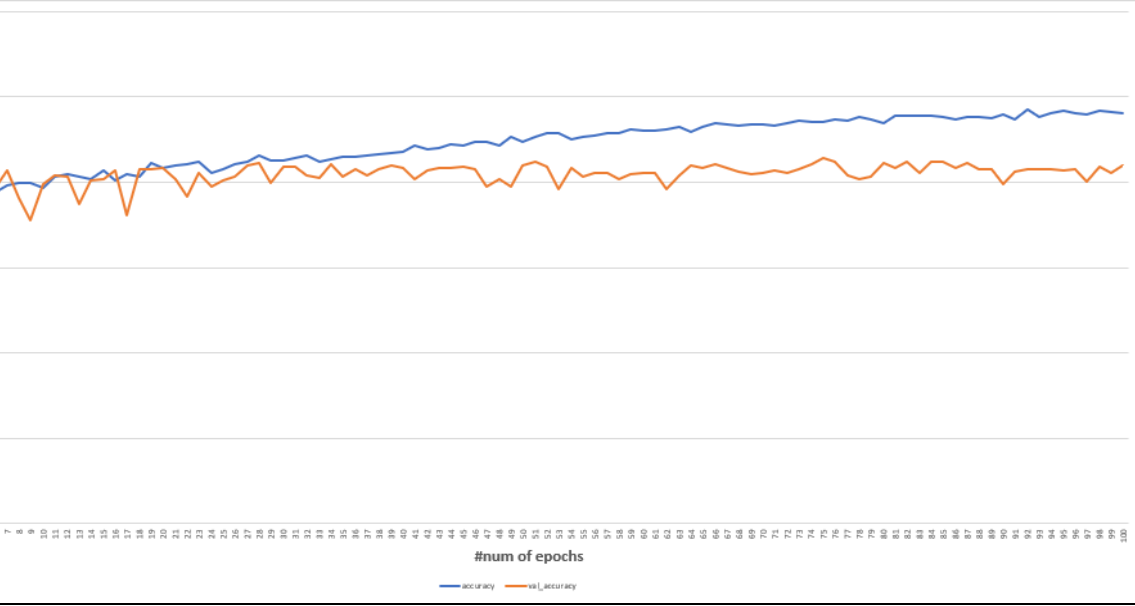
حالة SHOULDER و FINGER و HAND أما الأنواع الأربعة الباقية أعطت نتائج بدقة أعلى.

Study Type	Ref. Precision	Precision
WRIST	0.86	0.862
ELBOW	0.85	0.8778
HUMERUS	0.82	0.8819
SHOULDER	0.79	0.7964
FINGER	0.78	0.7969
FOREARM	0.77	0.8194
HAND	0.73	0.7969
ALL	-	0.8191

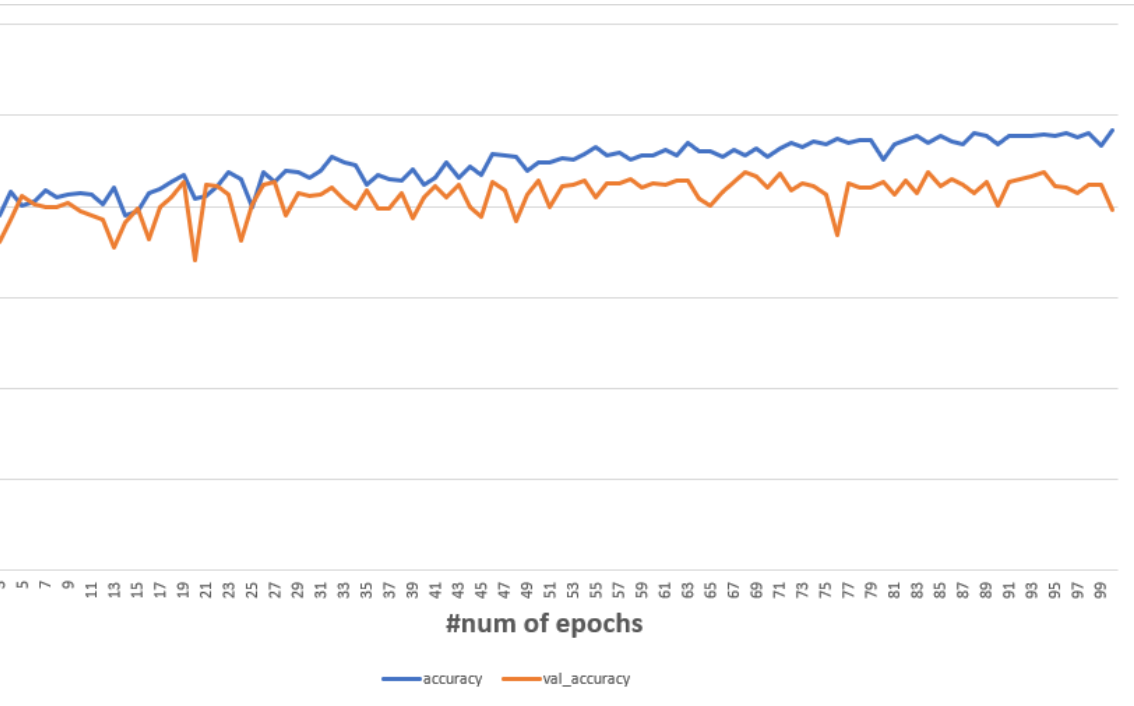
نلاحظ من جدول المقارنة بين النتائج التي حصلنا عليها بتطبيق التعلم بالنقل من خلال شبكة DenseNet201 والنتائج التي حصل عليها الباحثون في المرجع المعتمد لعينة البيانات أننا حصلنا على دقة أعلى في كل أنواع الدراسة وكان فرق الدقة التي حصلنا عليها في نوع HUMERUS الأفضل بين الأنواع الأخرى فقد كان تحسين الدقة من 82% إلى 88%، كما نلاحظ التحسين الثاني والذي يعتبر تحسناً قوياً هو في نوع HAND بحيث أننا حصلنا على دقة بنسبة 79% مقابل دقة 73% كانت في نتائج المرجع، كما أن هناك تحسين في الأنواع الأخرى ولو كان بنسبة غير عالية جداً كما في النوعين المذكورين.

قمنا أيضاً بتدريب النموذج على عينة البيانات كاملة بغض النظر عن النموذج وحصلنا على دقة تقارب 82% وهي تعتبر نسبة جيدة خاصة أنها تفوقت على بعض الحالات الخاصة المدروسة مثل FINGER و SHOULDER اللتين حققنا نسبة 79% تقريباً.

نوضح في المخططين الآتيين نتائج تدريب النموذج على عينة البيانات لنوعين من أنواع الدراسة.



المخطط (1): نتائج تدريب النموذج على عينة البيانات من نوع Elbow
 نلاحظ من المخطط (1) أن دقة التدريب استمرت في التصاعد واقتربت من الدقة
 المثالية أثناء الاستمرار في التدريب إلى 100 دورة بحيث وصلت إلى 98% في نوع
 Elbow بينما وصلت دقة التحقق إلى 85.6%، بحيث بدأت بقيمة صغيرة 72%
 من ثم تصاعدت ووصلت لدقة 85.6% والتي قمنا بتسجيلها عند الدورة 75 ولم
 تتجاوز هذه الدقة في الدورات الباقية.



المخطط (2): نتائج تدريب النموذج على عينة البيانات من نوع Humerus
 نلاحظ من المخطط (2) أن دقة التدريب استمرت في التصاعد واقتربت من الدقة
 المثالية أثناء الاستمرار في التدريب إلى 100 دورة ووصلت إلى حوالي 97%، أما
 دقة التحقق بدأت بقيمة صغيرة 68% من ثم تصاعدت ووصلت لدقة 88% التي
 قمنا بتسجيلها عند أكثر من نقطة (دورة).

8- النتائج والتوصيات:

يمكننا تقسيم النتائج إلى نتائج عامة، ونتائج خاصة بنقل التعلم إلى عينة البيانات
 الخاصة بالبحث.

النتائج العامة:

بعد العمل على مجموعة البيانات المذكورة وتطبيق طرق التعلم بالنقل عليها في
 محاولة من خلال بحثنا للحصول على نتائج أفضل، وجدنا ما يلي:

- يتطلب تدريب نموذج التعلم العميق Deep Learning باستخدام الصور الطبية العديد من الاختيارات (اختيار البارامترات الصحيحة والبارامترات الفائقة) وبدون وحدة معالجة الرسومات السريعة، يستغرق الأمر عدة أيام للحصول على نتيجة (استخدمنا Google Colab). يمكن اعتبار هذه النقطة هي نتيجة عامة لمعظم الدراسات المتعلقة بمجال تعلم الآلة والتعلم العميق والتعلم بالنقل، إلا أننا قمنا بتجريبها في هذه الدراسة ولاحظنا هذه النقطة.
- عدم وجود معرفة بمجال الصور الطبية (حالة الدراسة) ليست مشكلة في الحصول على نموذج بمستوى جيد ولكن هذه المعرفة الطبية تصبح ضرورية لإنشاء تطبيق على المستوى المهني.

أما فيما يتعلق بالنتائج الخاصة بتطبيق نقل التعلم على عينة البيانات الخاصة ببحثنا فقد تم توضيحها بالتفصيل في قسم دراسة النتائج العملية من حيث الحصول على نتائج أفضل في عدة أقسام من المجال المدروس على عينة البيانات التي تم الاعتماد عليها، والاستفادة من النماذج المدربة مسبقاً في عمليات التصنيف لمشاكل عملية جديدة لا يتوفر فيها الكم المطلوب من البيانات اللازمة لإنجاز التصنيف والحصول على نتائج مرضية.

9- الأعمال المستقبلية:

يمكن توسيع هذا البحث ليشمل تطبيق أنواع أخرى من خوارزميات نقل التعلم على عينة البيانات ومقارنة نتائجها مع نتائج البحث الحالي.

10-المراجع:

- 1- Gao J., Jiang Q., Zhou B., Chen D., 2019 - Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: An overview, College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China.
- 2- Yuliana J.- Gaona, Maria J. R.- Alvarez, Vasudevan L., 2020 – Deep Learning Based Computer-Aided Systems for Breast Cancer Imaging: A Critical Review, Department of Systems Design Engineering, Physics, and Electrical and Computer Engineering, University of Waterloo, ON N2L3G1, Canada.
- 3- Dhungel N., Carneiro G., Bradley A. P., 2015 - Automated Mass Detection from Mammograms using Deep Learning and Random Forest, 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA).
- 4- Adelaide, SA, Australia Scott R., Stone N., Kendall, C., 2016 - Relationships between pathology and crystal structure in breast calcifications: an in situ X-ray diffraction study in histological sections. npj Breast Cancer 2, 16029.
- 5- Scott R., Kendall C., Stone N., Rogers K., 2017 - Elemental vs. phase composition of breast calcifications, Sci Rep.
- 6- Mendoza P., Lacambra M., Tan PH., Tse GM., 2011 - Fine Needle Aspiration Cytology of the Breast: The Nonmalignant Categories, Patholog Res Int.
- 7- Se-woon C., Kokeb D., Gelan A., 2021 - Transfer Learning in Breast Cancer Diagnoses via Ultrasound Imaging, Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi 39253, Korea.
- 8- Khan, A., Sohail, A., Zahoora, U. et al. ,2020 - A Survey of the Recent Architectures of Deep Convolutional Neural Networks, Deep Learning Lab, Center for Mathematical Sciences, PIEAS, Nilore, Islamabad 45650, Pakistan, Artif Intell Rev 53, 5455–5516.

- 9- Yamashita, R., Nishio, M., Do, R.K.G, 2018 - Convolutional neural networks: an overview and application in radiology, Insights Imaging 9, 611–629.
- 10-Jason Y., Jeff C., Yoshua B., Hod L., 2014 - How transferable are features in deep neural networks?, arXiv:1411.1792.
- 11-Weiming Z., Zhenghao C.Henry W. F. Y., Zhicheng L., 2017 - Layer Removal for Transfer Learning with Deep Convolutional Neural Networks, Conference: International Conference on Neural Information Processing.
- 12-Niall O' M., Sean C., Anderson C., Suman H., Gustavo V. H., Lenka K., Daniel R., Joseph W., 2019 - Deep Learning vs. Traditional Computer Vision, IMaR Technology Gateway, Institute of Technology Tralee, Tralee, Ireland.
- 13-van Engelen, J.E., Hoos, H.H., 2020 - A survey on semi-supervised learning, Mach Learn 109, 373–440.
- 14-Vrbančič G., podgorelec V., 2020 - Transfer Learning with Adaptive Fine-Tuning, IEEE, Faculty of Electrical Engineering and Computer Science, University of Maribor, 2000 Maribor, Slovenia.
- 15-Ravi K S., Heang-P.C. , Lubomir H. , Mark A H. , Jun W., Kenny C. , 2016 - Mass detection in digital breast tomosynthesis: Deep convolutional neural network with transfer learning from mammography, Med Phys.
- 16-Xiang Y., Nianyin Z., Shuai L., Yu-Dong Z., 2019 - Utilization of DenseNet201 for diagnosis of breast abnormality - Machine Vision and Applications.
- 17-Natalie B., Jordan O., Erik J. L. , 2020 - Exploring the efficacy of transfer learning in mining image-based software artifacts, Journal of Big Data.
- 18-Zhongling H., Zongxu P., Bin L. ,2017 - Transfer Learning with Deep Convolutional Neural Network for SAR Target Classification with Limited Labeled Data - Institute of Electronics, Chinese Academy of Sciences, Beijing, China.
- 19-www.usbji.org
- 20-[Musculoskeletal conditions \(who.int\)](http://www.who.int)
- 21-Pranav R., Jeremy I. , Aarti B. , Daisy D., Tony D., Hershel M., Brandon Y., Kaylie Z., Dillon L., Robyn L. B., Curtis

- L., Katie S., Matthew P. L., Andrew Y. Ng, 2018 - MURA: Large Dataset for Abnormality Detection in Musculoskeletal Radiographs, 1st Conference on Medical Imaging with Deep Learning.
- 22- <https://stanfordmlgroup.github.io/competitions/mura>
- 23-Sedrick C. , Michele Z. – 2020, Transfer Learning in Image Classification: how much training data do we really need?
- 24-Francois C., 2016 - Building powerful image classification models using very little data, The Keras Blog.
- 25-Jason Y., Jeff C., Bengio Y., Hod L. , 2014 -How transferable are features in deep neural networks?, Conference: Advances in Neural Information Processing Systems.
- 26-Rupesh K. S., Klaus G., Jürgen S., 2015 - Training Very Deep Networks, The Swiss AI Lab IDSIA / USI / SUPSI
- 27- Olaide N. O., Absalom E. E., 2017 - A State-of-the-Art Survey on Deep Learning Methods for Detection of Architectural Distortion from Digital Mammography, IEEE

