Improved Adaptive Pooling Techniques for Convolutional Neural Networks in Image Classification

Prepared by

Rasha Issa Yousef Foudeh

Supervisor By:

Prof. Nameer N. El.Emam

Abstract

This thesis aims to develop and implement an Adaptive Convolutional Neural Network (ACNN) architecture that integrates structured Mask-Pooling processing techniques to enhance the adaptability and efficiency of pooling mechanisms. The data used in this thesis is the CIFAR -10 dataset. The adaptive mask-pooling approach selects the pooling method based on the variance of the current 3x3 mask compared to the overall variance of the input image, improving feature extraction and classification performance across diverse image datasets. The thesis aims to enhance the CNN model's performance and reduce training time by adjusting the stride value based on threshold conditions.

The performance of the ACNN model is evaluated using a separate test dataset to assess critical metrics such as accuracy, precision, recall, and F1-score. Techniques like Mean Squared Error (MSE) and cross-validation are employed to evaluate the model's precision, effectiveness, and

robustness in image classification tasks. The research advances the understanding of feature representation and optimization in image classification by incorporating novel adaptive mask-pooling techniques into CNNs. The proposed ACNN model with adaptive mask-pooling can deliver quantifiable improvements in classification accuracy, robustness, and computational efficiency, benefiting real-world applications such as medical imaging, autonomous vehicles, and security systems. The effectiveness of the proposed methods is evaluated through extensive experiments on the collected datasets, providing insights into their potential impact on CNN execution and informing future advancements in adaptive mask-pooling strategies for improved image recognition capabilities. Furthermore, the results of the adaptive mask-pooling show that the optimal results of accuracy, F1-score, Precision, Recall, and MSE are 0.70, 0.70, 0.75, 0.64, and 0.054, respectively. In addition, the results of adaptive mask-pooling are better than traditional techniques by about 62%, 70%, and 89% from Max, Average, and Frequency pooling.

تحسين تقنيات التجميع التكيفي للشبكات العصبية التلافيفية في تصنيف الصور

اعداد

رشا عيسى يوسف فوده

اشراف

الأستاذ الدكتور نمير ناظم الإمام

الملخص

تهدف هذه الدراسة إلى تطوير وتنفيذ بنية للشبكة العصبية التلافيفية التكيفية (ACNN) تدمج تقنيات المعالجة القائمة على القناع الهيكلي لتعزيز القدرة على التكيف والكفاءة في آليات التجميع. تم استخدام مجموعة البيانات CIFAR −10 في هذه الدراسة. يعتمد نهج التجميع التكيفي على اختيار طريقة التجميع بناءً على تباين القناع 3X3 الحالي مقارنةً بالتباين العام للصورة المدخلة، مما يحسن استخراج الميزات الخاصة بالصور وأداء التصنيف الخاص لهذه المجموعة من الصور المتنوعة. يتم تقييم أداء نموذج ACNN باستخدام مجموعة بيانات الاختبار التي تكون جزء من مجموعة البيانات المختارة لهذه الدراسة لتقييم المقاييس الرئيسية مثل نسبة الدقة و نسبة التمثيل الدقيق ونسبة التذكر أو الاسترجاع و نتيجة درجة F1. يتم استخدام تقنيات مثل تحليل المصفوفة، ومتوسط الخطأ التربيعي(MSE) ، والتحقق المتقاطع لتقييم دقة النموذج وفعاليته وقوته في مهام تصنيف الصور. تُعزز هذه الدراسة فهم تمثيل الميزات والخصائص والتحسين في تصنيف الصور من خلال دمج تقنيات التجميع التكيفي الجديدة في الشبكات العصبية التلافيفية (CNNs). يمكن أن يوفر نموذج ACNN المقترح مع التجميع التكيفي تحسينات قابلة للقياس في دقة التصنيف، والقوة، والكفاءة الحاسوبية، مما قد يعود بالفائدة على تطبيقات العالم الحقيقي مثل التصوير الطبي،

والمركبات ذاتية القيادة، وأنظمة الأمن والأمان. تم تقييم فعالية النموذج المقترح خلال هذه الدراسة عن طريق إستخدام تجارب واسعة على مجموعات البيانات المختارة، مما يوفر رؤيا حول تأثيرها المحتمل على تنفيذ الشبكات العصبية التلافيفية ويقدم إرشادات للتطوير المستقبلي في استراتيجيات التجميع التكيفي لتحسين قدرات التعرف على الصور. علاوة على ذلك، تظهر نتائج التجميع التكيفي أن النتيجة النهائية للدقة، ودرجةF1 ، و نسبة التمثيل، والاسترجاع، و MSEهي 0.70، 0.70، 0.64، 0.75، و0.054 على التوالي. بالإضافة إلى ذلك، فإن نتائج التجميع التكيفي أفضل مقارنة مع التقنيات التقليدية بنسبة حوالي 62%، و70%، و89% مثل التجميع الأقصى، والتجميع المتوسط، وتجمع التردد.

الكلمات المفتاحية: الشبكات العصبية التلافيفية (CNNs)، التجميع التكيفي، المعالجة القائمة على القناع، تصنيف الصور، التنفيذ القائم على Python.