

نظام آلي للتنميط النووي للكروموسومات للتعرف على حالات التشوهات الخلقية الأكثر انتشاراً باستخدام التعلم العميق

منى سالم زين الله الخراز

إشراف

د. لمياء الرفاعي و د. مي فاضل

المستخلص

يعد تحليل الكروموسومات أحد أهم المهام في علم الوراثة الخلوية، حيث يمكّن علماء الوراثة الخلوية من تشخيص الاضطرابات الصحية والخلقية. النمط النووي هو تقنية أساسية في تحليل الكروموسومات حيث يصنف الصورة الطورية إلى ٢٤ صنف من الكروموسومات. هناك نوعان رئيسيان من أنواع اضطراب الكروموسومات: اضطراب في عدد الكروموسومات من خلال إما وجود نسخة إضافية من أحد الكروموسومات أو فقدان كروموسوم من أحد الأزواج، والنوع الآخر اضطراب في هيكل الكروموسوم حيث يكون هناك تغيير في بنية الكروموسوم. النمط النووي اليدوي معقد حيث يتطلب خبرة عالية في المجال ويستهلك الكثير من الوقت.

في هذا البحث قمنا بإنشاء نظام آلي باستخدام التعلم العميق للتعرف على التشوهات الخلقية الأكثر انتشاراً الناتجة عن زيادة أو نقص في عدد الكروموسومات على مجموعة من الصور الطورية الغير متداخلة من مركز التميز البحثي في علوم الجينوم الطبي بجامعة الملك عبد العزيز والمحتوية على ١٤٧ صورة طورية. في هذا النظام تمر الصور الطورية بثلاثة مراحل. المرحلة الأولى هي تجزئة الكروموسوم عن طريق الشبكة العصبية التلافيفية YOLO v2 متنوعة ببعض المعالجة اللاحقة للكروموسومات. حققت هذه المرحلة دقة تجزئة ١٠٠٪ ونتيجة تقاطع على الاتحاد ٠,٨٤ ومتوسط دقة ٠,٩٩٢٣. المرحلة الثانية هي استخراج الميزات والتصنيف حيث قمنا بضبط شبكة VGG19 باستخدام منهجين مختلفين. في المنهج الأول تمت إضافة طبقة/طبقات fully connected إضافية وفي المنهج الثاني تم استبدال طبقات fully connected بطبقة global average pooling. حققت هذه المرحلة نتيجة دقة ٩٥,٠٤٪. المرحلة الأخيرة هي اكتشاف التشوهات وحصلت هذه المرحلة على دقة اكتشاف ٩٦,٦٧٪.

لمزيد من التحقق من صحة طريقة التصنيف المقترحة، قمنا بتصنيف مجموعة بيانات مختبر التصوير الطبي الحيوي المتوفرة على الإنترنت وحققت دقة ٩٤,١١٪.

Automated System for Chromosome Karyotyping to Recognize the Most Common Numerical Abnormalities using Deep Learning

Mona Salem Zainallah Al-Kharraz

**Supervised By
Dr. Lamiaa Elrefaei & Dr. Mai Fadel**

ABSTRACT

Chromosome analysis is an essential task in a cytogenetics lab, where cytogeneticists can diagnose whether there are abnormalities or not. Karyotyping is a standard technique in chromosome analysis that classifies metaphase image to 24 chromosome classes. The main two categories of chromosome abnormalities are structural abnormalities that are changing in the structure of chromosomes and numerical abnormalities which include different types like monosomy (missing one chromosome) or trisomy (extra copy of the chromosome). Manual karyotyping is complex and requires high domain expertise, as it takes an amount of time. With these motivations, in this research, we automate karyotyping by borrowing the latest ideas from deep learning and recognize the most common numerical abnormalities (monosomy and trisomy) on a dataset contains 147 non-overlapped metaphase images collected from Center of Excellence in Genomic Medicine Research at King Abdulaziz University. The metaphase images go through three stages in our study. The first one is individual chromosomes detection using YOLOv2 Convolutional Neural Network followed by some chromosome post-processing. This step achieved 0.84 mean IoU, 0.9923 AP, and 100% individual chromosomes detection accuracy. The second stage is feature extraction and classification where we fine-tune VGG19 network using two different approaches, one by adding extra fully connected layer(s) and another by replacing fully connected layers with the global average pooling layer. The best accuracy obtained is 95.04%. The final step is detecting abnormality and this step obtained 96.67% average abnormality detection accuracy. To further validate the proposed classification method, we examined the Biomedical Imaging Laboratory dataset which is publicly available online and achieved 94.11% accuracy.