جامعة دمشق
كلية الهندسة الميكانيكية والكهربائية
قسم هندسة الالكترونيات والاتصالات

استرجاع المعلومات الموسيقية باستخدام معالجة الإشارة الرقمية والتعلم العميق

Musical Information Retrieval Using Digital Signal Processing and Deep Learning

ابراهيم صالح وهبي

د. حسين طياوي بحبوح

الملخص

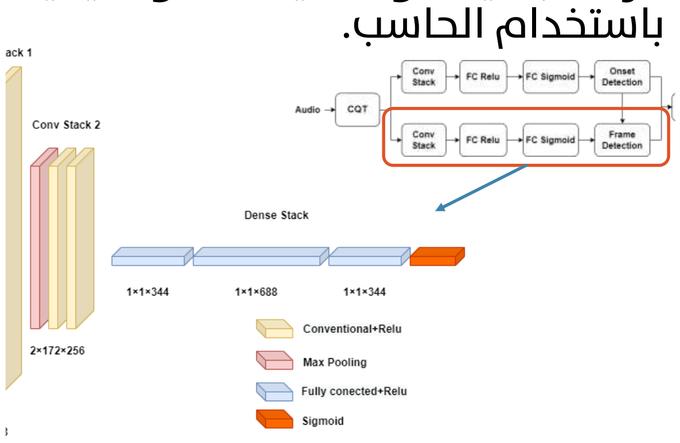
قدم البحث نموذج كامل لتتبع النغمات المتعددة للموسيقا العربية والغربية على حد سواء، والذي يتميز بقدرته على اكتشاف بدايات النغمات وامتدادها لمقطوعات البيانو الموسيقية وتحويلها إلى تدوين موسيقي. يجمع هذا النموذج كل من تقانات معالجة الإشارة الرقمية والتعلم العميق لزيادة دقة النتائج. بداية يُطبّق تحويل عامل الجودة الثابت على ملف الدخل الصوتي، وتمرير نتيجة التحويل للمرحلة الثانية التي تحتوي على شبكتين عصبونيتين التفافيتين. صممت الشبكة الأولى لاكتشاف بداية النغمة الموسيقية، في حين صممت الشبكة الثانية لاكتشاف امتداد هذه النغمة.

القسم العملي

يتم توليد عينات التدريب للموسيقى الشرقية بناء على عينات تدريب الموسيقى الغربية المتوافرة وذلك بتبديل بعض العلامات الموسيقية بعلامات أخرى (أرباع علامات) ثم تتم عملية المعالجة المسبقة لعينات التدريب وذلك بإنشاء ثلاثة ملفات؛ الأول صوتي بصيغة WAV والثاني يمثل زمن بداية كل العلامات الموسيقية والثالث يمثل امتداد العلامات الموسيقية على محور الزمن . يُطبّق تحويل عامل الجودة الثابت CQT على عينات التدريب ذات صيغة WAV وتعامل معاملة الدخل الذي نود استخراج العلامات الموسيقية منه بينما يمثل ملفا امتداد النغمات وبداية النغمات العلامات الموسيقية الصحيحة الموافقة لملف WAV آنف الذكر.

القسم العملي

تقوم الشبكة العصبونية بمقارنة النتائج المستخرجة من ملف WAV في كل دورة تدريب مع البيانات الموجودة في ملفي بداية النغمات وامتداد النغمات الموسيقية ويتم تغيير الأوزان من قبل الشبكة العصبونية وفق تابع تنشيط مناسب وذلك لتحسين النتائج قدر الإمكان. بعد انتهاء التدريب تصبح الشبكة العصبونية قادرة على اكتشاف كل من بداية النغمة الموسيقية وامتداد العلامة الموسيقية وتقوم بتحويلهم إلى ملف MIDI والذي يمكن تحويله إلى نوطة موسيقية أو عزفه بأي صوت لأي آلة موسيقية باستخدام الحاسب.



القسم النظري

تعد عملية استرجاع المعلومات الموسيقية عملية معقدة ولا سيّما في المقطوعات الموسيقية متعددة النغمات، فضلاً عن ذلك فإنّ معظم الدراسات السابقة لم تأخذ خطأ ربع العلامة الموسيقية في الحسبان علماً بأنّ ربع العلامة الموسيقية تعد السمة الأهم في الموسيقا العربية . يهدف هذا البحث إلى بناء نموذج كامل لتتبع النغمات المتعددة للموسيقا العربية والغربية على حد سواء، والذي يتميز بقدرته على اكتشاف بدايات النغمات وامتدادها لمقطوعات البيانو الموسيقية ويجمع هذا النموذج كل من تحويل عامل الجودة الثابت CQT والشبكات العصبونية الالتفافية CNN لزيادة دقة النتائج.

النتائج والمناقشة

قمنا في هذا البحث بالتأكد من قدرة الشبكات العصبونية الالتفافية على استرجاع المعلومات الموسيقية والكتابة الآلية للموسيقا في الموسيقا الشرقية. وجدنا أن النموذج المصمم حسّن من الأداء على كل من صعيدي اكتشاف بداية النغمة الموسيقية وامتداد الإطار، وذلك عندما يكون خطأ ربع العلامة محسوباً. حصلنا على درجة F1 تبلغ قيمتها 75.28% لبداية النغمة الموسيقية و 47.37% للإطار كاملا.

المراجع

- [1] Elowsson, A. (2020). Polyphonic pitch tracking with deep layered learning. The Journal of the Acoustical Society of America, 148(1), 446-468.
- [2] Alan, M. (2019), Middle Eastern Violin Method A Method for Teaching and Transcribing Middle Eastern Music. Institutionen för Folkmusik.
- [3] Lee, J.-H., Kang, J., Shim, W., Chung, H.-S., & Sung, T.-E. (2020). Pattern Detection Model Using a Deep Learning Algorithm for Power Data Analysis in Abnormal Conditions.
- [4] Meng, Z., & Chen, W. (2020). Automatic music transcription based on convolutional neural network, constant Q transform and MFCC. Journal of Physics.
- [5]] Kelz, R., Matthias Dorfer, Korzeniowski, F., Böck, S., Arzt, A., & Widmer, G. (2016). On the Potential of Simple Framewise Approaches to Piano Transcription. ArXiv.org.
- [6] Hawthorne, C. (n.d.). Onsets and Frames: Dual-Objective Piano Transcription. arXiv preprint arXiv:1710.11153, 2017