

UNIVERSIDAD DE LOS ANDES
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA

PRESENTACIÓN DE PROPUESTA DE PROYECTO DE GRADO

SEMESTRE: 2021-2
FECHA: 21 de junio de 2021

PROYECTO O TESIS DE GRADO PARA OPTAR EL TÍTULO DE:
Ingeniero electrónico

ESTUDIANTE: Christian David Florez Pinillos **CÓDIGO:** 201423499

TÍTULO DE LA TESIS O PROYECTO:
Automatic music transcription, aplicaciones a melodías de piano

DECLARACIÓN:

Soy consciente que cualquier tipo de fraude en esta Tesis es considerado como una falta grave en la Universidad. Al firmar, entregar y presentar esta propuesta de Tesis o Proyecto de Grado, doy expreso testimonio de que esta propuesta fue desarrollada de acuerdo con las normas establecidas por la Universidad. Del mismo modo, aseguro que no participé en ningún tipo de fraude y que en el trabajo se expresan debidamente los conceptos o ideas que son tomadas de otras fuentes.

Soy consciente de que el trabajo que realizaré incluirá ideas y conceptos del autor y el Asesor y podrá incluir material de cursos o trabajos anteriores realizados en la Universidad y por lo tanto, daré el crédito correspondiente y utilizaré este material de acuerdo con las normas de derechos de autor. Así mismo, no haré publicaciones, informes, artículos o presentaciones en congresos, seminarios o conferencias sin la revisión o autorización expresa del Asesor, quien representará en este caso a la Universidad.

Firma (Estudiante)

Código: 201423499
CC: 1020828212

Vo.Bo. ASESOR (Firma)
Nombre: Fernando Lozano

Vo.Bo. COASESOR (Firma)
Nombre: Carolina Higuera

1. Caracterización del problema (Justificación)

El desarrollo de la notación musical es una herramienta muy importante en la música. Permite a los compositores transmitir a los músicos el contenido de su creación, sin necesidad de que estos tengan que oír la pieza, algo que resultaba particularmente importante en tiempos pasados en donde los medios de comunicación eran mucho más limitados.

Adicionalmente, está el proceso de transcripción musical, en la cual una persona oye una melodía y es capaz de plasmarla en notación musical. Esta tarea resulta de gran complejidad, pues no puede ser desarrollada por cualquier persona, debe ser alguien con alto conocimiento musical y con un oído muy educado. Sumado a su complejidad está que requiere una gran cantidad de tiempo para llevarse a cabo.

En consecuencia, dados todos los nuevos adelantos tecnológicos, resulta natural buscar una solución computacional a este problema. En este marco el problema se conoce como *transcripción automática de música* (AMT por su sigla en inglés, *Automatic Music Transcription*), en el cual se busca desarrollar algoritmos que transformen una señal de audio de música en alguna clase de notación musical (generalmente una partitura o un archivo MIDI). Hoy en día esta es un área activa de investigación, por lo que el problema aún no ha sido solucionado.

La resolución del mismo tendría impacto en múltiples áreas como musicología, enseñanza musical, creación musical, búsqueda de música, etc. Más aún, este problema se encuentra enmarcado en área más general de investigación conocida como *recuperación de información musical*, y la solución del mismo impactaría otras áreas de investigación asociadas a la recuperación de información musical como la generación automática de música, detección de covers, entre otras.

2. Marco teórico

Varias aproximaciones han sido tomadas para abordar el problema de transcripción automática de música, entre ellas están: procesamiento de señales, modelamiento probabilístico, factorización en matrices no negativas, redes neuronales, entre otras. Las dos principales aproximaciones y en las que se han obtenido resultados más prominentes son las basadas en redes neuronales y en factorización en matrices no negativas [BDDE19].

Por otro lado, la solución de este problema conlleva múltiples desafíos, entre los cuales se destacan:

- Polifonía: es la presencia de varias voces musicales en la melodía, estas pueden estar interpretadas por el mismo instrumento (por ejemplo la mano izquierda y derecha en un piano llevan melodías diferentes) o directamente por instrumentos diferentes. La separación de las diferentes fuentes, así como sobrellevar la interferencia que se presenta entre estas resulta ser un desafío importante.

- Superposición de armónicos: los sonidos musicales consisten de una frecuencia fundamental acompañados por diversos armónicos, sonidos a frecuencias que son múltiplos enteros de la frecuencia fundamental. Una consecuencia de este comportamiento es que dos frecuencias base diferentes pueden compartir armónicos o una nota base estar a la misma frecuencia que un armónico de otra, dificultando así la identificación de las notas y su intensidad.
- Dependencia de eventos: en problemas similares donde se tienen diversas fuentes sonoras se asume independencia de los sonidos de estas (por ejemplo, en el *cocktail party problem*), sin embargo esto no se puede asumir en el contexto musical debido a la fuerte estructura en la que están enmarcado lo diversos sonidos presentes.
- Dificultad en la generación de datasets: como se mencionó previamente, actualmente la transcripción de música es una tarea larga y que requiere de un oído experto, razón por la cual construir grandes datasets anotados para aprendizaje supervisado es una tarea muy costosa, difícil y laboriosa.

Debido a estas dificultades se decidió concentrarse en el caso de un instrumento, concretamente el piano por ser uno de los instrumentos más completos y porque para este instrumento existe una buena cantidad de datos públicamente disponibles.

3. Caracterización del proyecto

3.1. Objetivo general

Implementar modelos basados tanto en redes neuronales como en descomposición en matrices no negativas de transcripción automática de música, entrenarlos y probarlos. También, se quiere evaluar cómo se comportan los modelos con grabaciones de menor calidad de audio a las del dataset utilizado para entrenar el modelo.

3.2. Objetivos específicos

- Implementar un modelo basado en redes neuronales para hacer transcripción automática de música.
- Implementar un modelo basado en descomposición en matrices no negativas para hacer transcripción automática de música.
- Entrenar y evaluar los modelos implementados.
- Estudiar el comportamiento de los modelos en grabaciones de baja calidad.
- Comparar y documentar los resultados obtenidos.

3.3. Alcance

Se busca entregar un documento donde se expone el diseño y entrenamiento de los modelos implementados de transcripción automática de música, también se incluirán y analizarán los resultados obtenidos en pruebas que se realizarán sobre estos con los audios del dataset y los de baja calidad.

4. Contexto del proyecto y tratamientos

4.1. Restricciones

- El proyecto debe ser terminado durante el periodo 2021-1.

4.2. Factores de riesgo

- Problemas al tratar de replicar las implementaciones de la literatura por falta de detalles en los artículos.
- Recursos computacionales insuficientes para manejar los datos y entrenar los modelos en el tiempo requerido.

5. Cronograma

5.1. Identificación de los hitos

- Revisión bibliográfica y del estado de arte de la transcripción automática de música.
- Selección de modelos a implementar.
- Selección de la tecnología a utilizar para la implementación y entrenamiento de los modelos.
- Implementación de los modelos.
- Entrenamiento de los modelos.
- Evaluación de los modelos.
- Pruebas de los modelos en audios de calidad reducida.
- Documentación del desarrollo y los resultados obtenidos.
- Entrega final de el proyecto.

5.2. Diagrama de Grant

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16
Revisión bibliográfica y del estado de arte																
Selección de modelos a implementar																
Selección de la tecnología a utilizar																
Implementación de los modelos																
Entrenamiento de los modelos																
Evaluación de los modelos																
Pruebas de los modelos en audios de calidad reducida.																
Documentación del desarrollo y los resultados obtenidos																
Entrega final de el proyecto																

6. Recursos

6.1. Software

El proyecto se llevará a cabo en Python. Posiblemente se acuda adicionalmente a software gratuito para manejo de formatos de audio y grabación.

Referencias

- [BDDE19] E. Benetos, S. Dixon, Z. Duan, and S. Ewert. Automatic music transcription: An overview. *IEEE Signal Processing Magazine*, 36(1):20–30, 2019.
- [HES⁺17] Curtis Hawthorne, Erich Elsen, Jialin Song, Adam Roberts, Ian Simon, Colin Raffel, Jesse H. Engel, Sageev Oore, and Douglas Eck. Onsets and frames: Dual-objective piano transcription. *CoRR*, abs/1710.11153, 2017.
- [HSR⁺19] Curtis Hawthorne, Andriy Stasyuk, Adam Roberts, Ian Simon, Cheng-Zhi Anna Huang, Sander Dieleman, Erich Elsen, Jesse Engel, and Douglas Eck. Enabling factorized piano music modeling and generation with the maestro dataset, 2019.
- [OP17] C. O’Brien and M. D. Plumbley. Automatic music transcription using low rank non-negative matrix decomposition. In *2017 25th*

European Signal Processing Conference (EUSIPCO), pages 1848–1852, 2017.



**TRABAJO DE GRADO
AUTORIZACIÓN DE SU USO A FAVOR DE LA
UNIVERSIDAD DE LOS ANDES**

Yo , mayor de edad, vecino de Bogotá D.C., identificado con la Cédula de Ciudadanía N° de , actuando en nombre propio, en mi calidad de autor del trabajo de tesis, monografía o trabajo de grado denominado:

haré entrega del ejemplar respectivo y de sus anexos del ser el caso, en formato digital o electrónico (CD-ROM) y autorizo a LA UNIVERSIDAD DE LOS ANDES, para que en los términos establecidos en la Ley 23 de 1982, Ley 44 de 1993, Decisión Andina 351 de 1993, Decreto 460 de 1995 y demás normas generales sobre la materia, utilice y use en todas sus formas, los derechos patrimoniales de reproducción, comunicación pública, transformación y distribución (alquiler, préstamo público e importación) que me corresponden como creador de la obra objeto del documento. PARÁGRAFO: La presente autorización se hace extensiva no sólo a las facultades y derechos de uso sobre la obra en formato o soporte material, sino también para formato virtual, electrónico, digital, óptico, usos en red, internet, extranet, intranet, etc., y en general para cualquier formato conocido o por conocer.

EL AUTOR - ESTUDIANTES, manifiesta que la obra objeto de la presente autorización es original y la realizará sin violar o usurpar derechos de autor de terceros, por lo tanto la obra es de su exclusiva autoría y tiene la titularidad sobre la misma. PARÁGRAFO: En caso de presentarse cualquier reclamación o acción por parte de un tercero en cuanto a los derechos de autor sobre la obra en cuestión, EL ESTUDIANTE - AUTOR, asumirá toda la responsabilidad, y saldrá en defensa de los derechos aquí autorizados; para todos los efectos la Universidad actúa como un tercero de buena fe.

EL AUTOR - ESTUDIANTE.

(Firma)

Nombre

C.C. N° de