



Búsqueda Avanzada

Nuevos resultados

[Sigue esta preimpresión](#)

CityNet - Herramientas de aprendizaje profundo para la evaluación ecoacústica urbana

AJ Fairbrass, M. Firman, C. Williams, GJ Brostow, H. Titheridge, KE Jones

doi: <https://doi.org/10.1101/248708>

Ahora publicado en *Métodos en Ecología y Evolución* doi: [10.1111/2041-210x.13114](https://doi.org/10.1111/2041-210x.13114)



Abstracto

Texto completo

Información/Historial

Métrica

Vista previa de PDF

RESUMEN

1. Las ciudades sustentan comunidades ecológicas únicas y valiosas, pero la comprensión de la vida silvestre urbana es limitada debido a las dificultades para evaluar la biodiversidad. Los estudios ecoacústicos son una forma útil de evaluar hábitats, donde el sonido biótico medido a partir de grabaciones de audio se utiliza como indicador de la biodiversidad. Sin embargo, se ha demostrado que los algoritmos existentes para medir el sonido biótico están sesgados por sonidos no bióticos en las grabaciones, típicos de entornos urbanos.

2. Desarrollamos CityNet, un sistema de aprendizaje profundo que utiliza redes neuronales convolucionales (CNN), para medir la actividad acústica audible biótica (CityBioNet) y antropogénica (CityAnthroNet) en las ciudades. Las CNN se entrenaron en un gran conjunto de datos de grabaciones de audio anotadas recopiladas en el Gran Londres, Reino Unido. Utilizando un conjunto de datos de prueba disponible, comparamos la precisión y la recuperación de CityBioNet y CityAnthroNet por separado con los mejores algoritmos alternativos disponibles: cuatro índices acústicos (IA): índice de complejidad acústica, índice de diversidad acústica, índice bioacústico e índice de paisaje sonoro de diferencia normalizada, y una CNN (bulbul) de detección de cantos de pájaros de última generación. También comparamos el efecto de los sonidos no bióticos en las predicciones de CityBioNet y bulbul. Finalmente aplicamos CityNet para describir patrones acústicos del paisaje sonoro urbano en dos sitios a lo largo de un gradiente de urbanización.

3. CityBioNet fue el algoritmo con mejor rendimiento para medir la actividad biótica en términos de precisión y recuperación, seguido de bulbul, mientras que las IA obtuvieron el peor rendimiento. CityAnthroNet superó el índice de paisaje sonoro de diferencia normalizada, pero por un margen menor que el logrado por CityBioNet frente a los algoritmos de la competencia. Las predicciones de CityBioNet se vieron afectadas por sonidos mecánicos, mientras que el tráfico aéreo y los sonidos del viento influyeron en las predicciones de bulbul. A lo largo de un gradiente de urbanización, mostramos que CityNet produjo patrones diarios realistas de actividad acústica biótica y antropogénica a partir de datos de audio urbanos del mundo real.
4. Con CityNet, es posible medir automáticamente la actividad acústica biótica y antropogénica en las ciudades a partir de grabaciones de audio. Si se integra dentro de un sistema de detección autónomo, CityNet podría producir datos ambientales para ciudades a gran escala y facilitar la investigación de los impactos de las actividades antropogénicas en la vida silvestre. Los algoritmos, el código y los modelos previamente entrenados están disponibles gratuitamente en combinación con dos conjuntos de datos de audio urbano anotados por expertos para facilitar la vigilancia ambiental automatizada en las ciudades.

INTRODUCCIÓN

Más de la mitad de la población humana mundial vive ahora en ciudades (**UN-DESA 2016**) y la biodiversidad urbana puede proporcionar a las personas una multitud de beneficios para la salud y el bienestar, incluida una mejor salud física y psicológica (**Natural England 2016** ; **Crouse *et al.* 2017**). Las ciudades pueden sustentar una alta biodiversidad, incluidas especies endémicas nativas (**Aronson *et al.* 2014**), y actuar como refugios para la biodiversidad que ya no puede persistir en paisajes agrícolas intensamente gestionados que rodean las ciudades (**Hall *et al.* 2016**). Sin embargo, nuestra comprensión de la biodiversidad urbana sigue siendo limitada (**Faeth, Bang & Saari 2011** ; **Beninde, Veith & Hochkirch 2015**). Una de las razones de esto son las dificultades asociadas con la evaluación de la biodiversidad, como el acceso repetido a los sitios de estudio y la intensidad de recursos de los métodos tradicionales (**Farinha-Marques *et al.* 2011**). Esto inhibe nuestra capacidad para realizar la evaluación a gran escala necesaria para comprender los ecosistemas urbanos.

Los estudios ecoacústicos han surgido como un método útil de cuantificación a gran escala de las comunidades ecológicas y sus hábitats (**Sueur y Farina 2015**). Los equipos de grabación acústica pasiva facilitan la recopilación de datos de audio durante largos períodos de tiempo y grandes escalas espaciales con menos recursos que los métodos de encuesta tradicionales (**Digby *et al.* 2013**). Se han desarrollado varios métodos automatizados para medir el sonido biótico en los grandes volúmenes de datos acústicos que normalmente se producen mediante estudios ecoacústicos (**Sueur & Farina 2015**). Por ejemplo, los índices acústicos (IA) utilizan las características espectrales y temporales de la energía acústica en grabaciones de sonido para producir medidas de sonido biótico y antropogénico de toda la comunidad (**Sueur *et al.* 2014**). Sin embargo, se ha demostrado que varias IA de uso común están sesgadas por sonidos no bióticos (**Towsey *et al.* 2014** ; **Fuller *et al.* 2015** ; **Gasc *et al.* 2015a**), y

no son adecuadas para su uso en el entorno urbano sin la autorización previa. eliminación de ciertos sonidos no bióticos de las grabaciones (**Fairbrass et al. 2017**).

El aprendizaje automático (ML) se aplica cada vez más a la evaluación y el monitoreo de la biodiversidad porque facilita la detección y clasificación de señales ecoacústicas en datos de audio (**Acevedo et al. 2009** ; **Walters et al. 2012** ; **Stowell & Plumbley 2014**). Utilizando conjuntos de datos de audio anotados de especies soníferas, se puede entrenar un modelo de ML para que reconozca sonidos bióticos en función de múltiples características o características acústicas, y para asociar estas características con clasificaciones taxonómicas, y luego puede asignar una clasificación probabilística a los sonidos dentro de las grabaciones. Las IA solo utilizan un número limitado de características acústicas en sus cálculos, como la entropía espectral dentro de bandas de frecuencia definidas (**Boelman et al. 2007** ; **Villanueva-Rivera et al. 2011** ; **Kasten et al. 2012**) o cambios de entropía a lo largo del tiempo (**Pieretti, Farina y Morri 2011**). Además, la relación entre las características y los resultados del algoritmo la elige un ser humano, en lugar de aprenderla automáticamente a partir de un conjunto de datos anotado. Por el contrario, los algoritmos de ML pueden utilizar muchas más funciones en sus cálculos y la relación entre entradas y salidas se determina automáticamente en función de los datos de entrenamiento anotados proporcionados. Las redes neuronales convolucionales, CNN (o aprendizaje profundo) (**LeCun, Bengio & Hinton 2015**) pueden incluso elegir, en función de las anotaciones en el conjunto de datos de entrenamiento, las características que mejor discriminan diferentes clases en los conjuntos de datos sin especificarlas a priori, y pueden aprovecharlas. de grandes cantidades de datos de entrenamiento donde su capacidad para superar los algoritmos definidos por humanos aumenta a medida que hay más datos etiquetados disponibles.

Se han desarrollado algoritmos de aprendizaje automático específicos de cada especie para identificar automáticamente los sonidos emitidos por una variedad de organismos soníferos, incluidos pájaros (**Stowell y Plumbley 2014**), murciélagos (**Walters et al. 2012** ; **Zamora-Gutiérrez et al. 2016**), anfibios (**Acevedo et al. 2009**) e invertebrados (**Chesmore y Ohya 2004**). Sin embargo, estos algoritmos se centran en una pequeña cantidad de especies, lo que limita su utilidad para tareas de clasificación amplias en comunidades. Más recientemente, se están desarrollando algoritmos que detectan grupos taxonómicos completos, por ejemplo, sonidos de pájaros en grabaciones de audio del Reino Unido y la zona de exclusión de Chernobyl (**Grill & Schlüter 2017**), pero estos algoritmos aún no se han probado con datos de audio ruidosos de entornos urbanos. Actualmente no existen algoritmos que produzcan medidas de sonido biótico para toda la comunidad que se sepa que sean adecuados para su uso en entornos urbanos acústicamente complejos.

Aquí, desarrollamos el sistema de análisis acústico CityNet, que utiliza dos CNN para medir la actividad acústica audible (0-12 kHz) biótica (CityBioNet) y antropogénica (CityAnthroNet) en grabaciones de audio de entornos urbanos. Utilizamos este rango de frecuencia ya que contiene la mayoría de los sonidos emitidos por especies soníferas audibles en el entorno urbano (**Fairbrass et al. 2017**). Las

CNN se entrenaron utilizando CitySounds2017, un conjunto de datos de sonidos urbanos anotados por expertos recopilados en el Gran Londres, Reino Unido, que desarrollamos aquí. Comparamos el rendimiento de CityNet utilizando un conjunto de datos reservado comparando la precisión y la recuperación de los algoritmos con cuatro IA de uso común: Índice de Complejidad Acústica (ACI) (**Pieretti, Farina y Morri 2011**), Índice de Diversidad Acústica (ADI) (**Villanueva- Rivera *et al.* 2011**), índice bioacústico (BI) (**Boelman *et al.* 2007**), índice de paisaje sonoro de diferencia normalizada (NDSI) (**Kasten *et al.* 2012**) y bulbul, un algoritmo de última generación para detectar sonidos de aves para resumir la actividad acústica de las aves (**Grill & Schlüter 2017**). Como el enfoque principal del estudio fue el desarrollo de algoritmos para la evaluación ecoacústica de la biodiversidad en las ciudades, llevamos a cabo un análisis adicional de los dos algoritmos de mejor rendimiento para medir el sonido biótico, CityBioNet y bulbul, investigando el efecto de los sonidos no bióticos en la exactitud de los algoritmos. Finalmente, aplicamos CityNet para investigar patrones diarios de sonido biótico y antropogénico en el paisaje sonoro urbano.

MATERIALES Y MÉTODOS

Desarrollamos dos modelos CNN, CityBioNet y CityAnthroNet dentro del sistema CityNet para generar medidas de sonido biótico y antropogénico, respectivamente. El proceso de CityNet (**Figura 1**) constaba de 7 pasos principales, de la siguiente manera:

1. *Grabar audio*: Las grabaciones de audio .wav de frecuencia audible (0-12 kHz) se realizaron utilizando una grabadora acústica pasiva.
2. *Conversión de audio a espectrograma Mel*: cada archivo de audio se convirtió automáticamente a una representación de espectrograma Mel con 32 contenedores de frecuencia, representados como filas en el espectrograma, utilizando una resolución temporal de 21 columnas por segundo de audio sin procesar. Antes de su uso en el clasificador, cada espectrograma S se convirtió a una representación en escala logarítmica, utilizando la fórmula $\log(A + B * S)$. Para la detección de sonido biótico se utilizaron los parámetros $A = 0,001$ y $B = 10,0$, mientras que para la detección de sonido antropogénico se utilizaron los parámetros $A = 0,025$ y $B = 2,0$.
3. *Extraer ventana del espectrograma*: una única entrada a la CNN comprendía un fragmento corto de espectrograma W_s , de 21 columnas de ancho, que representaba 1 segundo de audio.
4. *Aplicar diferentes estrategias de normalización*: existen muchos métodos diferentes para preprocesar espectrogramas antes de usarlos en ML; por ejemplo, blanqueamiento (**Lee *et al.* 2009**) y resta de valores medios a lo largo de cada intervalo de frecuencia (**Aide *et al.* 2013**). Las CNN pueden aceptar entradas con múltiples canales de datos, por ejemplo, los canales rojo, verde y azul de una imagen en color. Explotamos la capacidad de múltiples canales de entrada de nuestra CNN al proporcionar como entrada cuatro espectrogramas, cada uno de ellos preprocesado utilizando una estrategia de normalización diferente (ver Métodos complementarios), lo que brindó

mejoras considerables a la precisión de la red por encima de cualquier esquema de normalización único de forma aislada. Luego de aplicar diferentes estrategias de normalización, la entrada a la red consistió en un tensor de $32 \times 21 \times 4$.

5. *Aplicar el clasificador CNN:* como se describió anteriormente, la clasificación se realizó con una CNN, cuyos parámetros se aprendieron a partir de los datos de entrenamiento. La CNN constaba de una serie de capas, cada una de las cuales modificaba sus datos de entrada con operaciones matemáticas parametrizadas que se optimizaron para mejorar el rendimiento de la clasificación durante el entrenamiento (consulte Métodos complementarios para obtener más detalles). La capa final produjo la predicción de la presencia o ausencia de sonido biótico o antropogénico.
6. *Haga predicciones para cada momento en el tiempo:* en el momento de la prueba, los pasos (3) a (5) se repitieron cada segundo a lo largo del archivo de audio, para dar una medida de la actividad biótica o antropogénica a lo largo del tiempo. Las predicciones para cada fragmento de audio se hicieron de forma independiente.
7. *Resumir:* cuando correspondía, las predicciones a nivel de fragmentos se resumieron para obtener información sobre las tendencias en el tiempo y el espacio. Por ejemplo, los niveles de actividad previstos para cada ventana de media hora podrían promediarse para inspeccionar el nivel de actividad biótica y antropogénica en diferentes momentos del día.

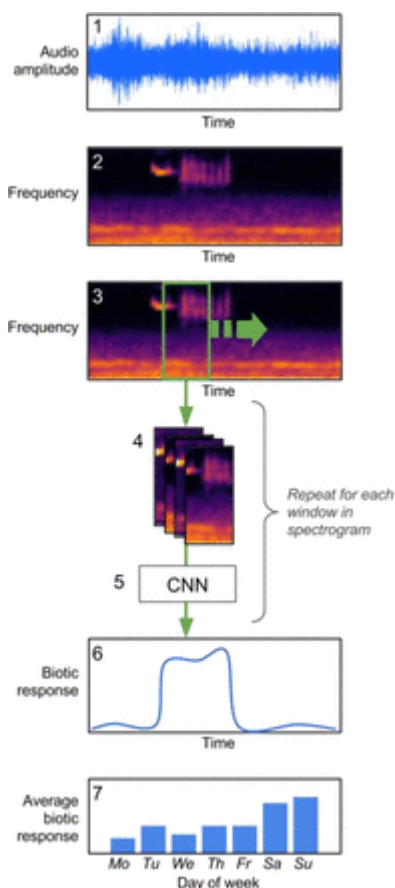


Figura 1.

El canal de análisis CityNet para medir la actividad acústica biótica y antropogénica. El audio sin procesar (1), grabado en el campo, se convierte en una representación de espectrograma (2). Se ejecuta una ventana deslizante a lo largo de la dimensión temporal y se extrae una ventana del espectrograma en cada paso (3). Esta ventana de espectrograma se preprocesa con cuatro estrategias de normalización diferentes y los resultados se concatenan. Esta pila de espectrogramas pasa a través de una CNN (5), que fue entrenada en *el tren* CitySounds2017. La CNN proporciona, en cada paso de tiempo de 1 segundo, una predicción de la presencia/ausencia de actividad acústica biótica o antropogénica (6). Finalmente, estas medidas por paso de tiempo se pueden agregar para brindar resúmenes en el tiempo o el espacio (7).

La canalización de ML se escribió en Python v.2.7.12 (**Python Software Foundation 2016**) usando Theano v.0.9.0 (**The Theano Development Team *et al.* 2016**) y Lasagne v.0.2 (**Dieleman *et al.* 2015**) para ML y librosa v.0.4.2 (**McFee *et al.* 2015**) para procesamiento de audio.

Conjunto de datos acústicos

Seleccionamos 63 sitios de infraestructura verde (IG) en el Gran Londres (Reino Unido) y sus alrededores para recopilar datos de audio para entrenar y probar los algoritmos de CityNet. Estos sitios representan una variedad de IG en el Gran Londres y sus alrededores en términos de tipo de IG, tamaño e intensidad urbana. Cada sitio fue muestreado sistemáticamente durante 7 días consecutivos durante los meses de mayo a octubre entre 2013 y 2015 (**Figura 2** , **Tabla S1**). En cada ubicación, se implementó un sensor de campo de audio digital Song Meter SM2+ (Wildlife Acoustics, Inc., Concord, Massachusetts, EE. UU.), que grabó sonido entre 0 y 12 kHz a una frecuencia de muestreo de 24 kHz. El sensor estaba equipado con un único micrófono omnidireccional (respuesta de frecuencia: -35 ± 4 dB) orientado horizontalmente a una altura de 1 m. Los archivos se guardaron en formato .wav en una tarjeta SD. El audio se grabó en fragmentos manejables computacionalmente de 29 minutos de cada 30

minutos (23,2 horas de grabación por día), que se dividieron en archivos de audio de 1 minuto usando Slice Audio File Splitter (NCH Software Inc. 2014), lo que generó un total de 613.872 minutos discretos de grabación de audio (9.744 minutos para cada uno de los 63 sitios). Esto constituyó el conjunto de datos CitySounds2017.

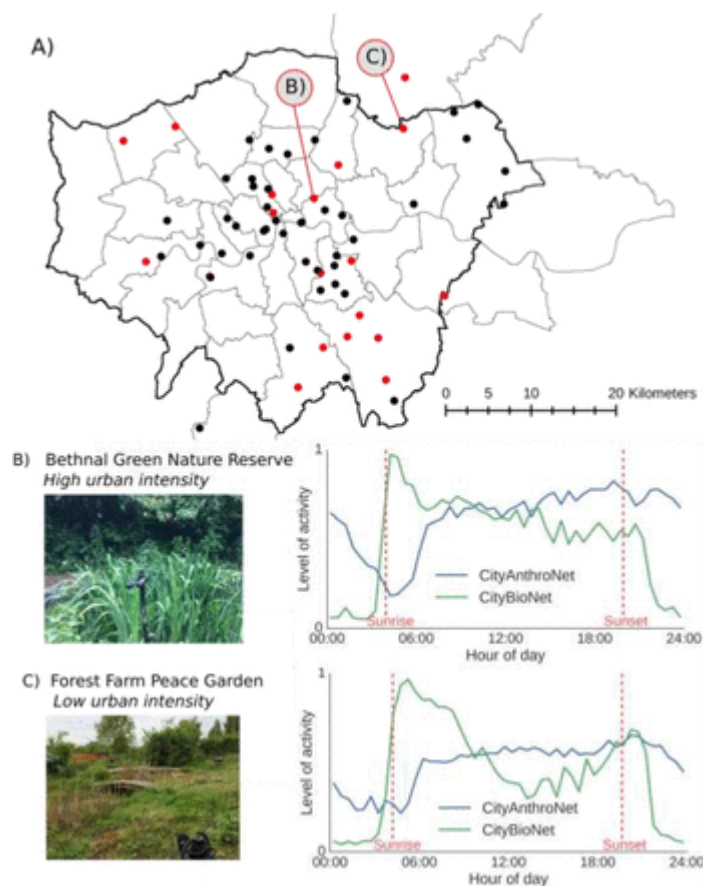


Figura 2.

Ubicación de los sitios de estudio y patrones acústicos diarios promedio en dos sitios a lo largo de un gradiente de urbanización. Los puntos en (A) representan las ubicaciones utilizadas para el conjunto de datos de entrenamiento, *el tren* CitySounds2017 (negro) y el conjunto de datos de prueba, *la prueba* CitySounds2017 (rojo). Aquí, CityNet se ejecutó durante los 7 días completos de grabación en dos sitios de alta (B) y baja (C) intensidad urbana para predecir la presencia/ausencia de sonido biótico y antropogénico en cada segundo de la semana utilizando un umbral de probabilidad de 0,5. El número previsto de segundos que contenían sonido biótico y antropogénico para cada período de media hora se promedió durante la semana para producir patrones diarios promedio de actividad acústica. El límite del Gran Londres se indica con una línea en negrita. Datos de límites del censo del Reino Unido (<http://www.ons.gov.uk/> , consultado el 11/04/2014).

Conjunto de datos de entrenamiento acústico

Para crear nuestro conjunto de datos de entrenamiento (*el tren* CitySounds2017), seleccionamos al azar veinticinco grabaciones de 1 minuto del 70% de los sitios de estudio (44 sitios, 1100 grabaciones). AF anotó manualmente los espectrogramas de cada grabación, calculados como la magnitud logarítmica de una transformada de Fourier discreta (tamaño de ventana de Hamming no superpuesta = 720 muestras = 10 ms), utilizando AudioTagger (disponible en <https://github.com/groakat/AudioTagger>). Los espectrogramas se anotaron localizando las bandas de tiempo y frecuencia de sonidos discretos dibujando cuadros delimitadores lo más ajustados visualmente posible dentro de los espectrogramas mostrados en un monitor LED Dell UltraSharp de 61 cm. Los tipos de sonido, como "invertebrado",

"lluvia" y "tráfico", se identificaron buscando patrones típicos en espectrogramas (**Figura S1**) y escuchando las muestras de audio representadas en las partes anotadas del espectrograma. Luego, las categorías de sonidos se agruparon en clases bióticas, antropogénicas y geofónicas siguiendo a **Pijanowski *et al.* (2011)** , donde definimos bióticos como sonidos generados por organismos bióticos no humanos, antropogénicos como sonidos asociados con actividades humanas y geofónicos como sonidos ambientales no biológicos, por ejemplo, viento y lluvia.

Conjunto de datos y evaluación de pruebas acústicas

Para evaluar el rendimiento de los algoritmos de CityNet, creamos un conjunto de datos de prueba (CitySounds2017 _{test}) seleccionando estratégicamente 40 grabaciones de CitySounds2017 del 30% restante de los sitios (19 sitios) que contenían una variedad de actividad acústica tanto biótica como antropogénica. La prueba CitySounds2017 se tomó como muestra de diferentes sitios de grabación para entrenar CitySounds2017 para demostrar que los algoritmos de CityNet se generalizan a los sonidos grabados en nuevas ubicaciones del sitio (**Figura 2** , **Tabla S1**). Para optimizar la calidad de las anotaciones en la prueba CitySounds2017 , seleccionamos cinco etiquetadores humanos para anotar por separado los sonidos dentro de las grabaciones de audio (usando los mismos métodos anteriores) para crear un único conjunto de datos de prueba anotado. Los conflictos se resolvieron mediante la regla de la mayoría, y en los casos en los que no había mayoría, utilizamos nuestro propio criterio sobre la clasificación más adecuada. Nuestros conjuntos de datos de prueba y capacitación anotados de CitySounds2017 están disponibles en <https://figshare.com/s/adab62c0591afaeafedd> .

Utilizando el conjunto de datos _{de prueba} CitySounds2017 , evaluamos por separado el rendimiento de los dos algoritmos CityNet, CityBioNet y CityAnthroNet, utilizando dos medidas: precisión y recuperación. Los algoritmos CityBioNet y CityAnthroNet dan una estimación probabilística del nivel de actividad acústica biótica o antropogénica para cada fragmento de audio de 1 segundo como un número entre 0 y 1. Se podrían usar diferentes umbrales para convertir estas probabilidades en asignaciones de categorías de sonido (por ejemplo, 'sonido'. presente' o 'sonido ausente'). En cada umbral, se calculó un valor de precisión y recuperación, donde la precisión fue la fracción de fragmentos de 1 segundo identificados correctamente que contienen el sonido según las anotaciones en la _{prueba} CitySounds2017 , y la recuperación fue la fracción de fragmentos de 1 segundo etiquetados como que contienen el sonido que fue recuperado por el algoritmo por debajo de ese umbral. A medida que el umbral se movió entre 0 y 1, los valores resultantes de precisión y recuperación se trazaron como una curva de precisión-recuperación. Se calcularon estadísticas resumidas para la precisión promedio bajo todos los valores de umbral y la recuperación cuando el umbral elegido dio una precisión de 0,95. Utilizando un umbral de 0,5 en las predicciones, se calcularon matrices de confusión que muestran cómo se clasificó cada momento en relación con las anotaciones. Estos análisis se realizaron en Python v.2.7.12 (**Python Software Foundation 2016**) utilizando Scikit-learn v.0.18.1 (**Pedregosa *et al.* 2011**) y Matplotlib v.1.5.1 (**Hunter 2007**).

Algoritmos en competencia

También comparamos la precisión y la recuperación de los algoritmos de CityNet con las medidas acústicas producidas por cuatro IA: Índice de Complejidad Acústica (ACI) (**Pieretti, Farina y Morri 2011**), Índice de Diversidad Acústica (ADI) (**Villanueva-Rivera et al. 2011**), Índice bioacústico (BI) (**Boelman et al. 2007**) e índice de paisaje sonoro de diferencia normalizada (NDSI) (**Kasten et al. 2012**). El NDSI genera una medida de perturbación antropogénica según la fórmula

$$NDSI = \frac{NDSI_{bio} - NDSI_{anthro}}{NDSI_{bio} + NDSI_{anthro}} \quad \text{Equation 1}$$

donde $NDSI_{bio}$ y $NDSI_{anthro}$ son la actividad acústica biótica y antropogénica total en cada grabación, respectivamente. En lugar de comparar CityNet con el NDSI, comparamos los elementos bióticos ($NDSI_{bio}$) y antropogénicos ($NDSI_{anthro}$) del NDSI con las medidas producidas por CityBioNet y CityAnthroNet, respectivamente, ya que eran más comparables. Como todas las IA están diseñadas para brindar un resumen de la actividad acústica de un archivo completo, se analizaron en el conjunto de datos *de prueba* CitySounds2017 tratando cada fragmento de audio de 1 segundo como un archivo de sonido separado para permitir comparaciones directas con CityNet. Las medidas de IA no tienen un umbral natural para la clasificación en sonido biótico/no biótico, lo que significa que no pudimos calcular matrices de confusión. Sin embargo, se utilizó un umbral entre su valor más bajo y su valor más alto en combinación con el rango de precisión y valores de recuperación para formar curvas de recuperación de precisión. Todas las IA se calcularon en R v.3.4.1 (**R Core Team 2017**) utilizando 'seewave' v.1.7.6 (**Sueur, Aubin & Simonis 2008**) y 'soundecology' v.1.2 (**Villanueva-Rivera & Pijanowski 2014**) paquetes.

La precisión y la recuperación de CityBioNet también se compararon con bulbul (**Grill & Schlüter 2017**), un algoritmo para detectar sonidos de aves en grabaciones de audio completas con el fin de resumir la actividad acústica de las aves, que fue la propuesta ganadora en el desafío Bird Audio Detección 2016-7 (**Stowell et al. 2016**). Al igual que CityNet, bulbul es un clasificador basado en CNN que utiliza espectrogramas como entrada. Sin embargo, no utiliza las mismas estrategias de normalización que CityNet y no se entrenó con datos de entornos urbanos ruidosos. Se aplicó Bulbul a cada segundo de datos de audio en *la prueba* CitySounds2017, utilizando el modelo previamente entrenado proporcionado por los autores junto con su código.

Impacto de los sonidos no bióticos

Realizamos análisis adicionales sobre los sonidos no bióticos que afectan las predicciones de CityBioNet y bulbul, ya que se descubrió que eran los algoritmos de mejor rendimiento para medir el sonido biótico. Para hacer esto, creamos subconjuntos del conjunto de datos *de prueba* CitySounds2017 que comprenden todos los segundos que contenían una variedad de sonidos no bióticos, por ejemplo, un subconjunto de datos de tráfico que contiene todos los segundos en la *prueba* CitySounds2017 donde el sonido del tráfico estaba presente. Luego utilizamos una prueba de Chi-cuadrado para identificar

diferencias significativas en la proporción de segundos en los que cada algoritmo predijo correctamente la presencia/ausencia de sonido biótico en el umbral 0,5 en los conjuntos de datos completos y en el subconjunto, y se utilizó la estadística V de Cramer para evaluar el tamaño del efecto de las diferencias (**Cohen 1992**). Estos análisis se realizaron en R v.3.4.1 (**R Core Team 2017**).

Aplicación ecológica

Utilizamos CityNet para generar patrones promedio diarios de actividad acústica biótica y antropogénica para dos sitios de estudio a lo largo de un gradiente de urbanización (sitios E29RR e IG62XL con urbanización alta y baja respectivamente, **Tabla S1**). Para controlar la fecha de grabación; ambos sitios fueron encuestados entre mayo y junio de 2015. CityNet se ejecutó durante los 7 días completos de grabaciones de cada sitio para predecir la presencia/ausencia de sonido biótico y antropogénico por cada fragmento de audio de 1 segundo utilizando un umbral de probabilidad de 0,5. Se crearon medidas de actividad biótica y antropogénica para cada ventana de media hora entre medianoche y medianoche promediando el número previsto de segundos que contienen sonido biótico o antropogénico dentro de esa ventana durante toda la semana.

RESULTADOS

Rendimiento acústico

CityBioNet tuvo una precisión promedio de 0,934 y una recuperación de 0,710 con una precisión de 0,95, mientras que CityAnthroNet tuvo una precisión promedio de 0,977 y una recuperación de 0,858 con una precisión de 0,95 (**Tabla 1** , **Figura 3**). En comparación, ACI, ADI, BI y NDSI_{bio} tuvieron una precisión promedio más baja (0,663, 0,439, 0,516 y 0,503, respectivamente) y una recuperación más baja de 0,95 (todas menos de 0,01). CityBioNet también superó a bulbul, que tuvo una precisión promedio de 0,872 y una recuperación de 0,95 de 0,398 (**Tabla 1**). En comparación con CityAnthroNet, el NDSI_{anthro} tuvo una precisión promedio más baja (0,975) y una recuperación más baja con una precisión de 0,95 (0,815). Cuando el sonido biótico estaba presente en las grabaciones, CityBioNet predijo correctamente la presencia de sonido biótico (Verdaderos Positivos) en una mayor proporción de datos de audio que el bulbul (33,2% en comparación con el 18,5%, para CityBioNet y bulbul respectivamente) (**Figura 4**). Sin embargo, CityBioNet no pudo predecir correctamente la presencia de sonido biótico (falsos negativos) en el 1,7% de las grabaciones, en comparación con el 1,0% de predicciones incorrectas de bulbul. Cuando el sonido biótico estuvo ausente en las grabaciones, CityBioNet predijo correctamente la ausencia de sonido biótico (Verdaderos Negativos) en el 51,6% de los datos de audio en comparación con el 52,6% para bulbul, y CityBioNet no pudo predecir correctamente la ausencia de sonido biótico (Falsos Positivos) en el 13,5% de los datos de audio en comparación con el 20,0% de predicciones incorrectas de bulbul (**Figura 4**).

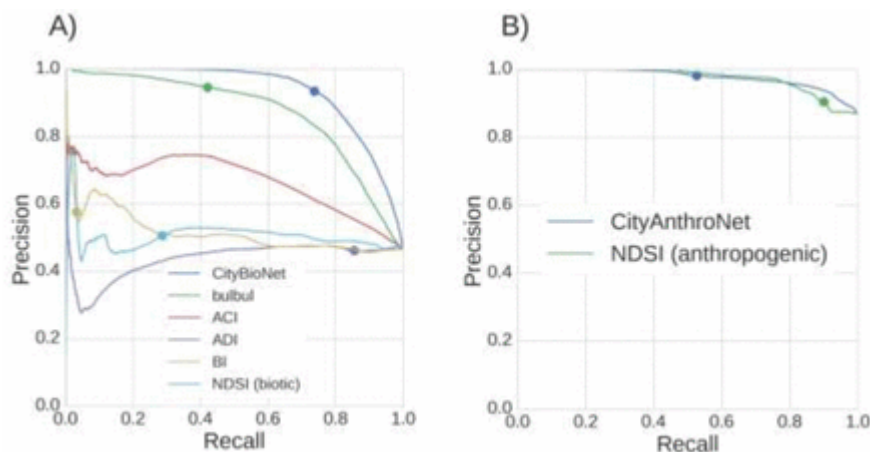


Figura 3.

Curvas de recuperación de precisión para CityNet y algoritmos de la competencia que predicen A) actividad acústica biótica y B) antropogénica para cada fragmento de audio de 1 segundo en el conjunto de datos *de prueba* CitySounds2017. Los puntos indican los valores de precisión y recuperación en un valor umbral de 0,5. ACI representa el índice de complejidad acústica, el índice de diversidad acústica ADI, el índice bioacústico BI y el índice de paisaje sonoro de diferencia normalizada *bio y antrobiótica* y antropogénica NDSI, respectivamente.

Tabla 1.

Resultados promedio de precisión y recuperación para CityNet y algoritmos de la competencia para cada fragmento de audio de 1 segundo en el conjunto de datos *de prueba* CitySounds2017. Los resultados de la recuperación se presentan con una precisión de 0,95. Los valores más altos son mejores para ambas métricas. Los valores más altos de cada sección se muestran en negrita. ACI representa el índice de complejidad acústica, el índice de diversidad acústica ADI, el índice bioacústico BI y el índice de paisaje sonoro de diferencia normalizada *bio y antrobiótica* y antropogénica NDSI, respectivamente.

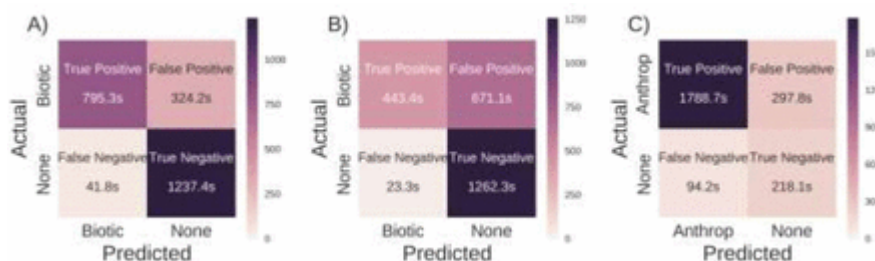


Figura 4.

Matrices de confusión que comparan la actividad acústica prevista de A) CityBioNet, B) bulbul y C) CityAnthroNet para cada fragmento de audio de 1 segundo en el conjunto de datos *de prueba* CitySounds2017. Los números en cada celda informan la cantidad de clips de audio de 1 segundo en el conjunto de datos *de prueba* de CitySounds2017 predichos correctamente (verdaderos positivos y verdaderos negativos) o incorrectamente (falsos positivos y falsos negativos) por contener bióticos (A y B) o antropogénicos (C). sonido. Para crear las matrices de confusión, las predicciones probabilísticas de los clasificadores se convierten en clasificaciones binarias utilizando un umbral que da una precisión de 0,95.

Impactos de los sonidos no bióticos

CityBioNet se vio fuertemente afectado negativamente (tamaño del efecto V de Cramer $>0,5$) por el sonido mecánico (la presencia/ausencia de sonido biótico se predijo correctamente en un 28,60 % menos de los datos cuando los sonidos mecánicos también estaban presentes) (**Tabla 2**). Bulbul se vio moderadamente afectado (tamaño del efecto V de Cramer 0,1-0,5) por el sonido del tráfico aéreo y el viento (la presencia/ausencia de sonido biótico se predijo correctamente en un 5,34% y un 6,93%

menos de los datos cuando se predijeron los sonidos del tráfico aéreo y del viento). también presentes en grabaciones, respectivamente).

Tabla 2.

Impacto de los sonidos no bióticos en CityBioNet y las predicciones de bulbul. Los valores representan diferencias en la proporción de fragmentos de audio de 1 segundo en el conjunto de datos *de prueba* completo de CitySound2017 (40 minutos) y los conjuntos de datos del subconjunto (tamaño en tiempo indicado en la columna de la izquierda) en los que la presencia/ausencia de sonido biótico fue predicha correctamente por ambos algoritmos (estadística de prueba de chi cuadrado para la diferencia en proporciones de éxitos en cada conjunto de datos y medida del tamaño del efecto V de Cramer). Los tamaños del efecto se indican como $<0,1$ (*), $0,1-0,3$ (**) y $>0,5$ (***) .

Aplicación ecológica

CityNet produjo patrones realistas de actividad acústica biótica y antropogénica en el paisaje sonoro urbano en dos sitios de estudio de baja y alta intensidad urbana (**Figura 2B y C**). En ambos sitios, la actividad acústica biótica alcanzó su punto máximo justo después del amanecer y disminuyó rápidamente después del atardecer. Un segundo pico de actividad acústica biótica se registró al atardecer en el sitio de baja intensidad urbana, pero no en el sitio de alta intensidad urbana. En ambos sitios la actividad acústica antropogénica aumentó bruscamente después del amanecer, permaneció constante durante todo el día y disminuyó después del atardecer.

DISCUSIÓN

Tanto CityBioNet como CityAnthroNet superaron a los algoritmos de la competencia en el conjunto de datos *de prueba* CitySound2017 . CityBioNet obtuvo mejores resultados que bulbul en grabaciones ruidosas del entorno urbano; era resistente a sonidos más no bióticos, incluido el tráfico rodado, el tráfico aéreo y la lluvia. Ser resistente al sonido del tráfico rodado respalda la idoneidad de CityBioNet para su uso en ciudades, ya que el paisaje sonoro urbano está dominado por el sonido del tráfico rodado (**Fairbrass et al. 2017**), lo que se ha demostrado que influye en varias de las IA probadas aquí (**Fuller y otros 2015** ; **Fairbrass y otros 2017**). También se ha demostrado que el sonido de la lluvia influye en varias IA (**Depraetere et al. 2012** ; **Gasc et al. 2015b** ; **Fairbrass et al. 2017**) y el desarrollo de un método que sea resistente a este sonido es una contribución considerable al campo. de la ecoacústica. El paisaje sonoro biótico urbano está dominado por los sonidos emitidos por las aves (**Fairbrass et al. 2017**), y el buen desempeño del bulbul, un algoritmo para medir exclusivamente los sonidos de las aves, en el conjunto de datos *de prueba de* CitySounds2017 , lo confirma. Las aves se utilizan como especies indicadoras en los esquemas existentes de monitoreo de la biodiversidad urbana (**Kohsaka et al. 2013**) utilizando datos recopilados de formas tradicionales de estudio de la biodiversidad. Los algoritmos desarrollados aquí podrían usarse para respaldar los esquemas existentes al facilitar la recopilación de datos sobre estos taxones indicadores.

CityNet es el único método disponible actualmente para medir la actividad acústica tanto biótica como antropogénica utilizando un único sistema en datos de audio ruidosos de entornos urbanos. Cada vez

hay más pruebas de que el ruido antropogénico afecta a la vida silvestre de diversas maneras, incluida la alteración del comportamiento de comunicación (**Gil y Brumm 2014**) y el uso del hábitat (**Deichmann *et al.* 2017**). Sin embargo, estas investigaciones tienen una escala limitada por el uso de métodos intensivos en recursos para medir el sonido biótico y antropogénico en el medio ambiente o a partir de datos de audio. Otros dependen de la IA (**Pieretti y Farina 2013**), que se ha demostrado que no son confiables en entornos con perturbaciones acústicas (**Fairbrass *et al.* 2017**). CityNet podría facilitar la investigación de los impactos de las actividades antropogénicas en las poblaciones de vida silvestre a escalas que actualmente no son posibles con los métodos tradicionales de análisis acústico.

CityBioNet superó claramente a todas las IA probadas, pero la diferencia de rendimiento entre CityAnthroNet y el algoritmo competidor para medir la actividad acústica antropogénica (NDSI_{anthro}) fue mucho menos marcada. Estos resultados sugieren que la medición del sonido biótico en datos de audio ruidosos de entornos urbanos requiere algoritmos más sofisticados que la medición del sonido antropogénico. Posiblemente los sonidos antropogénicos sean más fácilmente separables de otros sonidos en el espacio de frecuencias, una teoría que es la base de varias IA (**Boelman *et al.* 2007** ; **Kasten *et al.* 2012**), lo que facilita el uso de algoritmos definidos por humanos como el NDSI_{anthro} . Considerando que, debido a que los sonidos bióticos ocurren en un espacio de frecuencia compartido con sonidos antropogénicos y geofónicos (**Fairbrass *et al.* 2017**), los algoritmos como las IA, que solo utilizan una pequeña cantidad de características para discriminar sonidos, no son suficientes para su uso en las ciudades. Por lo tanto, los algoritmos de ML que pueden utilizar una mayor cantidad de funciones para discriminar sonidos, como las CNN implementadas en el sistema CityNet, son más capaces de detectar sonidos bióticos en grabaciones que también contienen sonidos no bióticos. Un método reciente no supervisado desarrollado por **Lin, Fang y Tsao (2017)** para separar sonidos biológicos de grabaciones largas podría usarse como paso de preprocesamiento para mejorar aún más el rendimiento de CityNet.

Los sensores acústicos y algoritmos de bajo costo para la medición automática del sonido biótico en datos de audio están facilitando la evaluación y el monitoreo de la biodiversidad a grandes escalas temporales y espaciales (**Sueur & Farina 2015**), pero hasta la fecha esta tecnología solo se ha implementado en áreas no urbanas. ambientes (por ejemplo, **Aide *et al.* 2013**). En las ciudades, la disponibilidad de red eléctrica y conexiones Wifi está apoyando el desarrollo del Internet de las cosas (IoT) urbano utilizando sensores integrados en la infraestructura existente para monitorear factores ambientales como la contaminación del aire, los niveles de ruido y el uso de energía (**Zanella *et al.* 2014**). El sistema CityNet podría integrarse en una red de detección de IoT para facilitar la evaluación ambiental urbana a gran escala. La implementación a gran escala de algoritmos como CityNet requiere un bajo uso de energía y tiempos de ejecución rápidos. Una forma de ayudar a lograr este objetivo sería combinar las dos redes (CityBioNet y CityAnthroNet) en una CNN que prediga simultáneamente la actividad acústica biótica y antropogénica.

Una expansión de CityNet a frecuencias ultrasónicas aumentaría la generalidad de la herramienta, ya que podría usarse para monitorear especies en ciudades que emiten sonidos en frecuencias superiores a 12 kHz, como murciélagos y algunos invertebrados. Los murciélagos se utilizan con frecuencia como indicadores ecológicos porque son sensibles a los cambios ambientales (**Walters *et al.* 2013**). Los métodos acústicos se utilizan comúnmente para monitorear las poblaciones de murciélagos utilizando grabadores ultrasónicos pasivos, lo que significa que los investigadores y conservacionistas de murciélagos se enfrentan al desafío de extraer información significativa de grandes volúmenes de datos de audio. El desarrollo de métodos automatizados para medir los llamados de los murciélagos en datos ultrasónicos se ha centrado hasta la fecha en la identificación de los llamados de las especies de murciélagos y muchos algoritmos son propietarios (p. ej., **Szewczak 2010** ; **Wildlife Acoustics 2017**). El desarrollo de un algoritmo de código abierto que produzca medidas de murciélagos a nivel comunitario sería una valiosa adición a la caja de herramientas de los investigadores y conservacionistas de murciélagos.

Reentrenar CityNet con datos de audio etiquetados de otras ciudades haría posible utilizar el sistema para monitorear la actividad acústica antropogénica y biótica urbana de manera más amplia. Sin embargo, como Londres es una ciudad grande y heterogénea, CityNet se ha entrenado utilizando un conjunto de datos que contiene sonidos que caracterizan una amplia gama de entornos urbanos. Nuestra recopilación de datos se limitó a una sola semana en cada sitio de estudio, lo que limita nuestra capacidad para evaluar la capacidad del sistema CityNet para detectar cambios ambientales. El trabajo futuro debería centrarse en la recopilación de datos acústicos longitudinales para evaluar la sensibilidad de los algoritmos para detectar cambios ambientales. Nuestro uso de etiquetadores humanos habría introducido subjetividad y sesgo en nuestro conjunto de datos. La tarea de anotar grandes conjuntos de datos de audio de entornos urbanos acústicamente complejos requiere una gran cantidad de recursos, un problema que se ha abordado recientemente con científicos ciudadanos para crear los conjuntos de datos UrbanSounds y UrbanSound8k utilizando datos de audio de la ciudad de Nueva York, EE. UU. (**Salamon, Jacoby & Bello 2014**). Se componen de breves fragmentos de 10 sonidos urbanos diferentes, como martillos neumáticos, motores en ralentí y disparos. Estos conjuntos de datos no representan completamente las características de los paisajes sonoros urbanos por tres razones. En primer lugar, suponen que sólo hay una clase de sonido presente en cada momento, mientras que en realidad pueden estar presentes varios tipos de sonido al mismo tiempo (considere el canto de un pájaro mientras un avión sobrevuela). En segundo lugar, solo incluyen sonidos antropogénicos, mientras que CityNet mide tanto sonidos antropogénicos como bióticos. Finalmente, cada archivo en estos conjuntos de datos tiene un sonido presente, mientras que los paisajes sonoros urbanos contienen muchos períodos de silencio o sonidos geofónicos, dos estados importantes que no están presentes en UrbanSounds y UrbanSound8k. Debido a estos factores, estos conjuntos de datos no son adecuados para el propósito de este proyecto de investigación, aunque un trabajo reciente ha superado algunas de estas deficiencias utilizando datos de paisajes sonoros sintetizados (**Salamon *et***

al. 2017). Esto resalta la necesidad de un esfuerzo coordinado internacionalmente para crear un conjunto de datos de audio etiquetados consistentemente de las ciudades para apoyar el desarrollo de sistemas automatizados de evaluación ambiental urbana con aplicación internacional.

Conclusiones

El sistema CityNet para medir la actividad acústica biótica y antropogénica en datos de audio urbanos ruidosos superó a los algoritmos de última generación para medir el sonido biótico y antropogénico en grabaciones de audio completas. Integrado en una red de IoT para grabar y analizar datos de audio en las ciudades, podría facilitar la evaluación ambiental urbana a mayores escalas de lo que ha sido posible hasta la fecha utilizando métodos tradicionales de evaluación de la biodiversidad.

Hacemos que nuestro sistema esté disponible en código abierto en combinación con dos conjuntos de datos de paisajes sonoros urbanos anotados por expertos para facilitar el desarrollo de investigaciones futuras en este campo.

DECLARACIÓN DE CONTRIBUCIÓN DEL AUTOR

AF, MF, HT y KJ concibieron ideas y diseñaron metodología; AF recopiló los datos; AF y MF analizaron los datos y dirigieron la redacción del manuscrito. Todos los autores contribuyeron críticamente a los borradores y dieron la aprobación final para su publicación.

ACCESIBILIDAD DE DATOS

Todas las grabaciones y anotaciones del conjunto de datos CitySounds2017 y todo el código Python subyacente a los algoritmos CityNet están disponibles en Figshare (<https://figshare.com/s/adab62c0591afaeafedd>).

EXPRESIONES DE GRATITUD

Agradecemos a varios propietarios y administradores de sitios por apoyar el estudio brindando acceso a los sitios de grabación, y a múltiples anotadores acústicos y a un experto en transporte por ayudarnos a crear el conjunto de datos CitySounds2017. Recibimos el apoyo financiero de una subvención catalizadora de recursos sostenibles para ciudades sostenibles de BHP Billiton y del Consejo de Investigación en Ciencias Físicas y de Ingeniería (EPSRC) a través de una subvención de formación doctoral (EP/G037698/1) para HT y una subvención de EPSRC (EP/K015664/ 1) a KEJ, GB y MF

REFERENCIAS

- ↵ Acevedo, MA, Corrada-Bravo, CJ, Corrada-Bravo, H., Villanueva-Rivera, LJ & Aide, TM (2009) Clasificación automatizada de cantos de aves y anfibios mediante aprendizaje automático: una comparación de métodos . *Informática Ecológica* , **4** , 206 – 214 . **Google Académico**
- ↵ Aide, TM, Corrada-Bravo, C., Campos-Cerqueira, M., Milan, C., Vega, G. & Alvarez, R. (2013) *Monitoreo bioacústico en tiempo real e identificación automatizada de especies* . **1** , e103 . Disponible: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3719130/pdf/peerj-01-103.pdf> Consultado: 19/12/2016 **Google Académico**
- ↵ Aronson, MFJ, La Sorte, FA, Nilon, CH, Katti, M., Goddard, MA, Lepczyk, CA, Warren, PS, Williams, NSG, Cilliers, S., Clarkson, B., Dobbs, C., Dolan, R., Hedblom, M., Klotz, S., Kooijmans, JL, Kühn, I., MacGregor-Fors, I., McDonnell, M., Mörtberg, U., Pyšek, P., Siebert, S., Sushinsky, J., Werner, P. & Winter, M. (2014) *Un análisis global de los impactos de la urbanización en la diversidad de aves y plantas revela factores antropogénicos clave* . **281** , 20133330 . Disponible: <http://rspb.royalsocietypublishing.org/content/281/1780/20133330.abstract> Consultado: 12/02/2016 **Google Académico**
- ↵ Beninde, J., Veith, M. & Hochkirch, A. (2015) La biodiversidad en las ciudades necesita espacio: un metaanálisis de los factores que determinan la variación de la biodiversidad intraurbana . *Cartas de Ecología* , **18** , 581 – 592 . **Google Académico**
- ↵ Boelman, NT, Asner, GP, Hart, PJ y Martin, RE (2007) Resistencia a la invasión multitrófica en Hawái: bioacústica, estudios de campo y teledetección aérea . *Aplicaciones ecológicas* , **17** , 2137 – 2144 . **PubMed Google Académico**
- ↵ Chesmore, E. y Ohya, E. (2004) Identificación automatizada de canciones grabadas en el campo de cuatro saltamontes británicos mediante reconocimiento de señales bioacústicas . *Boletín de Investigaciones Entomológicas* , **94** , 319 – 330 . **PubMed Google Académico**
- ↵ Cohen, J. (1992) Análisis de poder estadístico . *Direcciones actuales en la ciencia psicológica* , **1** , 98 – 101 . **Referencia cruzada Google Académico**
- ↵ Crouse, DL, Pinault, L., Balram, A., Hystad, P., Peters, PA, Chen, H., van Donkelaar, A., Martin, RV, Ménard, R., Robichaud, A. y Villeneuve, PJ (2017) Verdor urbano y mortalidad en las ciudades más grandes de Canadá: un estudio de cohorte nacional . *The Lancet Planetary Health* , **1** , e289 – e297 . **Google Académico**
- ↵ Deichmann, JL, Hernández-Serna, A., Delgado C, JA, Campos-Cerqueira, M. & Aide, TM (2017) El análisis del paisaje sonoro y el monitoreo acústico documentan los impactos de la exploración de gas natural en la biodiversidad en un bosque tropical . *Indicadores ecológicos* , **74** , 39 – 48 . **Google Académico**
- ↵ Depraetere, M., Pavoine, S., Jiguet, F., Gasc, A., Duvail, S. & Sueur, J. (2012) Monitoreo de la diversidad animal mediante índices acústicos: implementación en un bosque templado . *Indicadores ecológicos* , **13** , 46 – 54 . **Referencia cruzada Google Académico**
- ↵ Dieleman, S., Schlüter, J., Raffel, C., Olson, E., Sønderby, SK, Nouri, D., Maturana, D., Thoma, M., Battenberg, E., Kelly, J., De Fauw, J., Heilman, M., de Almeida, DM, McFee, B., Weideman, H., Takács, G., de Rivaz, P., Crall, J., Sanders, G., Rasul, K., Liu, C., French, G. & Degraeve, J. (2015) *Lasaña* . Disponible: <http://dx.doi.org/10.5281/zenodo.27878> Consultado: 19/09/2017 **Google Académico**
- ↵ Digby, A., Towsey, M., Bell, BD y Teal, PD (2013) Una comparación práctica de métodos manuales y autónomos para el monitoreo acústico . *Métodos en Ecología y Evolución* , **4** , 675 – 683 . **Google Académico**
- ↵ Faeth, SH, Bang, C. & Saari, S. (2011) Biodiversidad urbana: patrones y mecanismos . *Año en Ecología y Biología de la Conservación* , **1223** , 69 – 81 . **Google Académico**
- ↵ Fairbrass, AJ, Rennett, P., Williams, C., Titheridge, H. & Jones, KE (2017) Sesgos de los índices acústicos que miden la biodiversidad en áreas urbanas . *Indicadores ecológicos* , **83** , 169 – 177 . **Google Académico**

- ↵Farinha-Marques , P., Lameiras , J., Fernandes , C., Silva , S. & Guilherme , F. (2011) Biodiversidad urbana: una revisión de conceptos actuales y contribuciones a enfoques multidisciplinares . *Innovación: Revista europea de investigación en ciencias sociales* , **24** , 247 – 271 . **Google Académico**
- ↵Fuller , S., Axel , AC , Tucker , D. & Gage , SH (2015) Conectando el paisaje sonoro con el paisaje: ¿Qué índice acústico describe mejor la configuración del paisaje ? *Indicadores Ecológicos* , **58** , 7.07 – 7.15 **Google Académico**
- ↵Gasc , A., Pavoine , S., Lellouch , L., Grandcolas , P. & Sueur , J. (2015a) Índices acústicos para evaluaciones de biodiversidad: análisis de sesgos basados en conjuntos de aves simulados y recomendaciones para estudios de campo . *Conservación biológica* , **191** , 306 – 312 . **Google Académico**
- ↵Gasc , A., Pavoine , S., Lellouch , L., Grandcolas , P. & Sueur , J. (2015b) Índices acústicos para evaluaciones de biodiversidad: análisis de sesgos basados en conjuntos de aves simulados y recomendaciones para estudios de campo . *Conservación biológica* , **191** , 306 – 312 . **Google Académico**
- ↵Gil , D. & Brumm , H. (2014) Comunicación acústica en el entorno urbano: patrones, mecanismos y posibles consecuencias de los ajustes del canto de las aves . *Ecología urbana aviar* (eds D. Gil & H. Brumm), págs. 69 – 83 . Oxford University Press , Oxford, Reino Unido . **Google Académico**
- ↵Grill , T. & Schlüter , J. (2017) Dos redes neuronales convolucionales para la detección de aves en señales de audio . 25ª Conferencia Europea de Procesamiento de Señales (EUSIPCO2017) . Cos , Grecia . **Google Académico**
- ↵Hall , DM , Camilo , GR , Tonietto , RK , Smith , DH , Ollerton , J., Ahrné , K., Arduser , M., Ascher , JS , Baldock , KC & Fowler , R. (2016) La ciudad como refugio para insectos polinizadores . *Biología de la conservación* , **31** , 24 – 29 . **Google Académico**
- ↵Hunter , JD (2007) Matplotlib: un entorno de gráficos 2D . *Computación en ciencia e ingeniería* , **9** , 90 – 95 . **Referencia cruzada Google Académico**
- Ioffe , S. y Szegedy , C. (2015) Normalización de lotes: acelerar el entrenamiento profundo de la red reduciendo el cambio de covariables interno . Actas de la 32ª Conferencia Internacional sobre Aprendizaje Automático , págs. 448 – 456 . Lille, Francia . **Google Académico**
- ↵Kasten , EP , Gage , SH , Fox , J. & Joo , WV. (2012) La biblioteca acústica del laboratorio de evaluación ambiental remota: un archivo para estudiar la ecología del paisaje sonoro . *Informática Ecológica* , **12** , 50 – 67 . **Google Académico**
- Kingma , D. & Ba , J. (2015) Adam: un método para la optimización estocástica . Actas de la Conferencia Internacional sobre Representaciones del Aprendizaje 2015 . San Diego, Estados Unidos . **Google Académico**
- ↵Kohsaka , R., Pereira , HM , Elmqvist , T., Chan , L., Moreno-Peñaranda , R., Morimoto , Y., Inoue , T., Iwata , M., Nishi , M. & da Luz Mathias , M. (2013) Indicadores para la gestión de la biodiversidad urbana y los servicios ecosistémicos: índice de biodiversidad de la ciudad . *Urbanización, biodiversidad y servicios ecosistémicos: desafíos y oportunidades* (eds T. Elmqvist , M. Fragkias , J. Goodness , B. Güneralp , PJ Marcotullio , RI McDonald , S. Parnell , M. Schewenius , M. Sendstad , KC Seto & C .Wilkinson) , págs.699-718. Springer , Países Bajos . **Google Académico**
- ↵LeCun , Y., Bengio , Y. & Hinton , G. (2015) Aprendizaje profundo . *Naturaleza* , **521** , 436 – 444 . **Referencia cruzada PubMed Google Académico**
- ↵Lee , H., Pham , P., Largman , Y. & Ng , AY (2009) Aprendizaje de funciones no supervisado para la clasificación de audio mediante redes convolucionales de creencias profundas . Actas de la 22ª Conferencia Internacional sobre Sistemas de Procesamiento de Información Neural , págs. 1096 – 1104 . Istanbul, Turquía . **Google Académico**

- ← Lin, T.-H., Fang, S.-H. & Tsao, Y. (2017) *Mejora de la evaluación de la biodiversidad mediante la separación no supervisada de sonidos biológicos de grabaciones de larga duración* . 7 . Disponible: <https://www.nature.com/articles/s41598-017-04790-7> Consultado: 19/09/2017 **Google Académico**
- Maas, AL, Hannun, AY y Ng, AY (2013) Las no linealidades del rectificador mejoran los modelos acústicos de redes neuronales . Actas de la 30ª Conferencia Internacional sobre Aprendizaje Automático . Atlanta, Estados Unidos . **Google Académico**
- ← McFee, B., Raffel, C., Liang, D., Ellis, DP, McVicar, M., Battenberg, E. y Nieto, O. (2015) librosa: Análisis de señales de audio y música en Python . Actas de la 14ª conferencia sobre Python en la ciencia , págs . 18-25 . Austin, Texas . **Google Académico**
- ← Natural England (2016) *Vínculos entre los entornos naturales y la salud mental: resumen de evidencia* . Disponible: <http://publications.naturalengland.org.uk> Consultado: 24/11/2017 **Google Académico**
- ← Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M. & Perrot, MD, E. (2011) Scikit-learn: aprendizaje automático en Python . *Revista de investigación sobre aprendizaje automático* , **12** , 2825 – 2830 . **Web de la Ciencia Google Académico**
- ← Pieretti, N. & Farina, A. (2013) Aplicación de un índice de complejidad acústica recientemente introducido a un paisaje sonoro de aves con ruido de tráfico . *Revista de la Sociedad Acústica de América* , **134** , 891-900 . **Referencia cruzada PubMed Google Académico**
- ← Pieretti, N., Farina, A. & Morri, D. (2011) Una nueva metodología para inferir la actividad de canto de una comunidad aviar: el Índice de Complejidad Acústica (ACI) . *Indicadores ecológicos* , **11** , 868 – 873 . **Referencia cruzada Google Académico**
- ← Pijanowski, BC, Villanueva-Rivera, LJ, Dumyahn, SL, Farina, A., Krause, BL, Napoletano, BM, Gage, SH & Pieretti, N. (2011) Ecología del paisaje sonoro: la ciencia del sonido en el paisaje . *Biociencia* , **61** , 203 – 216 . **Referencia cruzada Google Académico**
- ← Python Software Foundation (2016) *Referencia del lenguaje Python* . Disponible: <http://www.python.org> Acceso: 19/09/2017 **Google Académico**
- ← R Core Team (2017) *R: Un lenguaje y entorno para la informática estadística* . Disponible: <http://www.R-project.org> Acceso: 31/10/2014 **Google Académico**
- ← Salamon, J., Jacoby, C. & Bello, JP (2014) Un conjunto de datos y una taxonomía para la investigación del sonido urbano . *ACMMM'14* , págs. 1041 – 1044 . Asociación de Maquinaria de Computación, Orlando , EE.UU. **Google Académico**
- ← Salamon, J., MacConnell, D., Cartwright, M., Li, P. y Bello, JP (2017) Scaper: una biblioteca para síntesis y aumento de paisajes sonoros . Taller IEEE 2017 sobre aplicaciones del procesamiento de señales al audio y la acústica . New Paltz, Nueva York . **Google Académico**
- ← Srivastava, N., Hinton, GE, Krizhevsky, A., Sutskever, I. y Salakhutdinov, R. (2014) Abandono: una forma sencilla de evitar el sobreajuste de las redes neuronales .de *investigación sobre aprendizaje automático* , **15** , 1929-1958 . **Referencia cruzada Google Académico**
- ← Stowell, D. y Plumbley, MD (2014) *La clasificación automática a gran escala de los sonidos de las aves mejora considerablemente mediante el aprendizaje de funciones no supervisado* . **2** , e488 . Disponible: <http://dx.doi.org/10.7717/peerj.488> Consultado: 12/09/2016 **Google Académico**
- ← Stowell, D., Wood, M., Stylianou, Y. & Glotin, H. (2016) Detección de aves en audio: una encuesta y un desafío .de 2016 sobre aprendizaje automático para el procesamiento de señales , págs . IEEE, Vietri sul Mare , Italia . **Google Académico**

- ↩Sueur, J., Aubin, T. & Simonis, C. (2008) Revisión de equipos: seewave, una herramienta modular gratuita para análisis y síntesis de sonido . *Bioacústica* , **18** , 213 – 226 . **Referencia cruzada** **Web de la Ciencia** **Google Académico**
- ↩Sueur, J. & Farina, A. (2015) Ecoacústica: la investigación ecológica y la interpretación del sonido ambiental . *Biosemiótica* , **8** , 493 – 502 . **Google Académico**
- ↩Sueur, J., Farina, A., Gasc, A., Pieretti, N. & Pavoine, S. (2014) Índices acústicos para la evaluación de la biodiversidad y la investigación del paisaje . *Acta Acustica unida con Acustica* , **100** , 772 – 781 . **Referencia cruzada** **Google Académico**
- ↩Szewczak, JM (2010) *SonoBat* . Disponible: www.sonobat.com Acceso: 29/05/2014 **Google Académico**
- ↩El equipo de desarrollo de Theano, Al-Rfou, R., Alain, G., Almahairi, A., Angermueller, C., Bahdanau, D., Ballas, N., Bastien, F., Bayer, J. y Belikov, A. (2016) Theano: un marco de *Python* para el cálculo rápido de expresiones matemáticas . Disponible: <https://arxiv.org/abs/1605.02688> Consultado: 19/09/2017 **Google Académico**
- ↩Towsey, M., Wimmer, J., Williamson, I. y Roe, P. (2014) El uso de índices acústicos para determinar la riqueza de especies de aves en grabaciones de audio del medio ambiente . *Informática Ecológica* , **21** , 110 – 119 . **Referencia cruzada** **Google Académico**
- ↩ONU-DESA (2016) Las ciudades del mundo en 2016 . *Folleto de datos* . Disponible: <http://www.un.org/en/development/desa/population/> Consultado: 02/10/2017 **Google Académico**
- ↩Villanueva-Rivera, LJ & Pijanowski, BC (2014) Paquete 'sonecología' . *Ecología del paisaje sonoro* . Disponible: <http://cran.r-project.org/web/packages/soundecology/index.html> Consultado: 15/04/2015 **Google Académico**
- ↩Villanueva-Rivera, LJ, Pijanowski, BC, Doucette, J. & Pekin, B. (2011) Introducción al análisis acústico para ecólogos del paisaje . *Ecología del paisaje* , **26** , 1233 – 1246 . **Google Académico**
- ↩Walters, CL, Collen, A., Lucas, T., Mroz, K., Sayer, CA y Jones, KE (2013) Desafíos del uso de la bioacústica para monitorear globalmente los murciélagos . *Evolución, ecología y conservación de los murciélagos* , págs. 479 – 499 . Saltador: **Google Académico**
- ↩Walters, CL, Freeman, R., Collen, A., Dietz, C., Brock Fenton, M., Jones, G., Obrist, MK, Puechmaille, SJ, Sattler, T., Siemers, BM, Parsons, S & Jones, KE (2012) a escala continental para la identificación acústica de murciélagos europeos . *Revista de Ecología Aplicada* , **49** , 1064 – 1074 . **Referencia cruzada** **Google Académico**
- ↩Acústica de vida silvestre, I. (2017) *Software de análisis de caleidoscopio* . Disponible: <http://www.wildlifeacoustics.com/products/kaleidoscope-software-ultrasonic> Consultado: 24/08/2017 **Google Académico**
- ↩Zamora-Gutiérrez, V., López-González, C., MacSwiney González, MC, Fenton, B., Jones, G., Kalko, EK, Puechmaille, SJ, Stathopoulos, V. & Jones, KE (2016) Identificación acústica de murciélagos mexicanos basado en limitaciones taxonómicas y ecológicas en el diseño de llamadas . *Métodos en Ecología y Evolución* , **7** , 1082 – 1091 . **Google Académico**
- ↩Zanella, A., Bui, N., Castellani, A., Vangelista, L. & Zorzi, M. (2014) Internet de las cosas para ciudades inteligentes . *Revista IEEE Internet de las cosas* , **1** , 22 – 32 . **Referencia cruzada** **Google Académico**

REFERENCIAS

- ↩Ioffe, S. y Szegedy, C. (2015) Normalización de lotes: acelerar el entrenamiento profundo de la red reduciendo el cambio de covariables interno . *Actas de la 32ª Conferencia Internacional sobre Aprendizaje Automático* , págs. 448 – 456 . Lille, Francia . **Google Académico**

↩ Kingma , D. & Ba , J. (2015) Adam: un método para la optimización estocástica .Actas de la Conferencia Internacional sobre Representaciones del Aprendizaje 2015 . San Diego, Estados Unidos . **Google Académico**

↩ Maas ,AL , Hannun ,AY y Ng ,AY (2013) Las no linealidades del rectificador mejoran los modelos acústicos de redes neuronales . Actas de la 30ª Conferencia Internacional sobre Aprendizaje Automático .Atlanta, Estados Unidos . **Google Académico**

Srivastava , N. , Hinton , GE , Krizhevsky , A. , Sutskever , I. y Salakhutdinov , R. (2014) Abandono: una forma sencilla de evitar el sobreajuste de las redes neuronales .de *investigación sobre aprendizaje* automático, **15** , 1929-1958 . **Referencia cruzada**
Google Académico

⬆ Volver arriba

⬅ Anterior

Próximo ➡

Publicado el 16 de enero de 2018.

📄 **Descargar PDF**

✉ Correo electrónico

📄 **Opciones de impresión/guardado**

🔄 Compartir

🌐 Herramientas de citas

Post

Preimpresiones de COVID-19 SARS-CoV-2 de medRxiv y bioRxiv

Área temática

Ecología ▶

Áreas temáticas

Todos los artículos

Comportamiento y cognición animal

Bioquímica

Bioingeniería

Bioinformática

Biofísica

Biología del cáncer

Biología Celular

Ensayos clínicos*

Biología del desarrollo

Ecología

Epidemiología*

Biología evolucionaria

Genética

Genómica

Inmunología

Microbiología

Biología Molecular

Neurociencia

Paleontología

Patología

Farmacología y Toxicología

Fisiología

Biología vegetal

Comunicación Científica y Educación

Biología sintética

Biología de sistemas

Zoología

* Las categorías temáticas de Ensayos clínicos y Epidemiología ahora están cerradas a nuevas presentaciones luego de la finalización del proyecto piloto de investigación clínica de bioRxiv y el lanzamiento del servidor dedicado a ciencias de la salud medRxiv (submit.medrxiv.org). Los nuevos artículos que informen sobre los resultados de los ensayos clínicos ahora deben enviarse a medRxiv. La mayoría de los artículos nuevos sobre epidemiología también deben enviarse a medRxiv, pero si un artículo no contiene información relacionada con la salud, los autores pueden optar por enviarlo a otra categoría temática de bioRxiv (por ejemplo, genética o microbiología).

