Revista Cubana de Ciencias Informáticas
Vol. 18, No. 3, Julio-Septiembre, 2024
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 48-62

Tipo de artículo: Artículo original

Temática: Procesamiento de imágenes

Huella digital de audio aplicada al reconocimiento de formas

Audio fingerprint applied to shape recognition

Luis Rafael Baez Gonzalez [1*] https://orcid.org/0000-0003-3098-1121

[1] Desoft, División Territorial Santiago de Cuba. Calle 9 s/n e/ A y B Rpto. Santiago de Cuba, Cuba.

*Autor para la correspondencia. (lrbaez160@gmail.com)

Revista Cubana de Ciencias Informáticas
Vol. 18, No. 3, Julio-Septiembre, 2024
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 48-62

RESUMEN

La tecnología de reconocimiento de formas es crucial en la visión por computadora, permitiendo la identificación precisa de objetos en imágenes y secuencias de video. Su alcance es amplio, abarcando campos como la biología, medicina y aplicaciones militares. Históricamente, estos procesos eran manuales; sin embargo, con el avance de la informática, se han automatizado en gran medida. Este estudio presenta un innovador enfoque para el reconocimiento de formas en imágenes digitales, empleando espectrogramas de Mel como huella digital de audio, una técnica no convencional en este campo. Las imágenes utilizadas se encuentran ya segmentadas en las bases de datos. Cada imagen se transformó a una serie temporal guardada como archivo wav. Se utilizó el método de mapa de constelaciones usado por el sitio web Shazam para clasificar las formas, y se llevaron a cabo experimentos con dos bases de datos: MPEG-7 CE-Shape-1 Part B y ETU-10. Los resultados obtenidos muestran una precisión del 97.92% y del 100% en las bases de datos mencionadas, evidenciando la efectividad y el potencial de este novedoso enfoque en el reconocimiento de formas en imágenes digitales.

Palabras clave: espectrograma de Mel; mapa de constelación; reconocimiento de formas.

ABSTRACT

Shape recognition technology is crucial in computer vision, enabling accurate identification of objects in images and video sequences. Its scope is broad, covering fields such as biology, medicine and military applications. Historically, these processes were manual; However, with the advancement of computing, they have become largely automated. This study presents an innovative approach to shape recognition in digital images, using Mel spectrograms as an audio fingerprint, an unconventional technique in this field. The images used are already segmented in the databases. Each image was transformed to a time series saved as a wav file. The constellation map method used by the Shazam website was used to classify shapes, and experiments were carried out with two databases: MPEG-7 CE-Shape-1 Part B and ETU-10. The results obtained show an accuracy of 97.92% and 100% in the aforementioned databases, evidencing the effectiveness and potential of this novel approach in the recognition of shapes in digital images.

Revista Cubana de Ciencias Informáticas
Vol. 18, No. 3, Julio-Septiembre, 2024
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 48-62

Keywords: Mel spectrogram; constellation map; shape recognition.

Recibido: 02/05/2024

Aceptado: 10/06/2024

Introducción

La evolución tecnológica ha revolucionado la forma en que interactuamos con el mundo que nos rodea; desde la visión por computadora hasta la identificación de formas han experimentado un crecimiento exponencial a través de innovadoras técnicas como la inteligencia artificial u otros métodos de automatización en la actualidad. Según (Virginia Kindergarten Readiness Program, 2019), “Reconocer formas implica distinguir entre formas y asociarlas con nombres de formas. Seleccionar un cuadrado de un grupo de diferentes formas y proporcionar el nombre "cuadrado" es un ejemplo de cómo reconocer una forma.”. Si bien el problema más general en visión por computadora implica objetos arbitrarios, en el contexto de este trabajo se restringirá exclusivamente a la silueta del objeto. No se tendrán en cuenta ninguna de las propiedades que pudieran encontrarse dentro de la silueta como pueden ser: color, textura, oquedades u otras. Para métodos tradicionales en el reconocimiento de formas (Mingqiang et al., 2008) presentan una colección bastante exhaustiva de ellos. Si bien durante el trabajo se citan diversas fuentes que realizan el reconocimiento de formas, no se conoce en el momento de su redacción, ningún trabajo que utilice técnicas de audio aplicadas al reconocimiento de formas. Para el indexado y recuperación de audio (Kim, Moreau y Sikora, 2006) describen numerosos métodos, así como otros descriptores que pueden ser usados como huella digital de audio. El espectrograma tradicional se encuentra entre ellos. Este no es más que la descomposición en frecuencias en función del tiempo, presentadas como un gráfico de intensidad, usualmente calculado con la Transformada de Fourier de Tiempo Corto (conocida por sus siglas en inglés, STFT). Uno de los algoritmos que utilizan el espectrograma como base para encontrar similitudes entre ficheros de audio es el de mapa de constelaciones presentado por (Wang, 2003). Dicho algoritmo calcula un

mapa constelación. Un mapa de constelación es un espectrograma en el que se descartan todas las intensidades excepto las que superen un valor de intensidad dado. El objetivo principal es hacer una reducción de la dimensionalidad para luego generar mediante una función de troceo (*hash*, en inglés) posibles correlaciones entre los valores restantes.

El objetivo de este trabajo es valorar la factibilidad y precisión de la aplicación del mapa de constelaciones para el reconocimiento de formas, pero utilizando el espectrograma de Mel como huella digital de audio. Este cambio se sustenta en que la escala Mel, utilizada por este espectrograma, la relación entre frecuencias es logarítmica, no lineal. Aplicado en el contexto del trabajo significa que serán más evidentes las diferencias entre frecuencias distintas que entre aquellas muy cercanas.

El trabajo pretende justificar dicha aplicación debido a que el campo de reconocimiento de audio es un campo bien estudiado y en donde se tienen excelentes resultados. La aplicación en el sitio web Shazam del algoritmo de (Wang, 2003) es un ejemplo de ello. Los resultados experimentales se llevaron a cabo con las bases de datos MPEG-7 CE-Shape-1 Part B y ETU-10. Estas bases de datos contienen formas de color blanco de diferentes objetos ya segmentadas sobre fondo negro. No es objetivo del trabajo tratar con la segmentación y binarización de los objetos de interés, así que cualquier proceso que produzca imágenes como las contenidas en estas bases de datos es aplicable. Esto generalmente ocurre en la etapa de segmentación y/o preprocesamiento durante un proceso de reconocimiento de imágenes. La primera cuenta con 1400 imágenes distribuidas en 70 clases. La segunda tiene 10 clases con 720 imágenes en total. Ambas bases de datos están balanceadas en la cantidad de imágenes por clases.

Dos posibles aplicaciones del sistema que se propone, pueden ser las siguientes: identificación de objetos de interés en mapas, ver (**Fig. 1** – Ejemplos de aplicabilidad. a la izquierda, o identificación de especies vegetales teniendo una muestra de las hojas, ver (**Fig. 1** – Ejemplos de aplicabilidad. a la derecha.

Revista Cubana de Ciencias Informáticas
Vol. 18, No. 3, Julio-Septiembre, 2024
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 48-62

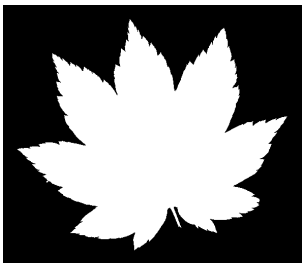

Fig. 1 – Ejemplos de aplicabilidad.

Métodos o Metodología Computacional

A continuación se detallarán las fases de la metodología computacional que se utilizó en el trabajo. Las fases consisten en la transformación a una serie temporal y luego el cálculo de la huella digital de audio usando el espectrogram de Mel y aplicación del mapa de constelación.

A. Conversión a serie temporal

Para la conversión a una serie temporal se utilizó como base el método descrito en (Baez Gonzalez, 2021). El proceso se describe brevemente a continuación. Se comienza por encontrar todos los puntos que componen el borde de la figura de interés. Sus coordenadas serán inicialmente la posición x, y donde aparecen cada uno de los píxeles en la imagen original. Se calcula el centroide de todos los puntos. Tomando como punto inicial el punto más alejado del borde, se procederá su conversión a coordenadas polares, tomándose como la distancia del punto más alejado la de 1 unidad y su ángulo 0°. Los puntos formarán una secuencia ordenada en dirección contraria a las agujas del reloj, comenzando con el más alejado. En el orden en que vayan apareciendo, sus coordenadas serán la distancia al centroide y el ángulo que se forme entre el centroide y el eje 0° (que corresponde con la alineación del punto más alejado). Para evitar una serie de discontinuidades que aparecen, todos los ángulos se expresarán de manera monótona ascendente, reemplazando cada ángulo que no cumple este criterio por uno equivalente con la fórmula $360k + a$, siendo a el ángulo en cuestión y eligiendo el menor k que mantenga la monotonía. Por último, se

hace un submuestreo por interpolación, hacia una cantidad arbitraria de puntos. Esta es la cantidad final de puntos por borde.

El proceso anteriormente descrito es tolerante a traslaciones, rotaciones, escalado, pequeños ruidos en el borde de la figura y parcialmente a reflexiones.

En este paso se realizarán dos modificaciones.

La primera será determinar la cantidad óptima de los puntos por borde, algo que (Baez Gonzalez, 2021) no especifica. Utilizando las *pipelines* descritas en (Brownlee, 2016), se creó un flujo automático para determinar los mejores parámetros para el clasificador k-NN, el porciento de la proporción de prueba y el número de iteraciones estratificadas en la validación cruzada. Para el número de puntos se hizo una búsqueda exhaustiva utilizando las recomendaciones de (scikit-learn developers, 2019). El rango de los parámetros fue el siguiente: para el k-NN de 2 a 5 vecinos, la proporción de prueba entre el 10 y el 33% con saltos del 5% y el número de iteraciones de 5 a 20 según la cantidad máxima de imágenes por clase en cada base de datos en saltos de 5. El k-NN mencionado aquí, es para mantener los parámetros del trabajo citado durante su reproducción al determinar de la cantidad óptima de puntos mencionada.

La segunda fue que la salida final sería directamente a ficheros wav con las siguientes características: audio monocanal, en formato flotante, con 64 bits de precisión y una frecuencia de muestreo igual al número óptimo de puntos por borde determinado en la primera modificación. Lo que haría que el fichero resultante tenga siempre 2 segundos, esto es debido a que la primera mitad del total de muestras contiene las distancias de los puntos que componen el borde al centroide y la otra mitad, los ángulos. Antes de generar el fichero wav resultante, se normalizará por separado las distancias y los ángulos y luego se procederá a su concatenación. El hecho de que la salida sea directamente a ficheros de audio wav es sólo una transformación de dominio de los datos para que las herramientas que proporcionan bibliotecas como librosa (McFee et al., 2015) estén directamente disponibles; en este caso en particular, el espectrograma de Mel. No es necesario tener en cuenta el teorema de muestreo de Nyquist-Shannon por dos motivos. El

primero y más importante, es que a la hora de generar la salida a fichero wav, ya la señal está muestreada satisfactoriamente según los parámetros que se eligieron (en este caso, la cantidad óptima de puntos por borde). El segundo, es que el fichero de audio generado no está pensado para ser reproducido, sino que como se mencionó, es sólo para poder utilizar cualquier herramienta o biblioteca que trabaje directamente con ficheros de audio, por lo tanto no hay que asegurar la reconstrucción de ninguna señal de audio original.

B. Cálculo de la huella digital y el mapa de constelación

Al concluir la fase anterior se procede a calcular la huella digital de cada uno de los archivos wav resultantes utilizando el espectrograma de Mel. Del espectrograma que se obtuvo, se procederá a calcular el 10% de los puntos con mayor valor y que constituirán los puntos que forman el mapa de constelación. La generación de combinación entre los puntos que forman el mapa de constelación se procederá al igual que en (Wang, 2003). El mapa de constelación obtenido es discreto y nunca se tienen más de 16 bandas de Mel que cubren el espectro de frecuencia. Se procedió así para mantener baja la cantidad total de puntos que aparecen en el mapa de constelación. Este número de bandas Mel es totalmente arbitrario y puede cambiarse de ser necesario. Debido a la baja resolución del mapa de constelación, cuando se tiene un punto de anclaje, y se buscan puntos dentro de una zona objetivo, se permite como error de las coordenadas de los puntos en la zona objetivo una 8-vecindad de las coordenadas del punto que se busca en particular.

La (**Fig. 2**) es un ejemplo de onda de audio, espectrograma de Mel (cuya función es servir de huella digital) y el mapa de constelación resultante para la imagen “apple-6” contenida en la base de datos MPEG.

Revista Cubana de Ciencias Informáticas
Vol. 18, No. 3, Julio-Septiembre, 2024
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 48-62

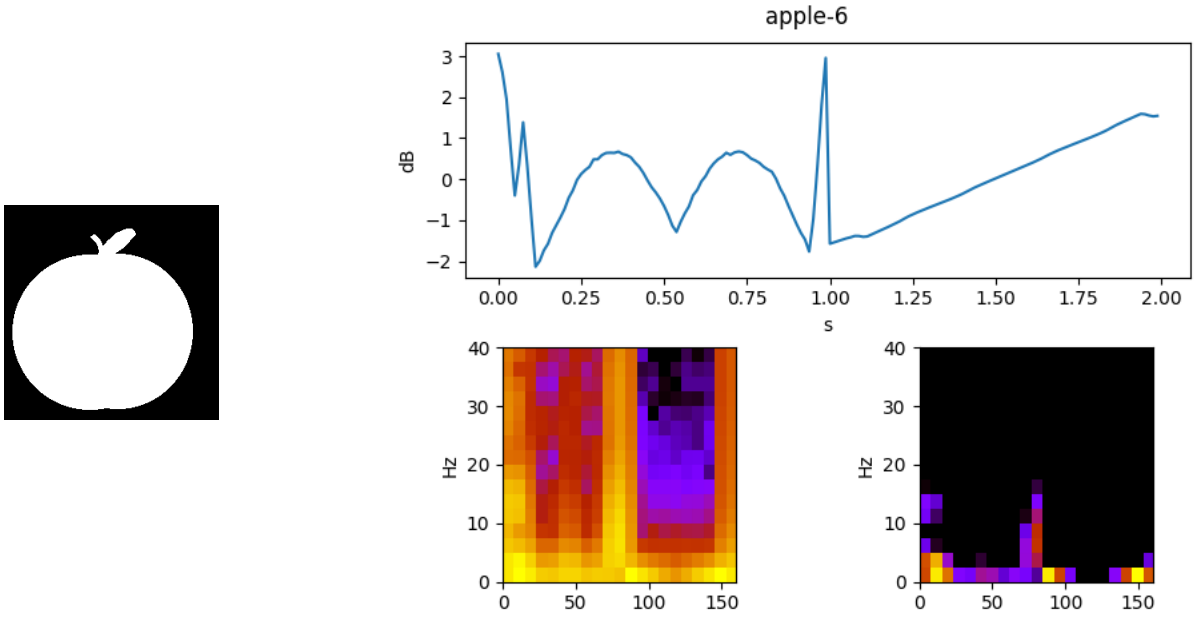

Fig. 2 – Ejemplo de onda de audio, huella digital y mapa de constelación para apple-6.

Resultados y discusión

Ésta sección se divide en 3 subsecciones. La primera, el Marco experimental, detalla la descripción del experimento, la medida utilizada para evaluar el experimento, la configuración de hardware utilizada y el software específico. Los resultados se presentan la subsección de Resultados. La subsección de Análisis y discusión presenta la explicación de los resultados obtenidos y su comparación con trabajos similares.

Marco experimental

El experimento consiste en aplicar la metodología descrita en la sección anterior a 2 bases de datos: MPEG-7 CE-Shape-1 Part B y ETU-10. La primera base de datos se eligió debido a que contiene por cada clase, imágenes similares que presentan transformaciones que incluyen traslación, rotación, escalado, volteado y ruido. En (Dalitz et al., 2013) se menciona que es prácticamente imposible lograr un 100% de etiquetado correcto debido a que algunas imágenes son más parecidas a las de otras clases que a la propia. La base de

datos ETU-10 presenta por cada clase, el mismo objeto, pero la rotación no es sobre la imagen final, sino sobre el objeto antes de hacerle la captura. Esto implica que cada clase es el mismo objeto pero las siluetas cambian para el mismo objeto. Se utilizarán todas las clases en ambas bases de datos. La medida principal a utilizar para evaluar el resultado será la precisión ya que cada clase en ambas bases de datos están balanceadas.

La experimentación se llevó a cabo utilizando una computadora de escritorio con un procesador Core i3-4130 a 3.4 GHz con 2 núcleos y 8GB de memoria RAM.

Dado que el objetivo del trabajo es analizar la precisión de los resultados y no el rendimiento en tiempo de ejecución, el lenguaje utilizado para la implementación fue Python en su versión 3.9.2. Las bibliotecas principales utilizadas fueron: librosa para el cálculo de la huella digital de audio y matplotlib para la visualización de los resultados.

Resultados

Los resultados de determinar la cantidad de puntos óptima y el rango de los parámetros fueron los siguientes: cantidad de puntos óptima, 80 para la base de datos MPEG y 32 para la ETU; la cantidad de vecinos del k-NN 2, para ambas bases de datos; la proporción de prueba del 33% para la MPEG y 10% para ETU y 20 iteraciones para la MPEG y 5 para la ETU.

Los resultados de precisión fueron: 100% para la base de datos ETU y 97.92% para la MPEG. Las clases que presentaron fallos en la base MPEG fueron: fish, camel, elephant, octopus y guitar. Siguiendo el orden mencionado, la cantidad de fallos respectivos fueron: 8, 7, 6, 6 y 2, para un total de 29 fallos en total.

Revista Cubana de Ciencias Informáticas
Vol. 18, No. 3, Julio-Septiembre, 2024
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 48-62

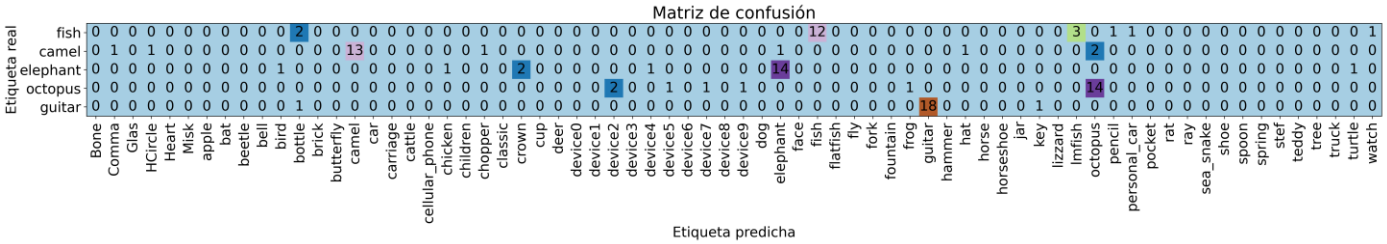

Fig. 3 – Clases con fallos en MPEG.

Análisis y discusión

De los resultados obtenidos, se puede explicar en el caso de la clase fish, por qué ocurren tantos fallos. Si dos clases tienen distintos nombres, pero tienen objetos con forma similar, es de esperar que haya fallos en la clasificación entre ambas. Se pondrá de ejemplo las imágenes de fish-15 y lmfish-17 para ilustrar este punto.

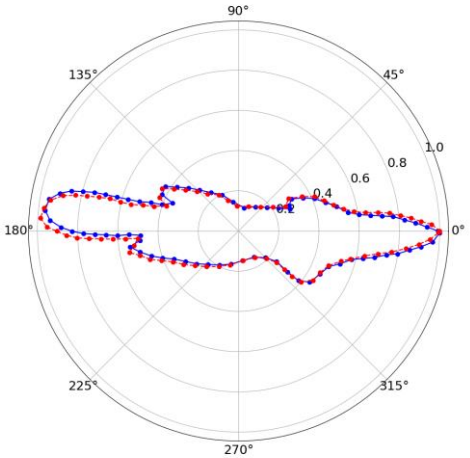

Fig. 4 – Comparación en coordenadas polares de fish-15 y lmfish-17.

Revista Cubana de Ciencias Informáticas
Vol. 18, No. 3, Julio-Septiembre, 2024
ISSN: 2227-1899 | RNPS: 2301
http://rcci.uci.cu
Pág. 48-62

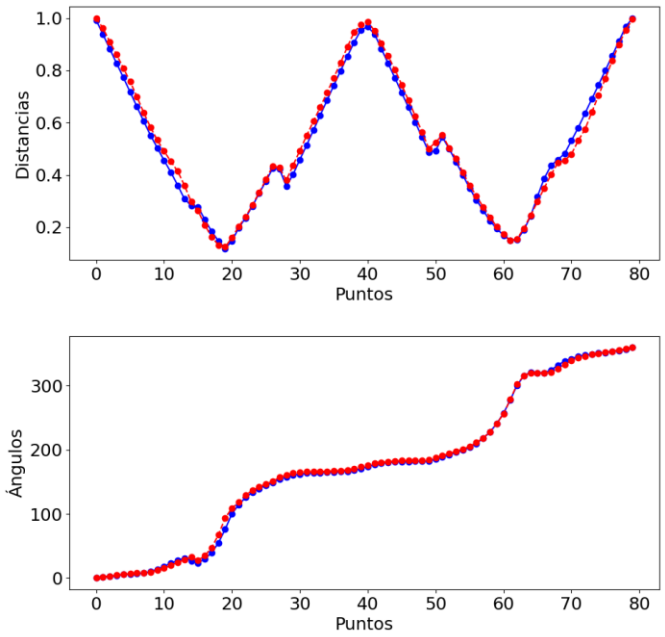

Fig. 5 – Distancias y ángulos que componen el fichero wav de fish-15 y lmfish-17.

Al ver las 2 figuras anteriores, es evidente que ambas figuras son casi idénticas. Cabría esperar entonces que ocurran fallos tanto durante el entrenamiento como en las pruebas, ya que se ha etiquetado a la misma figura en 2 clases distintas. Este constituye un ejemplo de lo mencionado por (Dalitz et al., 2013).

En lo referido a la comparación de resultados, el presente trabajo mejora los resultados obtenidos por (Gual-Arnau, Herold-García y Simó, 2013) con un 95% de precisión, el de (Goshvarpour, Ebrahimnezhad y Goshvarpour, 2013) con un 92.5% de precisión, el de (Abbas et al., 2019) con un 92.45% de precisión y el de (Baez Gonzalez, 2021) con un 92.86% de precisión. De los trabajos mencionados antes, la comparación debiera realizarse con (Abbas et al., 2019) y (Baez Gonzalez, 2021) utilizando la precisión como métrica, pues sólo estos utilizan el total de las 70 clases. Los otros 2 trabajos utilizan un subconjunto de las 70 clases. En el caso de (Gual-Arnau, Herold-García y Simó, 2013) 11 clases y para (Goshvarpour, Ebrahimnezhad y Goshvarpour, 2013) 8 clases. Se desestimó la comparación con los siguientes trabajos, pues estos utilizan como métrica de rendimiento la puntuación *bull's eye* (tiro al blanco, en español): (Yang, Koknar-Tezel y

Latecki, 2009), (Temlyakov et al. 2010) y (Pan, Chachada y Kuo, 2016). Estos trabajos obtienen, en el orden mencionado, puntuaciones del 93.32%, 95.60% y 100%.

Lo novedoso del presente trabajo es la aplicación de técnicas relacionadas con el audio para el reconocimiento de formas. La propuesta presentada tiene una precisión significativa, al obtener un resultado superior al 95%.

Conclusiones

Teniendo en cuenta los resultados obtenidos y habiéndose realizado un análisis y discusión de estos, se concluye que es factible utilizar el espectrograma de Mel como huella digital de audio, al aplicarle el mapa de constelaciones para lograr el reconocimiento de formas en imágenes. En ambas bases de datos los resultados son significativos teniendo en cuenta el uso de la precisión como métrica para evaluar los resultados. El presente trabajo mejora todos los resultados citados por (Baez Gonzalez, 2021), incluyéndolo. El presente trabajo puede utilizarse como base para probar otras técnicas de audio al reconocimiento de formas siempre que estas pueden ser transformadas a series temporales. Un posible trabajo futuro basado en el actual, pudiera consistir en evaluar el rendimiento en tiempo de ejecución de la propuesta presentada con implementaciones en Python y C++. Otro trabajo derivado pudiera ser la determinación de que tan escalable es la propuesta presentada con una cantidad de al menos 10 000 muestras o mayor.

Referencias

Abbas, S., Farhan, S., Fahiem, M.A. Y Tauseef, H., 2019. Efficient Shape Classification Using Zernike Moments And Geometrical Features On Mpeg-7 Dataset. Advances In Electrical And Computer Engineering, Vol. 19, No. 1,

Agarwaal, A., Kanaujia, P., Roy, S.S. Y Ghose, S., 2023. Robust And Lightweight Audio Fingerprint For Automatic Content Recognition. Arxiv Preprint Arxiv:2305.09559,

Akesbi, K., Desblancs, D. Y Martin, B., 2023. Music Augmentation And Denoising For Peak-Based Audio Fingerprinting. Arxiv Preprint Arxiv:2310.13388,

Ansari, M.I. Y Hasan, T., 2022. Spectnet: End-To-End Audio Signal Classification Using Learnable Spectrograms. Arxiv Preprint Arxiv:2211.09352,

Baez Gonzalez, L.R., 2021. Reconocimiento De Formas Mediante Transformada De Dominio. . Bayamo: S.N., Pp. 53.

Brownlee, J., 2016. Machine Learning Mastery With Python: Understand Your Data, Create Accurate Models, And Work Projects End-To-End. S.L.: Machine Learning Mastery.

Cetin, O., 2023. Accent Recognition Using A Spectrogram Image Feature-Based Convolutional Neural Network. Arabian Journal For Science And Engineering, Vol. 48, No. 2,

Chakravarty, N. Y Dua, M., 2024. Feature Extraction Using Gtcc Spectrogram And Resnet50 Based Classification For Audio Spoof Detection. International Journal Of Speech Technology,

Ciric, D.G., Peric, Z.H., Milenkovic, M. Y Vucic, N.J., 2022. Evaluating Similarity Of Spectrogram-Like Images Of Dc Motor Sounds By Pearson Correlation Coefficient. Elektronika Ir Elektrotechnika, Vol. 28, No. 3,

Dalitz, C., Brandt, C., Goebbels, S. Y Kolanus, D., 2013. Fourier Descriptors For Broken Shapes. Eurasip Journal On Advances In Signal Processing, Vol. 2013,

Developers, Scikit-Learn, 2019. Scikit-Learn User Guide Release 0.21.2. Mayo 2019. S.L.: S.N.

Franzoni, V. Y Otros, 2023. Cross-Domain Synergy: Leveraging Image Processing Techniques For Enhanced Sound Classification Through Spectrogram Analysis Using Cnns. Journal Of Autonomous Intelligence, Vol. 6, No. 3,

Gao, Y., Vuong, T., Elyasi, M., Bharaj, G. Y Singh, R., 2021. Generalized Spoofing Detection Inspired From Audio Generation Artifacts. Arxiv Preprint Arxiv:2104.04111,

Goshvarpour, Atefeh, Ebrahimnezhad, H. Y Goshvarpour, Ateke, 2013. Shape Classification Based On Normalized Distance And Angle Histograms Using Pnn. Journal Of Information Technology And Computer Science,

Gual-Arnau, X., Herold-García, S. Y Simó, A., 2013. Shape Description From Generalized Support Functions. Pattern Recognition Letters, Vol. 34, No. 6,

He, P., Li, Y., Chen, S., Xu, H., Zhu, L. Y Wang, L., 2021. Core Looseness Fault Identification Model Based On Mel Spectrogram-Cnn. Journal Of Physics: Conference Series. S.L.: Iop Publishing, Pp. 012060. Vol. 2137.

Kamuni, N., Chintala, S., Kunchakuri, N., Narasimharaju, J.S.A. Y Kumar, V., 2024. Advancing Audio Fingerprinting Accuracy Addressing Background Noise And Distortion Challenges. Arxiv Preprint Arxiv:2402.13957,

Kim, H.-G., Moreau, N. Y Sikora, T., 2006. Mpeg-7 Audio And Beyond: Audio Content Indexing And Retrieval. S.L.: John Wiley & Sons.

Kishor, K., Venkatesh, S. Y Koolagudi, S.G., 2023. Audio Fingerprinting System To Detect And Match Audio Recordings. International Conference On Pattern Recognition And Machine Intelligence. S.L.: Springer, Pp. 683-690.

Mcfee, B., Raffel, C., Liang, D., Ellis, D.P., Mcvicar, M., Battenberg, E. Y Nieto, O., 2015. Librosa: Audio And Music Signal Analysis In Python. Scipy. S.L.: S.N., Pp. 18-24.

Mingqiang, Y., Kidiyo, K., Joseph, R., Y Others, 2008. A Survey Of Shape Feature Extraction Techniques. Pattern Recognition, Vol. 15, No. 7,

Oo, M.M. Y Oo, L.L., 2020. Fusion Of Log-Mel Spectrogram And Glcm Feature In Acoustic Scene Classification. Software Engineering Research, Management And Applications,

Pan, X., Chachada, S. Y Kuo, C.-C.J., 2016. A Two-Stage Shape Retrieval (Tsr) Method With Global And Local Features. Journal Of Visual Communication And Image Representation, Vol. 38,

Serrano, S., Sahbudin, M.A.B., Chaouch, C. Y Scarpa, M., 2022. A New Fingerprint Definition For Effective Song Recognition. Pattern Recognition Letters, Vol. 160,

Temlyakov, A., Munsell, B.C., Waggoner, J.W. Y Wang, S., 2010. Two Perceptually Motivated Strategies For Shape Classification. 2010 Ieee Computer Society Conference On Computer Vision And Pattern Recognition. S.L.: Ieee, Pp. 2289-2296.

Ustubioglu, A., Ustubioglu, B. Y Ulutas, G., 2023. Mel Spectrogram-Based Audio Forgery Detection Using Cnn. Signal, Image And Video Processing, Vol. 17, No. 5,

Virginia Kindergarten Readiness Program, 2019. Shape Recognition And Properties [En Línea]. Marzo 2019. S.L.: S.N. [Consulta: 1 Mayo 2024]. Disponible En: Https://Vkrponline.Org/Wp-Content/Uploads/Sites/3/2020/03/Vkrp_Shape_Rec_Prop_K.Pdf.

Wang, A., 2003. An Industrial Strength Audio Search Algorithm. Ismir. S.L.: Washington, Dc, Pp. 7-13. Vol. 2003.

Yang, X., Koknar-Tezel, S. Y Latecki, L.J., 2009. Locally Constrained Diffusion Process On Locally Densified Distance Spaces With Applications To Shape Retrieval. 2009 Ieee Conference On Computer Vision And Pattern Recognition. S.L.: Ieee, Pp. 357-364.

Conflicto de interés

El autor autoriza la distribución y uso de su artículo.

Contribuciones de los autores

Conceptualización: Luis Rafael Baez Gonzalez

Curación de datos: Luis Rafael Baez Gonzalez

Análisis formal: Luis Rafael Baez Gonzalez

Adquisición de fondos: -

Investigación: Luis Rafael Baez Gonzalez

Metodología: Luis Rafael Baez Gonzalez

Administración del proyecto: Luis Rafael Baez Gonzalez

Recursos: Luis Rafael Baez Gonzalez

Software: Luis Rafael Baez Gonzalez

Supervisión: Luis Rafael Baez Gonzalez

Validación: Luis Rafael Baez Gonzalez

Visualización: Luis Rafael Baez Gonzalez

Redacción – borrador original: Luis Rafael Baez Gonzalez

Redacción – revisión y edición: Luis Rafael Baez Gonzalez