

Aprendizaje multietiqueta de patrones geométricos en objetos de herencia cultural

Profesor guía: Benjamín Bustos - bebustos@dcc.uchile.cl

Profesor coguía: Iván Sipirán - isipiran@dcc.uchile.cl

Alumno: Matías Vergara - matiasvergara@ug.uchile.cl

5 de julio de 2022

Abstract

Una tarea frecuente del dominio de la Arqueología es el etiquetado de patrones geométricos. En ella, se estudian y describen objetos de herencia cultural en búsqueda de extraer toda la información relevante y así obtener un mejor entendimiento de los contextos socioculturales, económicos y/o conductuales de la cultura a la cual pertenecen. Para ello, expertos del área analizan la superficie de los artefactos en búsqueda de patrones, los cuales describen mediante la asignación de “etiquetas” relativas a su contenido. Dicho proceso presenta dos grandes desafíos: por un lado, el representar una tarea tediosa en donde los expertos pueden invertir horas y no dar con un trabajo completo y, por otro, la falta de estándares al momento de escoger etiquetas a utilizar. La presente propuesta de tesis busca abordar este problema a través del desarrollo de una herramienta de apoyo al etiquetado que, a partir de un patrón geométrico, entregue un conjunto de etiquetas relevantes como sugerencia al experto, contribuyendo así a la simplificación del proceso al reducir la cantidad de etiquetas a idear y a la formación de un estándar mediante sugerencias de etiquetas de una misma taxonomía. Para ello se propone un enfoque de aprendizaje supervisado multietiqueta, considerando en primera instancia los métodos tradicionales y posteriormente aquellos de aprendizaje profundo. A la fecha, los experimentos del enfoque tradicional han permitido validar las principales hipótesis necesarias para lograr el objetivo y concluir a su vez que el camino natural a seguir es aquel de aprendizaje profundo, trabajo propuesto para Tesis II.

Aprobado por:

Benjamín Bustos, Ph.D.

Profesor guía

Aprobado por:

Iván Sipirán, Ph.D.

Profesor coguía

1. Introducción

El presente documento versa sobre la propuesta de tesis de Matías Vergara, estudiante de Ingeniería Civil en Computación y del Magíster en Ciencia de Datos. El tema lleva por nombre “Aprendizaje multietiqueta de patrones geométricos en objetos de herencia cultural” y se desarrolla con los profesores Benjamín Bustos e Iván Sipirán como guía y coguía, respectivamente.

El área de estudio a abordar en esta propuesta corresponde a la Arqueología, ciencia que busca estudiar la diversidad humana a través del registro material. En particular, el problema a tratar tiene por contexto el estudio y catalogación de *herencia cultural*, definida como aquel patrimonio de bienes tangibles o intangibles de un grupo o sociedad que se hereda de generaciones pasadas [1]. Dicho dominio juega un rol fundamental no solo en arqueología, si no también en ciencias como la paleontología, geología y antropología.

Más específicamente, la investigación se centra en los bienes tangibles, los cuales reciben la denominación de *objetos de herencia cultural*. Definidos como cualquier bien mueble de importancia cultural que requiera protección - trátase de artefactos arqueológicos, especímenes paleontológicos/geológicos, meteoritos o cualquier otro objeto con un significado cultural [2] -, dichos objetos son materia importante de descripción, en cuanto su estudio permite alcanzar un mayor entendimiento de la cultura que les subyace. Dicho proceso, sin embargo, acostumbra requerir de múltiples esfuerzos e iteraciones [3], en cuanto incluye diversos desafíos que obstaculizan la catalogación y descripción de los artefactos, dificultando de esta manera la comprensión de sus contextos socioculturales, económicos y/o conductuales [4], así como las labores de conservación y *musealización*.

Son estos obstáculos y la consecuente pérdida de información valiosa los que motivan al desarrollo de esta investigación. En concreto, se busca abordar aquellas dificultades presentes al momento de describir artefactos cuya superficie incorpora patrones geométricos, proceso que, por la forma en que se realiza - donde un arqueólogo experto asigna distintas “etiquetas” a cada figura - recibe el nombre de *etiquetado de patrones geométricos*. A modo de ejemplo, la Figura 1 presenta un patrón (imagen en alto) y sus etiquetas (texto bajo la imagen).

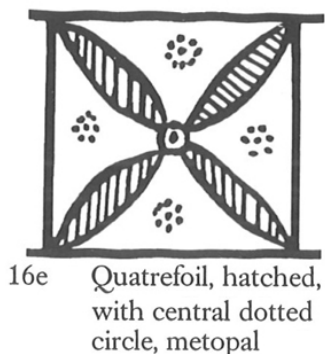


Figura 1: Ejemplo de patrón etiquetado [5]

Aún más, los objetos de herencia cultural presentan una característica de particular interés para el estudio. Se trata de su histórica simbiosis con el dominio de la computación gráfica y análisis tridimensional [6], lo cual ha demostrado que no solo se trata de un área abierta a la inclusión de herramientas de apoyo en forma de *software*, si no que además guarda cautivadores desafíos cuyo estudio desde un dominio específico no hace si no contribuir al avance del mismo.

Es así como el fin último de esta investigación se configura no solo como el desarrollo de una herramienta de apoyo, si no también como la ambiciosa simbiosis entre la Arqueología, la Computación y la Ciencia de Datos.

2. Problema a abordar

El etiquetado geométrico de patrones presenta dos grandes dificultades que hacen de él un proceso tedioso, a menudo incompleto y carente de estándares. Estas se encuentran, por un lado, en la naturaleza compleja de la tarea misma, donde la descripción de un patrón puede ser tan amplia como determine la motivación del experto etiquetador y, por el otro, en la falta de acuerdo entre expertos. Ambas razones se discutirán a continuación.

En primer lugar, la complejidad del proceso se debe principalmente a la gran cantidad de información geométrica que un patrón puede contener, dando así lugar a un vasto universo de etiquetas posibles. Esto se traduce en un trabajo extensivo y eventualmente incompleto, en el cual los expertos pueden invertir horas buscando asignar todas las etiquetas necesarias y aún así no lograr representar todas las características relevantes del patrón.

Aún más, esta búsqueda de las características “relevantes” de un patrón es de por sí subjetiva, en cuanto lo relevante para un experto no tiene por qué serlo para otro. Es ahí donde nace la falta de acuerdo, la cual encuentra sus orígenes en una diferencia de trasfondo que a menudo da lugar a variaciones en la consistencia de los datos [7]. Por ejemplo, un patrón podría recibir las etiquetas *ave*, *vertical* y *metopa* por un experto *A*, mientras que un experto *B* podría asignarle las etiquetas *pájaro* y *vertical*, evidenciando así dos problemas: la ausencia de la etiqueta *metopa* en el etiquetado de *B* y el uso de sinónimos para referirse a la misma característica (*pájaro* y *ave*).

Dichas dificultades han servido de motivación para que los expertos busquen, a través de su obra, sentar una taxonomía base sobre la cual sus colegas puedan guiarse. Uno de los intentos más famosos es el libro *Ornamente Geometrischer Vasen: Ein Kompendium* [5], autoría del prof. Dr. Norbert Kunisch, arqueólogo de renombre mundial que buscaba establecer una taxonomía de etiquetas a través de ejemplos provenientes de excavaciones en Grecia durante el siglo XX. Dicho libro lograría reconocimiento en el área, más no habría alcanzado a establecerse como una referencia del etiquetado. Una posible razón para ello se encuentra en la misma naturaleza tediosa del proceso, donde el incorporar una tarea adicional - tal como la revisión de la obra de un tercero - no hace más que disminuir la motivación del experto por alcanzar un trabajo completo [7].

Un punto importante a mencionar es que, si bien en este caso se busca trabajar en específico con el etiquetado de patrones geométricos, dichas dificultades son transversales a otros procesos de etiquetado en arqueología y han despertado esfuerzos desde otros dominios. Un ejemplo se encuentra en el etiquetado de estilos de vasijas de herencia cultural, donde se desarrollase una herramienta que buscaría reducir la probabilidad de observar etiquetados incompletos a través de aumentar la motivación del etiquetador al proponer el proceso como parte de un juego, tratándose entonces de un enfoque de *gamificación*. Si bien dicha herramienta habría logrado cumplir su objetivo, el estudio de sus resultados indicaría que también habría aumentado el nivel de desacuerdo entre expertos [7].

Los esfuerzos identificados representan un buen insumo para la investigación, en cuanto demuestran que la problemática es conocida por el área y que requiere una solución. Este trabajo buscará entonces alcanzar dicha solución desde un nuevo enfoque, a través del desarrollo de una herramienta de apoyo que sugiera etiquetas relevantes al etiquetador.

3. Solución propuesta

Se propone abordar el problema identificado desde un enfoque de aprendizaje supervisado, donde se buscará entrenar modelos en la tarea de etiquetar patrones geométricos y así posteriormente utilizar su salida como conjunto de etiquetas sugeridas. De esta manera, se buscará (1) reducir la dificultad del proceso al disminuir la cantidad de etiquetas a idear por el experto y (2) contribuir al alcance de un estándar al sugerir etiquetas de una misma taxonomía.

La elección de dicho enfoque se basa en la inherente pertenencia del problema a la familia de desafíos de *clasificación multietiqueta* (MLC), en donde la tarea consiste no en predecir una clase $l \in L$ si no más bien en predecir un subconjunto $Z \subseteq L$, donde L tal que $|L| > 1$ representa el conjunto de todas las etiquetas posibles. Dicho enfoque se opone a los planteamientos clásicos de clasificación singular como el caso binario o multiclase, ofreciendo una mejor adaptación a la naturaleza misma de una gran variedad de tareas que van desde la clasificación de funciones de proteínas hasta la categorización de imágenes. Por ejemplo, a una imagen de playa podrían corresponderle las etiquetas *atardecer*, *arena* y *mar*, en lugar de solo *playa*. Desde luego, la opción múltiple es mucho más informativa.

De esta manera, la solución propuesta lleva por hipótesis central la capacidad de reproducir, mediante técnicas de aprendizaje multietiqueta y un conjunto de datos apropiado, el etiquetado por expertos. No se pretende, sin embargo, que dicho resultado sea perfecto, pues la meta de la investigación no es el reemplazo del personal si no más bien el desarrollo de una herramienta de apoyo, la cual podrá sentar los precedentes para una futura automatización total del proceso.

Más en detalle, la solución propuesta consta de tres grandes etapas, siendo la primera de ellas la conformación de un conjunto de datos y su correspondiente análisis exploratorio y procesamiento. Para ello se toma como restricción el uso de los patrones disponibles en el libro *Ornamente Geometrischer Vasen: Ein Kompendium*, el cual cuenta con 776 patrones previamente etiquetados cuya imagen digitalizada es proporcionada como insumo por los profesores guía. Dicho conjunto será entonces analizado en búsqueda de sus características principales, con particular atención en aquellas que la literatura ha demostrado juegan un rol fundamental en la capacidad de aprendizaje de los modelos, tales como la distribución de las etiquetas [8] y el valor de métricas como *label average* (LA) [9] y *label density* (LD) [10].

La segunda etapa, por otro lado, consistirá en la búsqueda de modelos desde un enfoque de aprendizaje multietiqueta tradicional, utilizando para ello algoritmos basados en representaciones previamente construidas. En este sentido, se realizarán experimentos tanto con técnicas de *transformación del problema* - las cuales buscan convertir la tarea multietiqueta en muchas tareas singulares [11] - como con aquellas de *adaptación del algoritmo*, buscando encontrar la configuración que mejor se adapte al problema según un conjunto de métricas cuidadosamente seleccionadas, de las cuales *Hamming Loss* (HL) [9] se considerará como métrica principal.

Por último, la tercera etapa tendrá por fin la búsqueda de mejores resultados mediante el desarrollo y aplicación de modelos de aprendizaje profundo, los cuales se espera superen aquellos del enfoque tradicional gracias a la capacidad de incorporar el aprendizaje de representaciones en el entrenamiento [8]. Se buscará experimentar con técnicas del estado del arte tales como *Binary Relevance Neural Networks* (BRNN) y *Threshold Dependent Neural Networks* (TDNN) [12], además de arquitecturas propias basadas en la intuición obtenida en los experimentos previos.

4. Resultados preliminares

El trabajo realizado inició con la extracción de las etiquetas de cada patrón desde el libro escogido, proceso que se realizó de forma manual. Posteriormente se estudió cómo estas se relacionaban con los patrones en estudio, identificándose una distribución sumamente *skew* y un *label density* (LD) de 0.005, valor que representaría un obstáculo para el aprendizaje por ser demasiado bajo. Esto sirvió de motivación para realizar una *normalización* de las etiquetas, proceso a través del cual se aplicaría *lemmatization* y remoción de *stopwords* para quitar duplicados y separar etiquetas compuestas, dando así lugar a un conjunto refinado de datos con 340 etiquetas en lugar de las 586 originales y una LD de 0.015, tres veces mayor que la inicial.

Pese a que lo anterior significaría un progreso en cuanto al potencial de las etiquetas de ser “aprendidas” por un modelo MLC, el LD se mantenía aún muy bajo. Esto motivaría a realizar los experimentos no solo desde un enfoque método contra método, si no también aplicando un podado del problema mediante un umbral de frecuencias t . Dicho umbral permitiría simplificar el problema al mantener solamente las etiquetas con al menos t apariciones en los datos, sin embargo, debía aplicarse con cuidado de no caer en una trivialización tal que la solución pierda valor. Se decidió entonces mantener dicho umbral como una variable en los experimentos, donde se compararían los métodos sobre distintos valores de t y, de esta forma, se observaría también como distintos métodos se adaptan mejor a distintos escenarios.

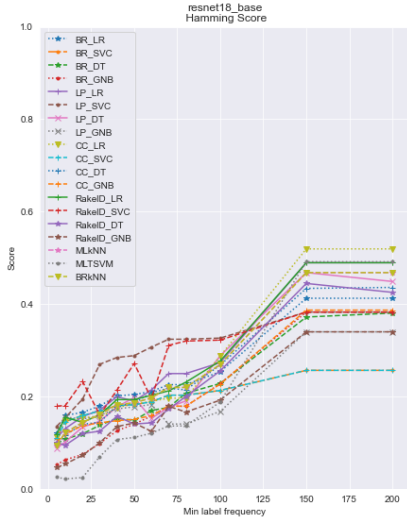


Figura 2: HS al variar t . Enfoque tradicional, descriptores ResNet18.

De esta manera, se procedería a realizar los primeros experimentos del enfoque tradicional, utilizando para ello descriptores obtenidos a partir de arquitecturas ResNet18 y ResNet50 entrenadas para la tarea de predecir el capítulo del libro al cual cada patrón pertenece, dentro de 5 categorías. Los resultados de dichos experimentos se presentan en la Figura 2.

Los resultados, aunque insuficientes, validarían la hipótesis central: aún en el desfavorable contexto de utilizar representaciones obtenidas para una tarea diferente, los métodos tradicionales eran capaces de aprender y asignar ciertas etiquetas correctamente. Aún más, experimentos posteriores confirmarían que la baja calidad de los resultados se debería en gran medida a este fenómeno, en cuanto descriptores obtenidos a partir de entrenamientos parciales de las arquitecturas ResNet ofrecerían mejores resultados, evidenciando así que la especificidad de la red generadora estaría perjudicando los resultados en MLC.

Es así como, tras haber explorado ampliamente los métodos tradicionales (y habiendo realizado, por ejemplo, experimentos con múltiples técnicas de *data augmentation*), se llegaría a la conclusión de que para obtener mejores resultados sería necesario cambiar a un paradigma End-To-End, donde el aprendizaje de representaciones fuese parte del entrenamiento. Esto coincidiría con el camino recomendado por la literatura, llegando de esta manera al cierre del estudio del enfoque tradicional para, tras haber validado la hipótesis central, pasar ahora a la experimentación desde un enfoque de aprendizaje profundo.

5. Conclusiones & Trabajo Futuro

Tras haber concluido exitosamente las etapas primera y segunda de la Solución Propuesta y habiendo logrado validar la hipótesis central a través de resultados preliminares, se confirma entonces la presente investigación como propuesta de Tesis para optar al grado de Magíster en Ciencia de Datos.

Así mismo, los hallazgos realizados permiten definir el siguiente Plan de Trabajo para el curso de Tesis II, en caso de ser aceptada la propuesta:

- Dar inicio al trabajo con un estudio más en profundidad de las técnicas de clasificación multietiqueta mediante aprendizaje profundo. En particular, estudiar conceptos tales como el *Challenging MLC* y *Extreme MLC* [8], categorías del área que, si bien no engloban el problema en cuestión, podrían aportar con un mayor entendimiento del desafío que el aprendizaje multietiqueta significa y así contribuir a la formación de nuevas ideas para los experimentos desde este nuevo enfoque.
- Realizar experimentos preliminares con las arquitecturas más ligeras del estado del arte, a fin de dimensionar la factibilidad de utilizarlas con (1) la cantidad de ejemplos con los que se cuenta y (2) el *hardware* disponible. En este sentido, se considera que modelos tales como *Canonical Correlated Autoencoder (C2AE)* [13] y arquitecturas basadas en redes *feed-forward* sobre redes recurrentes (*RNN*) representan un buen punto de partida.
- Los próximos pasos tendrán relación con el ajuste de hiperparámetros para encontrar aquellos más adecuados al problema, así como con la elección y/o definición de una función de pérdida apta para el desafío. En este punto cobrará especial relevancia el hecho de que los datos multietiqueta no admiten las técnicas de balanceo tradicionales - tales como el *subsampling* u *oversampling* - dada la presencia conjunta de las etiquetas en un mismo patrón, por lo cual de presentarse dicha necesidad, deberá ser satisfecha a través de la función de pérdida y/o técnicas más complejas, tales como un aprendizaje con pesos.
- Por último, el resultado de las etapas anteriores determinará la factibilidad de realizar pruebas con arquitecturas más profundas, para lo cual sin embargo se presume será necesario invertir esfuerzos en crear una cantidad considerable de datos sintéticos. Para ello se considera que ya hay gran parte del camino avanzado, en cuanto la exploración de técnicas de aumentado de imágenes realizada para el enfoque tradicional aporta una gran base sobre la cual trabajar.
- Finalmente, restará discutir la relación eficiencia - eficacia de los modelos alcanzados, en cuanto la herramienta de apoyo a concebir deberá no solo ser buena en cuanto a MLC se refiere, si no además realizar las sugerencias en un tiempo razonable (si un etiquetador experto debe esperar horas por las sugerencias de un único patrón, entonces la herramienta se vería limitada a un uso *off-line*, lo cual a priori no es lo que se busca).

Una vez llevados a cabo los pasos anteriores, la investigación concluirá con un análisis de los principales hallazgos y la evaluación de la herramienta de apoyo concebida. Aún más, se buscará que dicho análisis permita no solo comprender el porqué de los resultados encontrados, si no también sentar los lineamientos necesarios para una futura investigación en búsqueda del automatizado total del proceso.

Referencias

- [1] Silverman, H., Ruggles, D. F., y S., L. W., Closing Pandora's Box: Human Rights Conundrums in Cultural Heritage. Springer Science+Business Media, 2007.
- [2] NHCN, "Heritage objects: The national heritage council of namibia," 2005, <https://www.nhc-nam.org/heritage-objects>.
- [3] Castillo Luján, F., "Tipología y seriación de la cerámica proveniente del cementerio chimú de huaca de la luna, Perú," Boletín del Museo Chileno de Arte Precolombino, vol. 23, pp. 27 – 58, 2018, http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-68942018000300027&nrm=iso.
- [4] Biasotti, S., Moscoso Thompson, E., Bathe, L., Berretti, S., Giachetti, A., Lejembale, T., Mellado, N., Moustakas, K., Manolas, I., Dimou, D., Tortorici, C., Velasco-Forero, S., Werghi, N., Polig, M., Sorrentino, G., y Hermon, S., "Shrec'18 track: Recognition of geometric patterns over 3d models," 2018.
- [5] Kunisch, N., "Ornamente geometrischer vassen: Ein kompendium," 1998.
- [6] Sipiran, I., Lazo, P., Lopez, C., Jimenez, M., Bagewadi, N., Bustos, B., Dao, H., Gangisetty, S., Hanik, M., Ho-Thi, N., Holenderski, M., Jarnikov, D., Labrada, A., Lengauer, S., Licandro, R., Nguyen, D., Nguyen-Ho, T., Rey, L. A. P., Pham, B., Preiner, R., Schreck, T., Trinh, Q., Tonnaer, L., von Tycowicz, C., y Vu-Le, T., "SHREC 2021: Retrieval of cultural heritage objects," Comput. Graph., vol. 100, pp. 1–20, 2021, [doi:10.1016/j.cag.2021.07.010](https://doi.org/10.1016/j.cag.2021.07.010).
- [7] Lee, J., Yi, J. H., y Kim, S., "Cultural heritage design element labeling system with gamification," IEEE Access, vol. 8, pp. 127700–127708, 2020.
- [8] Liu W., Wang H., Shen X., Tsang I. "The Emerging Trends of Multi-Label Learning". To appear in IEEE Transactions on Pattern Analysis Machine Intelligence. DOI: 10.1109/T-PAMI.2021.3119334.
- [9] Tsoumakas, G. y Katakis, I., "Multi-label classification: An overview," Int. J. Data Warehous. Min., vol. 3, no. 3, pp. 1–13, 2007, [doi:10.4018/jdwm.2007070101](https://doi.org/10.4018/jdwm.2007070101).
- [10] Blanco, A., Casillas, A., Pérez, A., y de Ilarraza, A. D., "Multi-label clinical document classification: Impact of label-density," Expert Syst. Appl., vol. 138, 2019.
- [11] Tsoumakas, G., Katakis, I., y Vlahavas, I., "I.: A review of multi-label classification methods," en In: Proceedings of the 2nd ADBIS Workshop on Data Mining and Knowledge Discovery (ADMKD), pp. 99–109, 2006.
- [12] He, H. y Xia, R., "Joint binary neural network for multi-label learning with applications to emotion classification," en NLPCC (1), vol. 11108 de Lecture Notes in Computer Science, pp. 250–259, Springer, 2018.
- [13] Yeh, C.-K., Wu, W.-C., Ko, W.-J., y Wang, Y.-C. F., "Learning deep latent spaces for multi-label classification," 2017.