

Trademark Image Retrieval Using Deep Learning

Freddy Pesantez
freddy.pesantez@uchile.cl

Daniel Baeza
dabaeza@uchile.cl

Abstract—El reconocimiento de Imágenes de Marcas es un tema de mucho interés actualmente. Esto es debido al fuerte crecimiento de registros de logo y los conflictos de propiedad que pueden ocasionar dos logos similares. A pesar de haber varios intentos por dar solución a este problema, no existen metodologías computacionales que logren superar o igualar al análisis de un experto. En este documento se resumen algunos avances sobre el tema.

Index Terms—Reconocimiento de imágenes, Imágenes de marcas.

I. INTRODUCTION

En el mundo actual es muy común que las empresas utilicen algún tipo de imagen llamativa para su identificación en el mercado, de forma tal que les permita ser reconocidos fácilmente en el mercado. Estas imágenes tienen mucho valor porque les permite posicionarse de alguna manera en las preferencias de los usuarios, por lo tanto el uso de ellas constituye en factor importante en un ambiente de mercado.

Sin embargo, no es difícil imaginarse que el número actual de estas imágenes de las marcas se ha ido incrementado enormemente con el tiempo, y las bases de datos actuales son muy grandes. Por lo mencionado antes, estas imágenes están sujetas a derechos de autor en los países.

iiiiiii HEAD Para el registro de una nueva imagen se debe verificar que no existan otras imágenes lo suficientemente parecidas como para que exista confusión por un lado y para que no se violen los derechos de autor. ===== Sin embargo, el número actual de logos se ha ido incrementado enormemente con el tiempo, haciendo mas difícil el trabajo de los expertos en las oficinas de registro.

Por esta razones, en visión computacional, el reconocimiento de marcas ha tenido un creciente interés. En la actualidad no existen metodologías que tengan tasas de clasificación superiores a la de un experto, pero si han contribuido a facilitar el trabajo de estas personas.

En este trabajo se propone analizar el desempeño de tres de los métodos mas utilizados actualmente para el reconocimiento de marcas. El primero intenta encontrar patrones locales en los logos que permitan su discriminación. El segundo se basa en en el análisis de componentes principales espectral, extrayendo así características generales del logo. Por último se implementará un método de clasificación basado en la forma de las marcas.

Estos métodos serán puestos a prueba con un subconjunto de imágenes extraídas de la base de logos mas grande actualmente, orientada a medir desempeño en metodologías de detección y reconocimineto de marcas.

El presente informe se estructura de la siguiente forma. A continuación se muestra un estado del arte en el tema de reconocimiento de marcas desde el punto de vista de visión computacional. Posteriormente, en este reporte se realiza una descripción de los métodos seleccionados. Finalmente se describe la base de datos a utilizar, justificando su uso mediante comparaciones con el resto de las bases de datos disponibles.

iiiiiii 00b9f2fa05bc5d45654137357ff7d40941774d74

II. ESTADO DEL ARTE

La primera forma en que se enfrentó el reconocimiento y detección de marcas fue a través de una persona, la cual "memoriza" una imagen y busca similitudes en una base de datos. Esta solución, al tener bases de datos de miles de logos, comprende un tiempo muy largo de trabajo por parte de los expertos, además de errores de interpretación o sesgos que dependen de cada persona. En visión computacional, resulta muy desafiante este tipo de problemas ya que se necesita detección y clasificación de elementos muy diversos en color, forma, tamaño, inclusión de cateres y/o palabras, etc. En este mismo sentido, resulta necesario la construcción de medidas de similitud o distancia entre distintas imágenes de marcas.

En la actualidad existen varias metodologías desarrolladas para el reconocimiento y detección de logos en imágenes basadas en visión computacional. Un estudio del desempeño de las metodologías con mejor desempeño fue hecho por Tursun et al. [?]. En esta publicación los autores hacen comparaciones entre metodologías basadas en extracción de características con apoyo experto[[?], [?], [?], [?], [?], [?]] y las basadas en aprendizaje de máquina utilizando redes convolucionales[[?], [?]].

El resultado del estudio hecho por Tursun et al. [?] concluye un desempeño sobresaliente de los métodos basados en redes convolucionales, sin embargo, aún no alcanzan una tasa de clasificación comparable con un humano experto.

Además de los resultados de desempeño de las distintas metodologías, en el trabajo de Tursun et al. [?] se presenta la segunda versión de la librería METU. Esta librería, cuya primera versión fue desarrollada por Tursun & Kalkan [?], es actualmente la base de datos mas diversa y numerosa de todas las utilizadas en investigaciones relacionadas con reconocimiento y detección de marcas en imágenes. Mas detalles de esta base de datos se muestran en secciones posteriores de este informe.

III. METODOLOGÍAS SELECCIONADAS

A continuación se describen las metodologías seleccionadas para medir sus desempeños en el problema de reconocimiento y detección de logos. Además, se probará el desempeño bajo tres métricas: euclidiana, similitud coseno y distancia manhattan.

A. Características de la textura: Patrones locales binarios (LBP)

Un algoritmo que se ha demostrado es muy eficiente para esta tarea es conocido por sus siglas en inglés como LBP[?]. Este algoritmo lo que hace es establecer similitud entre píxeles basados en vecindario. La comparación se realiza mediante el índice calculado con la siguiente ecuación,

$$L_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)^{2^p} \quad (1)$$

Donde P es la cuenta de los vecinos en un círculo de radio R , y g_p y g_c son intensidades de el pixel p y el pixel central respectivamente.

B. Características Generales: GIST

GIST es un algoritmo que extrae características generales del entorno en el cual está la imagen. Está basado en el calculo de componentes principales de la energía espectral. En el trabajo hecho por Douze et al. [?] muestran el desempeño de GIST como herramienta para el control de propiedad intelectual. En este estudio muestra tener un mejor desempeño que algoritmos usados comúnmente usados como detección de texto BoVW [?] con descriptores locales como SIFT [?].

C. Forma general de la imagen

La forma de la imagen es quizás uno de las características más importantes en la detección y reconocimiento de logos. El método propuesto por [?] tiene un buen desempeño en este tipo de problemas.

La metodología general de extracción de formas es la siguiente:

- 1) Muestrear uniformemente n puntos de la parte interior y exterior de la forma
- 2) Asignar un histograma tipo log-polar a cada muestra
- 3) De acuerdo a la localización de las muestras en cada histograma logarítmico, generar n vectores del contorno de la forma

IV. BASE DE DATOS METU

La base de datos METU es introducida por Tursun & Kalkan [?]. LA base de datos METU Posee mas de 900.000 imágenes de logos en distintas resoluciones. La tabla I muestra las características de METU.

En el estudio hecho por Tursun et al. [?] se muestra además que METU es la base de datos mas numerosa y refinada entre las usadas actualmente en investigación. La tabla II muestra una comparación entre las distintas bases de datos usadas en reconocimiento de marcas.

TABLE I: Características de base de datos MATU

Aspect	Version 2	Version 1
# trademarks	923,343	930,328
# query sets	417	320
# unique registered firms	409,675	410,439
# unique trademarks	687,842	690,418
# trademarks containing text only	583,715	589,098
# trademarks containing figure only	19,214	19,387
# trademarks containing figure and text	310,804	311,986
# trademarks with unknown contents	9,610	9,857
# file format	JPEG	JPEG
# Max Resolution	1,800 × 1,800(px)	1,800 × 1,800(px)
# Min Resolution	30 × 30(px)	30 × 30(px)

TABLE II: Comparación entre distintas bases de datos

Dataset	Number of logos	Requires preprocessing?	Image type	Image size (px)
UM	106	No	BW	various
MPEG7 CE2B	3,621	No	BW	-
Wei et al.	1,003	No	BW	200 × 200
Alwis et al.	210	No	BW	-
Alwis et al.	1,000	No	BW	-
abdel et al.	63,718	No	BW	-
MPEG7 CE2B	1,400	No	BW	256 × 256
MPEG7	3,000	No	BW	-
Jain et al.	1,100	No	BW	200 × 200
UKTR	10,745	No	BW	-
Leung et al.	2,000	No	BW	-
Her et al.	2,020	No	RGB	64 × 64
USPTO	~1,500,000	Yes	RGB	various
METU	923,343	No	RGB	various

Para este proyecto METU será la base de datos en que se probarán las metodologías seleccionadas. Algunos ejemplos de logos incluidos en la base de datos METU se muestran en las Figura 1.



(a) Ejemplos de logos entrenamiento de METU



(b) Ejemplos de logos de validación de METU

Fig. 1: Ejemplos de logos en base de datos METU

REFERENCES

- [1] Tursun, O., Aker, C., & Kalkan, S. (2017). A large-scale dataset and benchmark for similar trademark retrieval. arXiv preprint arXiv:1701.05766.
- [2] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- [3] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).
- [4] Tursun, O., & Kalkan, S. (2015, May). Metu dataset: A big dataset for benchmarking trademark retrieval. In Machine Vision Applications (MVA), 2015 14th IAPR International Conference on (pp. 514-517). IEEE.