

## APLICACIÓN DE REDES NEURONALES CONVOLUCIONALES A ESTEGOANÁLISIS

Harold Brayan Arteaga Arteaga – Mario Alejandro Bravo Ortiz

Grupos de Investigación en Automática e Ingeniería del Software

Semillero en Bioinformática e Inteligencia Artificial

UNIVERSIDAD AUTÓNOMA DE MANIZALES

Harold.ArteagaA@autonoma.edu.co

TUTORES: Reinel Tabares Soto – Raúl Ramos Pollán – Gustavo Isaza – Simón Orozco Arias

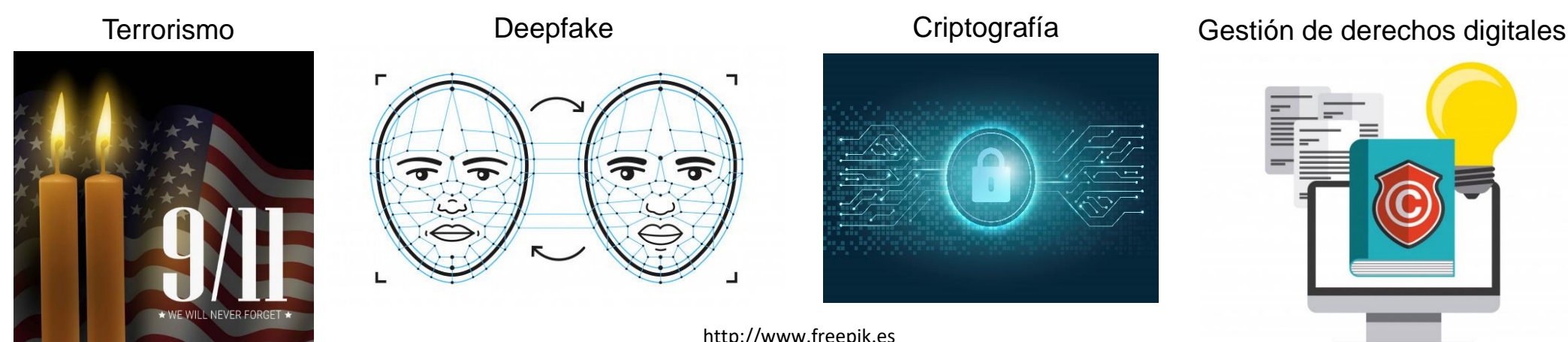
### PLANTEAMIENTO DEL PROBLEMA



### PREGUNTA DE INVESTIGACIÓN Y JUSTIFICACIÓN

¿Cuáles elementos computacionales y arquitecturas de CNNs son más apropiados para la detección de imágenes esteganográficas?

Las Redes Neuronales Convolucionales (CNNs) son herramientas muy potentes para la clasificación de imágenes esteganográficas [1]. Actualmente con los medios digitales, es posible enviar información oculta [2]. Esto se presenta para distintas actividades, como transmitir mensajes imperceptibles en Terrorismo. Modificación de imágenes tal que pasen desapercibidas como en Deepfake [7]. Y considerando que en algunos países está prohibida la criptografía, se puede usar esteganografía, contraparte del estegoanálisis. Además de la gestión de derechos digitales.



Con este proyecto se pretende aportar una nueva y más precisa propuesta de arquitectura [3]. Además de la presentación de una nueva base de datos, obtenida mediante selección objetiva con la red neuronal convolucional planteada.

### OBJETIVO GENERAL

Diseñar arquitecturas y elementos computacionales específicos de CNNs para la detección de imágenes esteganográficas.

### OBJETIVOS ESPECÍFICOS

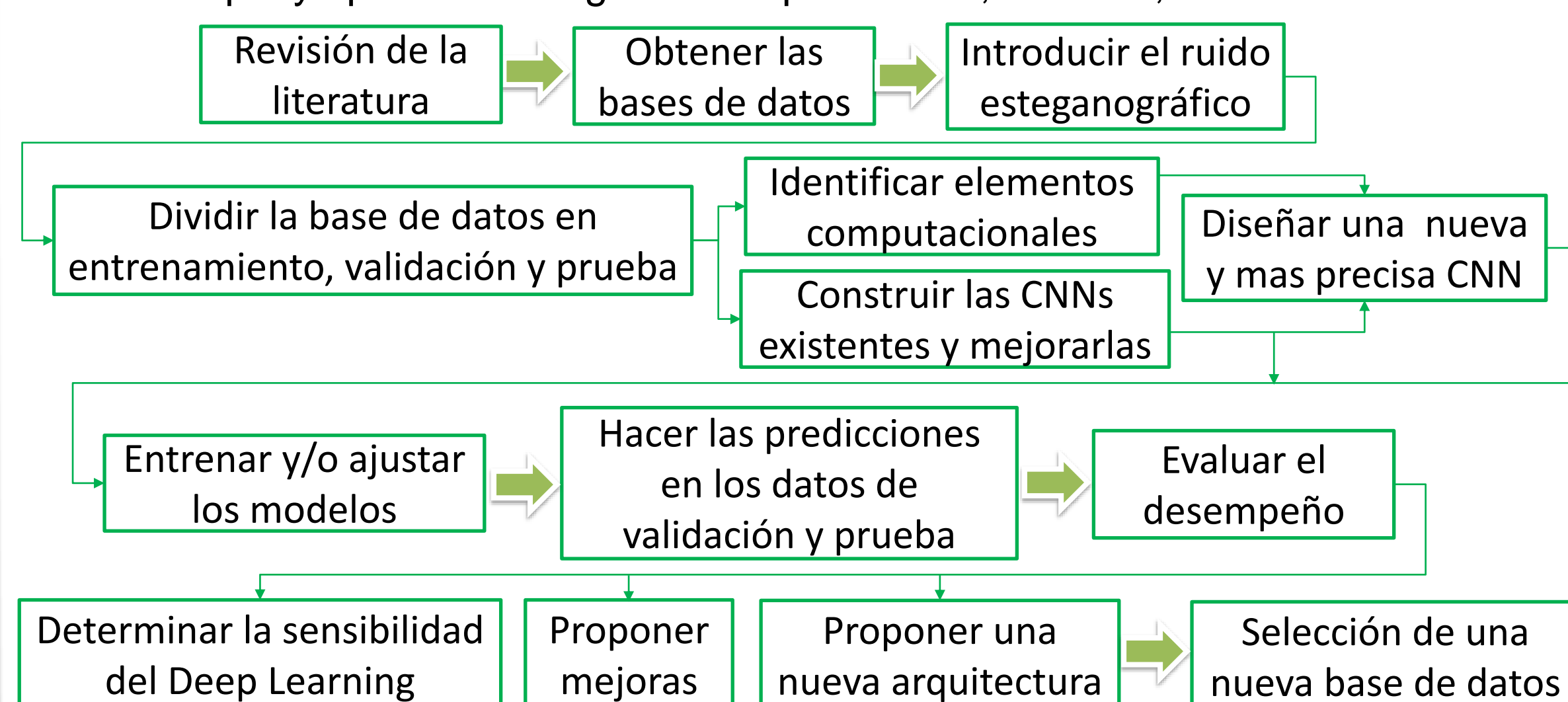
Analizar los resultados de las estrategias actuales para estegoanálisis de imágenes digitales en el dominio espacial.

Desarrollar elementos computacionales basados en CNNs respecto a las estrategias existentes para mejorar la precisión de detección de imágenes esteganográficas.

Diseñar arquitecturas de CNNs respecto a las estrategias existentes para mejorar la precisión de detección de imágenes esteganográficas.

### METODOLOGÍA PROPUESTA

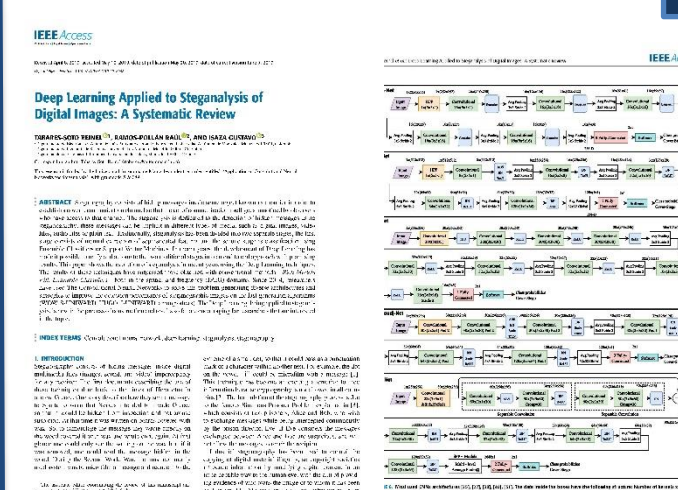
Enfoque y tipo de investigación: Experimental, Iterativo, Cuantitativo.



Bases de datos: BOSSbase 1.01 [5] y BOWS2 [6]

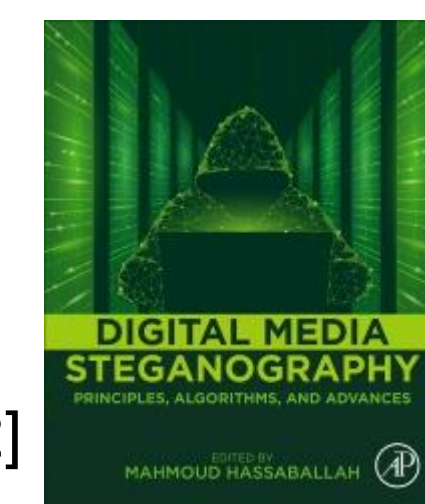
Estas bases de datos están conformadas cada una de 10000 imágenes que se llaman **cover**, las cuales NO tienen mensajes ocultos. Para evaluar los resultados se generan imágenes llamadas **stego**, incorporándoles información, con algoritmos esteganográficos [4] con una baja carga útil sobre los píxeles.

### RESULTADOS



Artículo científico de revisión de la literatura publicado en IEEE Access: **Deep Learning Applied to Steganalysis of Digital Images: A Systematic Review** [1]

Capítulo del libro publicado en Elsevier: **Digital Media Steganography: Principles, Algorithms, and Advances** [2]



Artículos científicos en proceso: Metodología para mejorar las arquitecturas existentes, tanto en precisión como en tiempo de convergencia. Y sobre la sensibilidad del Deep Learning aplicado a este problema.

Publicación en proceso con la propuesta de una nueva arquitectura: **GBRAS-Net**, que es más precisa que las anteriores. Convirtiéndose en el nuevo estado del arte.

ALGORITMO	WOW 0.2bpp	WOW 0.4bpp	S-UNIWARD 0.2bpp	S-UNIWARD 0.4bpp
(2020) GBRAS-Net	0.803	0.898	0.736	0.871
(2019) ZHU-Net	0.769	0.881	0.714	0.845
(2018) SR-Net	0.755	0.864	0.677	0.813
(2018) YEDROUDJ-Net	0.723	0.851	0.635	0.774
(2017) YE-Net	0.669	0.767	0.601	0.687
(2016) XU-Net	0.675	0.793	0.609	0.727
(2015) QIAN-Net	0.614	0.707	0.537	0.691
(2012) SRM+EC	0.635	0.745	0.634	0.753

Selección de una **base de datos** nueva para el dominio espacial, con precisión del 100% con la arquitectura propuesta, que consta de 28000 pares de imágenes cover, stego.

### BIBLIOGRAFÍA

- [1] Reinel, T.S., & Raúl G., & Gustavo I. (2019). Deep Learning Applied to Steganalysis of Digital Images: A Systematic Review. IEEE Access.
- [2] Mahmoud Hassaballah. (2020) Digital Media Steganography: Principles, Algorithms, and Advances. 1st Edition. Qena, Egypt. Elsevier Inc., p. 386. ISBN: 9780128194386.
- [3] R. Zhang, F. Zhu, J. Liu & G. Liu. (2019). Efficient feature learning and multisize image steganalysis based on CNN. 23, 534-540
- [4] V. Holub, J. Fridrich & T. Denemark. (2014). Universal distortion function for steganography in an arbitrary domain, EURASIP Journal on Information Security.
- [5] P. Bas, T. Filler, and T. Pevny. (2011) "Break Our Steganographic System": The Ins and Outs of Organizing BOSS," in INFORMATION HIDING, vol. 6958/2011 of Lecture Notes in Computer Science, (Czech Republic).
- [6] BOWS-2 Web Page. (2007). [Online]. Available: <http://bows2.ec-lille.fr/index.php?mode=VIEW&tmpl=index1>
- [7] Nguyen, Thanh & Nguyen, Cuong M. & Nguyen, Tien & Duc, Thanh & Nahavandi, Saeid. (2019). Deep Learning for Deepfakes Creation and Detection: A Survey.