



MAESTRÍA EN EXPLOTACIÓN DE DATOS Y GESTIÓN DEL
CONOCIMIENTO

Proyecto de Tesis

Franco Lianza

Abril 2022

Tema

Título del trabajo

(Tentativo) Manejo del sesgo en los datos: Detección de los dígitos escritos en los telegramas de las elecciones legislativas en Santa Fe mediante técnicas de adaptación de dominio.

Resumen

Resumen del área sobre la que se realizará el trabajo.

La rama de *Computer Vision* se encarga desarrollar herramientas para reconocer patrones complejos en imágenes en múltiples dominios. Se ha extendido exponencialmente a lo largo del tiempo, llegando a un punto en el cual se pueden detectar todo tipo de objetos con una precisión más que óptima [1].

Las técnicas desarrolladas en el área precisan de un gran volumen de datos para su entrenamiento. Esto implica que es de suma importancia de tener disponibles las *labels* (etiquetas) de los datos que se van a utilizar para entrenar los modelos. El etiquetado de los datos es una tarea costosa, ineficiente y hasta a veces resulta inviable de realizar [2].

Aún teniendo los *labels*, puede ocurrir que el *dataset* (conjunto de datos) donde se va a utilizar el modelo resulte diferente al que se utilizó para entrenarlo. Por mencionar, un modelo de detección de rostros entrenado en una etnia demográfica particular funcionará de manera errónea si se lo aplica a otra. Este fenómeno se conoce como *dataset bias* o *dataset shift* (sesgo en los datos). Dicho de otra manera, un modelo entrenado en un *dataset* puede no generalizar correctamente debido al *dataset bias*. Algunos autores afirman que el sesgo es un problema que no se puede evitar al momento de crear un *dataset* [3].

La detección de dígitos en los telegramas de elecciones en Argentina podría llevarse a cabo mediante un modelo entrenado en *datasets* de dígitos públicos como el *MNIST* [4]. Como no existe una única forma de escribir, el modelo estará sesgado a reconocer dígitos escritos de forma similar a los que se encontraban en el *dataset* de entrenamiento. No será capaz de generalizar lo aprendido en un dominio distinto.

En trabajos anteriores, se aplican distorsiones al conjunto de entrenamiento para intentar que el modelo pueda generalizar y aplicarse a los telegramas de elecciones de la Ciudad de Buenos Aires [5]. En el presente trabajo se utilizarán técnicas referidas al *transfer learning*, específicamente de *domain adaptation* para resolver el problema.

El *transfer learning* (transferencia de aprendizaje) es un área del *machine learning* que se encarga de almacenar el conocimiento ganado en un problema y aplicarlo a otro. *Domain adaptation* (adaptación de dominio) consiste en la habilidad de aplicar un modelo entrenado en uno o mas dominios de origen (*source domain*) en un uno distinto pero relacionado (*target domain*). La figura 1 muestra un ejemplo dos *datasets* de dígitos pero con dominios diferentes. Un modelo entrenado con *MNIST* no generalizará al conjunto *SVHN* por más que ambos sean dígitos.

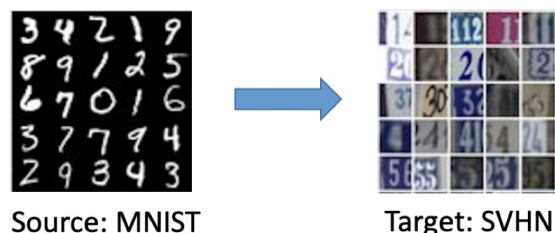


Figure 1: Ejemplo dominios diferentes: MNIST y SVHN

Director o Tutor

El nombre del director o tutor de la tesis o trabajo, si éste ya hubiese aceptado la tarea.

Pendiente determinar el director de la tesis.

Motivación e importancia del campo

Explicar la o las motivaciones que llevan a realizar el trabajo planteado y su importancia.

La principal motivación es obtener un modelo que, basado en un *dataset* similar, pueda interpretar los dígitos de los telegramas de las elecciones de Santa Fe sin tener de los *labels* de los mismos. De esta manera, el modelo podría utilizarse en futuras elecciones a fin de automatizar la digitalización de los mismos.

Problemas no resueltos

Problemas no resueltos detectados en el área y que el trabajo a realizar pretende resolver.

La digitalización de los telegramas de las elecciones sigue siendo de forma manual, lo que genera lentitud y desconfianza en el proceso. Detectar los votos de manera automática y reducir la intervención humana, aumentará la eficiencia de las elecciones y la confianza en ellas.

Objetivo del trabajo

Explicar claramente el objetivo del trabajo, especificando su alcance y limitaciones.

El objetivo del trabajo consiste en detectar los dígitos escritos en los telegramas de las elecciones legislativas de Santa Fe utilizando modelos entrenados mediante técnicas de adaptación de dominio en *datasets* públicos de dígitos.

Requerimientos y desafíos

Requerimientos y desafíos que plantea el trabajo a realizar.

El trabajo presenta como principal desafío poder extraer los votos de los telegramas y analizar distintos modelos utilizando las técnicas de *domain adaptation* disponibles. También constará de investigación respecto a las *Generative Adversarial Networks (GANs)* debido a que varios algoritmos de *domain adaptation* dependen de ellas [6, 7, 8].

Metodología

Metodología a emplear para el desarrollo del trabajo.

Los telegramas serán descargados desde la [página oficial del estado argentino](#). Luego se extraerán los dígitos de los votos de cada telegrama utilizando la librería *OpenCV* [9]. Una vez obtenido los dígitos, se entrenarán distintas redes convolucionales (LeNet [4], ResNet [10], etc) utilizando algoritmos de *domain adaptation*. Finalmente, se evaluarán los modelos y se seleccionará/n el/los mejor/es.

Plan de Trabajo

Especificar las distintas tareas a realizar con los tiempos que se estime que deberían insumir. Indicar las fechas estimadas de inicio y finalización del trabajo.

El trabajo se realizará como consecuencia de la realización de las siguientes etapas:

- Investigación respecto de técnicas de *GANs* y *domain adaptation*.
- Recolección de telegramas, extracción de dígitos y limpieza de los mismos.
- Elaboración de experimentos para el entrenamiento de las redes convolucionales.
- Evaluación y comparación de los modelos obtenidos.
- Redacción final de la tesis.

La fecha de inicio del proyecto es Febrero 2022 y como fecha de finalización estimada Septiembre 2022.

Bibliografía

- [1] Richard Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.
- [2] Pedro Miguel Lima de Sousa Reis. Data labeling tools for computer vision: a review. 2022.
- [3] Aditya Khosla, Tinghui Zhou, Tomasz Malisiewicz, Alexei A Efros, and Antonio Torralba. Undoing the damage of dataset bias. In *European Conference on Computer Vision*, pages 158–171. Springer, 2012.
- [4] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [5] Walter Marcelo Lamagna. *Lectura artificial de números manuscritos en datos abiertos de elecciones legislativas en la Ciudad de Buenos Aires*. PhD thesis, Universidad de Buenos Aires. Facultad de Ciencias Exactas y Naturales, 2016.
- [6] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
- [7] Ruijia Xu, Guanbin Li, Jihan Yang, and Liang Lin. Unsupervised domain adaptation: An adaptive feature norm approach. *arXiv preprint arXiv:1811.07456*, 2018.
- [8] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *The journal of machine learning research*, 17(1):2096–2030, 2016.
- [9] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- [10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.