

Marco Metodológico para el Desarrollo de un Sistema de Reconocimiento Biométrico Mediante Técnicas de Machine Learning

Silvia Estela Ruiz¹ y Carlos Eduardo Alvez²

Universidad Nacional de Entre Ríos

¹{silvia.ruiz, ²carlos.alvez}@uner.edu.ar

Abstract: El uso del Aprendizaje Automático en los sistemas de reconocimiento biométrico supone un gran paso en la evolución tecnológica. En el presente trabajo se exploran diferentes papers en los cuales se aplican distintos mecanismos del aprendizaje automático en los sistemas de reconocimiento biométrico mediante iris, los cuales han logrado resultados exitosos en los últimos años. Debido a que en la actualidad no existe una metodología estándar formal de proyectos que utilicen Aprendizaje Automático, se analizaron varias metodologías y modelos de procesos adoptadas en la actualidad. Por consiguiente, el presente trabajo de tesis, tiene como principal objetivo proponer un marco metodológico para el desarrollo de un sistema biométrico de reconocimiento de iris mediante Aprendizaje Automático.

Keywords: reconocimiento de iris, aprendizaje automático, metodología.

Fecha de inicio: Abril 2022.

1 Introducción

Se observa en la actualidad una tendencia muy marcada en la incorporación de técnicas biométricas, tales como rostro, huellas digitales, geometría de la mano, iris, patrones de retina, firma y voz, entre las más destacadas en los sistemas de identificación/autenticación de personas. Todas estas técnicas presentan diferentes grados de singularidad, permanencia, mensurabilidad, desempeño, aceptación del usuario y robustez.

Los sistemas de reconocimiento biométrico utilizan características fisiológicas o de comportamiento propias de cada individuo para identificarlo, es decir, se reconoce al usuario por lo que es en lugar de por lo que posee o sabe [1]. Los rasgos fisiológicos presentan una reducida variabilidad a lo largo del tiempo, pese a que su adquisición es más invasiva y requiere de la cooperación de los sujetos. Por el contrario, los rasgos de comportamiento resultan menos invasivos aunque la exactitud de la identificación es menor debido a la variabilidad de los patrones de comportamiento.

De esta forma, el objetivo de todos ellos será obtener, a partir de la captura de un rasgo biométrico, una representación de cada individuo que resulte lo suficientemente discriminante respecto a las de otros usuarios del sistema. En lo que respecta

a los distintos rasgos biométricos, el reconocimiento del iris se considera uno de los sistemas biométricos más confiables y precisos, a pesar de ser un campo de estudio muy reciente en relación a otros rasgos [2]. En diferentes trabajos se ha demostrado [1, 3] que el rasgo del iris tiene una serie de ventajas con respecto a otros rasgos biométricos (por ejemplo, cara, huella digital), que lo hacen comúnmente aceptado para su aplicación en sistemas biométricos precisos y de alta confiabilidad. Entre las distintas ventajas que presenta el iris se destacan: un patrón único en cada persona, la estructura permanece invariable durante toda la vida del ser humano y la adquisición de la imagen se ha simplificado con el avance tecnológico de distintos dispositivos de captura. Por los motivos precedentes el desarrollo de este trabajo de tesis se enfoca en sistemas de reconocimiento basados en iris. En la Figura 1, se representa un esquema con el proceso de un sistema biométrico de reconocimiento de iris. El primer paso captura la imagen y el segundo la reconoce, realiza el preprocesamiento y la extracción de características. La etapa de normalización tiene como objetivo obtener una imagen del iris que sea independiente del tamaño de la pupila y permita la comparación entre diferentes iris. Los valores de las características son almacenados en la base de datos como plantillas (templates) biométricas. Luego, la toma de decisión se realiza clasificando la entrada dada con las plantillas biométricas en la base de datos.



Fig. 1. Esquema de un sistema de reconocimiento biométrico de iris.

A pesar de las ventajas que presentan los sistemas de reconocimiento biométrico enfrentan diversos desafíos en cuanto a la precisión de los datos. En este sentido, a lo largo de los últimos años, se han comenzado a utilizar diferentes técnicas de Aprendizaje Automático o Machine Learning (en adelante ML) en las etapas que conforman el proceso de reconocimiento de iris [4]. En [5,6] se destacan principalmente las técnicas de ML tales como SVM (Máquina de Soporte Vectorial), WPNN (Redes Neuronales Probabilísticas Ponderadas), BPNN (Redes Neuronales de Retro Propagación) y RBFNN (Redes Neuronales con Función de Base Radial), las cuales son utilizadas generalmente en la etapa de clasificación de los patrones del iris, así como también para la normalización, segmentación y extracción de características.

En las últimas dos décadas, el término Aprendizaje Automático se ha convertido en una herramienta común en casi cualquier tarea que requiera la extracción de información de grandes conjuntos de datos. Estamos rodeados de una tecnología basada en el Aprendizaje Automático: motores de búsqueda, software antispam, software para detección de fraudes, etc. Una característica común de todas estas aplicaciones es que, en contraste con los usos más tradicionales de las computadoras, en estos ca-

sos debido a la complejidad de los patrones que necesitan ser detectados, un programador humano no puede proporcionar información explícita y detallada de cómo se deben ejecutar tales tareas. Muchas de las habilidades del ser humano se adquieren o refinan a través del aprendizaje de la experiencia (en lugar de seguir instrucciones explícitas). Las herramientas de Aprendizaje Automático se interesan en dotar a los programas de la capacidad de "aprender" y adaptarse.

Por otro lado, en toda implementación de un sistema es conveniente seguir un modelo de proceso o metodología [7] a fin de obtener mayor probabilidad de éxito en los resultados obtenidos. Según Pressman [7], un modelo de proceso define un conjunto de actividades organizadas para llevar a cabo una tarea determinada, mientras que una metodología además de definir las actividades y tareas de un proceso, define cómo llevar a cabo las mismas.

Al respecto, si bien en la actualidad se están desarrollando distintas líneas de investigación para definir una metodología estándar específica para utilizar en sistemas que utilicen ML [8] aún no existe una definición estándar formal. En este sentido, existen varias metodologías y modelos de procesos de minería de datos y ML, tales como KDD (Knowledge Discovery in Databases) y SEMMA (Sample, Explore, Modify, Model and Assess), no obstante, se suele adoptar como base el estándar de facto CRISP-DM (Cross Industry Standard Process for Data Mining) [9]. Este modelo de referencia es utilizado habitualmente como marco metodológico en distintos tipos de proyectos [10], principalmente de minería de datos.

En este sentido, el presente trabajo de tesis tiene como principal objetivo proponer un marco metodológico para el desarrollo de un sistema biométrico de reconocimiento de iris mediante ML. Para esto, se realizará un estudio de los fundamentos teóricos de la biometría, los sistemas biométricos y de las distintas técnicas de ML. Además, será necesario realizar una revisión sistemática de los distintos modelos de procesos de trabajo y/o metodologías adoptadas en el desarrollo de sistemas que utilicen ML, analizando las ventajas y desventajas de las mismas e identificando los desafíos que se encuentran en su aplicación. En base a la información recolectada, se realizará un análisis comparativo y se elaborará una propuesta de marco metodológico que pueda ser aplicado en un proyecto real de un sistema de reconocimiento de iris mediante ML.

2 Aportes realizados

A continuación se presentan los aportes parciales obtenidos en pos del objetivo general del trabajo de tesis. Se propone un marco metodológico específico para proyectos de sistemas de reconocimiento de iris mediante el uso de técnicas de ML. En la Fig. 2 se presenta el esquema metodológico propuesto y se detallan brevemente las fases generales.

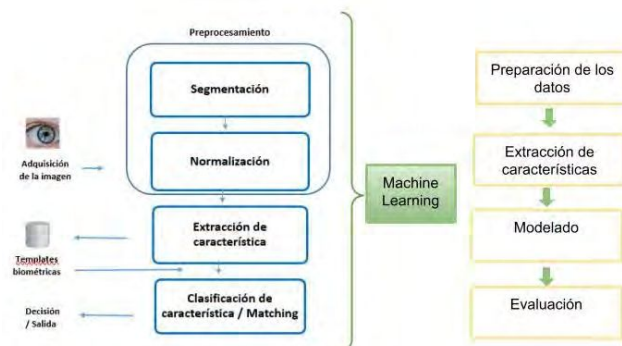


Fig. 2 Marco metodológico propuesto.

- **Preparación de los datos**

Implica las tareas de análisis, descripción, verificación y preprocesamiento de los datos. Se debe preprocesar la imagen para localizar los límites de la pupila y el iris y así poder normalizar el iris. Para ello se utilizan diferentes algoritmos como por ej. filtro gaussiano de paso bajo el cual permite suavizar y agudizar la imagen del ojo. Todo este proceso permite la adaptación de los datos para el uso posterior de las técnicas de ML seleccionadas.

- **Extracción de características**

En esta fase se procede a la extracción de características y su correspondiente codificación. Es importante destacar que, el éxito de cualquier sistema biométrico definido como un sistema de clasificación y reconocimiento depende principalmente de la eficiencia y robustez de las etapas de extracción y clasificación de características.

- **Modelado**

Esta fase abarca el diseño, construcción y evaluación del modelo. Aquí es necesario seleccionar las técnicas de modelado más apropiadas para la construcción del modelo. Los parámetros utilizados en la generación del modelo, dependen de las características de los datos y de la precisión que se desea lograr con el modelo. En general, se separan los datos en dos conjuntos, uno de entrenamiento y otro de prueba, para luego construir el modelo basado en el conjunto de entrenamiento y medir la calidad del modelo generado con el conjunto de prueba. La selección de los mejores parámetros es un proceso iterativo y se basa exclusivamente en los resultados generados.

- **Evaluación**

Una vez que se tienen los datos preparados, se puede empezar a entrenar el modelo. Para ello, se realizan diferentes entrenamientos con distintos algoritmos y parámetros de los mismos. En lo que respecta a los sistemas de reconocimiento de iris, en trabajos recientes [12,13] se ha demostrado que el uso Deep Learning y Redes Neuronales logran tasas de precisión significativamente alta mejorando aspectos de eficiencia y confiabilidad. Los resultados del modelo podemos verlos mediante la utilización de distintas herramientas que permitan la visualización del desempeño del algoritmo de aprendizaje utilizado.

3 Posibles líneas de investigación a futuro

Las actividades recomendadas en cada fase de este marco metodológico surgen de prácticas usuales de distintos proyectos de ML y minería de datos. Sin lugar a dudas, la adopción de un marco metodológico en este tipo de proyectos favorece la calidad del proceso para la obtención de datos confiables en forma eficiente. Como trabajo futuro se propone profundizar las fases de preparación de los datos y modelado, consideradas esenciales para el éxito de un sistema de reconocimiento biométrico, avanzando en una descripción más detallada de las tareas que las conforman definiendo sus objetivos específicos y los documentos entregables de cada una. Por otro lado, se pretende profundizar en la aplicación de diferentes formas de medir su éxito, es decir, utilizando métricas específicas y criterios de aceptación.

4 Bibliografía básica

1. A. K. Jain, A. Ross and S. Prabhakar. "An introduction to biometric recognition". IEEE Transactions on Circuits and Systems for Video Technology., 14(1), 2004.
2. J. Daugman. "High confidence visual recognition of persons by a test of statistical independence". IEEE Transactions on Pattern Analysis and Machine Intelligence. 1993.
3. A. Basit. "Iris Recognition: An Identification Biometric System". Lap Lambert Academic Publishing. 2010.
4. M. De Marsico, A. Petrosino, y S. Ricciardi. "Iris recognition through machine learning techniques: A survey". Pattern Recognit. Lett., vol. 82, pp. 106-115, oct. 2016, doi: 10.1016/j.patrec.2016.02.001.
5. A. A. Khan, S. Kumar, y M. Khan, "Iris Pattern Recognition using Support Vector Machines and Artificial Neural Networks". IJIREEICE, pp. 2208-2211, dic. 2014, doi: 10.17148/IJIREEICE.2014.21203.
6. K. Saminathan and T. Chakravarthy and M. Chithra Devi. "Iris Recognition based on kernels of support vector machine". ICTACT J. Soft Comput., vol. 05, n.o 02, pp. 889-895, ene. 2015, doi: 10.21917/ijsc.2015.0125.]
7. R. Pressman y B. Maxim, "Software engineering: a practitioner's approach". Novena Edición. McGraw-Hill Education. New York. 2020.
8. S. Studer, T. Bui, C. Drescher, A. Hanuschkin, L. Winkler, S. Peters, K. Mueller, "Towards CRISP-ML(Q): A Machine Learning Process Model with Quality Assurance Methodology". ArXiv200305155 Cs Stat, mar. 2020, Accedido: ene. 29, 2022. [En línea]. Disponible en: <http://arxiv.org/abs/2003.05155>.
9. C. Shearer, "The CRISP-DM model: the new blueprint for data mining". J Data Wareh., vol. 5, pp. 13-22, ene. 2000.
10. V. G. Cortina, "Aplicación de la Metodología CRISP-DM a un Proyecto de Minería de Datos en el Entorno Universitario". p. 120.
11. J. Daugman, "How iris recognition works". Proceedings of 2002 International Conference on Image Processing, Vol. 1, 2002.
12. A. Al-Waisy, R. Qahwaji, S. Ipson, A. Shumoos, T. Nagem "A multi-biometric iris recognition system based on a deep learning approach". Pattern Analysis and Applications. 21(3): 783-802. 2017.
13. S. Minace, A. Abdolrashidi, H. Su, M. Bennamoun, D. Zhang , "Biometric Recognition Using Deep Learning: A Survey". 2020.