

El modelo SNARE-CNN, red neuronal convolucional 2D| Capítulo 6.4

Autores:  Dr^a De la Morena Vacas, Elena |  Juan Antonio Lloret Egea. Miembro de 
Escrito: 26/12/2019 | Actualizado: 26/12/2019

➤ "Una pérdida funcional de proteínas SNARE se ha implicado en una variedad de enfermedades humanas" [C6.4-2]

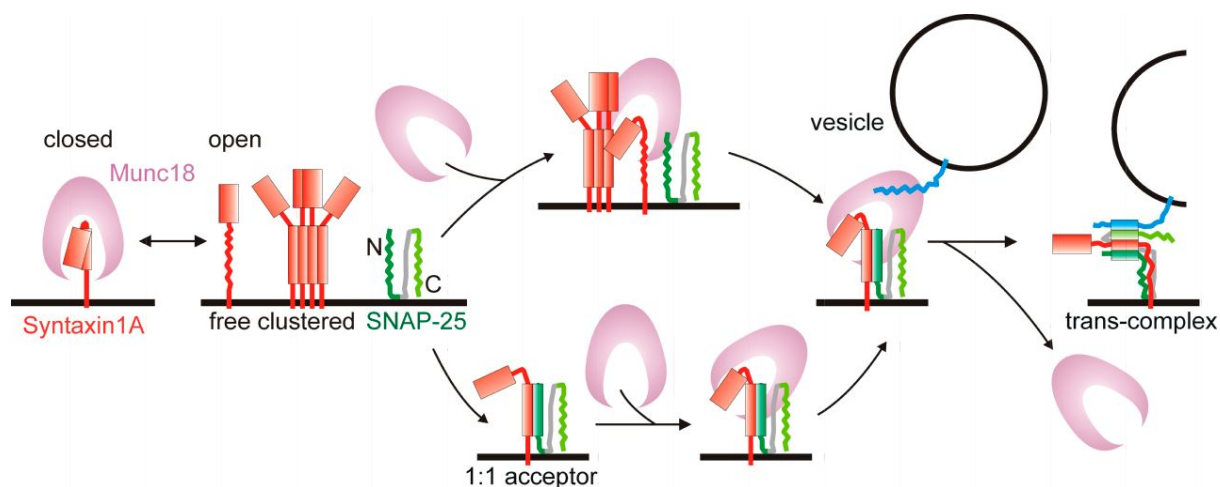


Fig. I A. C6.4-1- Representación de la formación de un complejo trans-SNARE. Indica cómo Munc18 puede interactuar con la proteína SNARE durante la formación del complejo. Crédito imág. URL

Índice

6.4.1.- Introducción o Resumen
6.4.2.- Aplicación del <i>deeplearning</i> a las proteínas
6.4.3.- El modelo SNARE-CNN es uno de los trabajos más citados en 2019, según PeerJ
6.4.4.- Abstract
6.4.5.- Nota final

6.4.1.- Introducción | Resumen

Las proteínas SNARE (Soluble NSF Attachment Protein) son un grupo de éstas cuyo objetivo más importante es facilitar la fusión de las vesículas encargadas del transporte de moléculas necesarias para el funcionamiento de las células con los compartimentos celulares apropiados [C6.4-1].

6.4.2.- Aplicación del *deep learning* a las proteínas

El modelo SNARE-CNN es una arquitectura de red neuronal convolucional 2D¹ para identificar proteínas SNARE a partir de datos de secuenciación de alto rendimiento. ("Sensibilidad alcanzada del 76,6%, especificidad del 93,5%, precisión del 89,7% y CCM de 0,7 en el conjunto de datos de validación cruzada" [C6.4-2]).

Este modelo está disponible *Open Access* en <https://github.com/khanhlee/snare-cnn/>

The screenshot shows the GitHub repository page for 'snare-cnn' by 'khanhlee'. At the top, there are buttons for 'Watch' (0), 'Star' (0), and 'Fork' (0). Below this, there are navigation tabs for 'Code', 'Issues' (0), 'Pull requests' (0), 'Projects' (0), 'Security', and 'Insights'. A prominent banner encourages users to 'Join GitHub today' with a 'Sign up' button. Below the banner, it states 'No description, website, or topics provided.' A summary bar shows '16 commits', '1 branch', '0 packages', '0 releases', and '1 contributor'. There are buttons for 'Branch: master', 'New pull request', 'Find file', and 'Clone or download'. The file list includes: 'dataset' (Delete a.txt, last year), 'figures' (Add files via upload, last year), 'README.md' (Update README.md, 10 months ago), 'snare_classification.py' (Update snare_classification.py, last year), 'snare_plot_optimizers.py' (Add files via upload, last year), and another 'README.md'.

Fig. I A. C6.4-2- Repositorio Github y el modelo snare-cnn. Crédito imág (My own snapshot). URL: <https://github.com/khanhlee/snare-cnn/>

¹ Las redes neurales tienen una arquitectura estratificada. Cada capa está formada por un número de nodos, y cada uno de éstos realiza de forma efectiva operaciones matemática sobre una entrada/s para calcular un resultado. La entrada de cualquier nodo es una suma ponderada de los resultados de la capa anterior (más un término de sesgo que normalmente es igual a uno o a cero). Éstas son las ponderaciones que el algoritmo aprende durante su capacitación. Para aprender estos parámetros el resultado de una ejecución de una capacitación se compara con el valor verdadero, y el error se retropropaga a lo largo de la red para actualizar las ponderaciones. La convolución es una operación matemática en la que una función se 'aplica' de alguna manera a otra función. El resultado se puede entender como una 'mezcla' de las dos funciones. [C6.4-1-1].

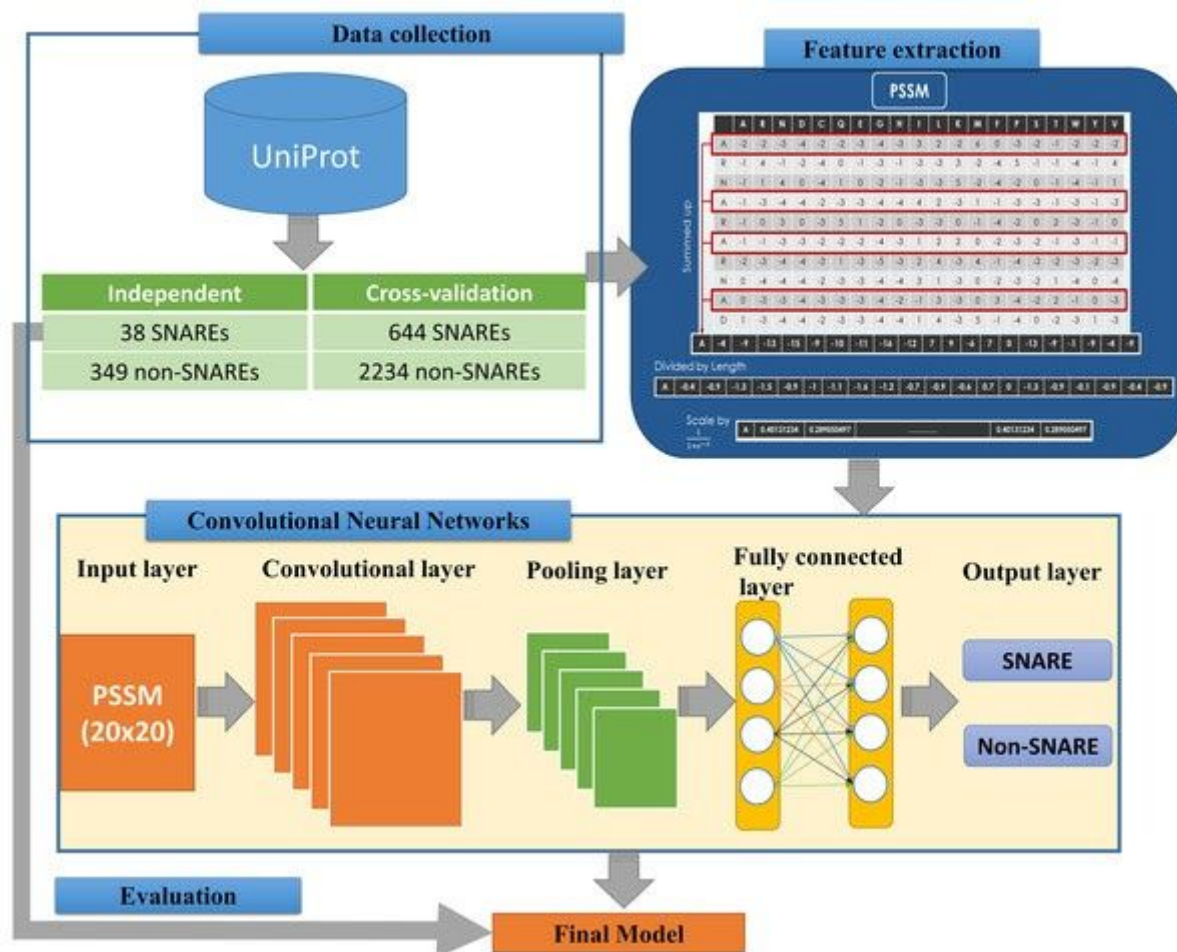


Fig. 1 A. C6.4-3- Diagrama de flujo para identificar proteínas SNARE utilizando redes neuronales convolucionales bidimensionales. Crédito imág (DOI: 10.7717/peerjcs.177/fig-1). URL: <https://dfzljdn9uc3pi.cloudfront.net/2019/cs-177/1/fig-1-1x.jpg>

6.4.3.- El modelo SNARE-CNN es uno de los trabajos más citados en 2019, según PeerJ

Desde 2019 PeerJ ha publicado más de 2.000 artículos revisados por pares y más de 8.000 publicaciones de *Open Access* desde que comenzaron a publicar en 2013, indican en su blog. Y de entre todas las áreas y artículos que se describen como los más citados, nosotros ponemos nuestro centro de atención sobre la inteligencia artificial (*deeplearning*) aplicada a la bioinformática y en este artículo: [SNARE-CNN: a 2D convolutional neural network architecture to identify SNARE proteins from high-throughput sequencing data](#) (by Le and Nguyen) ^[C6.4-3].

6.4.4.- Abstract

El aprendizaje profundo se ha utilizado cada vez más para resolver numerosos problemas en diversos campos con un rendimiento de vanguardia. También se puede aplicar en bioinformática para reducir el requisito de extracción de características y alcanzar un alto rendimiento. Este estudio intenta utilizar el aprendizaje profundo para predecir las proteínas SNARE, que es una de las funciones moleculares más vitales en las ciencias de la vida. Una pérdida funcional de proteínas SNARE se ha implicado en una variedad de enfermedades humanas (p. Ej., Neurodegenerativas, enfermedades mentales, cáncer, etc.). Por lo tanto, crear un modelo preciso

para identificar sus funciones es un problema crucial para comprender estas enfermedades y diseñar los objetivos farmacológicos. Nuestro modelo SNARE-CNN que utiliza redes neuronales convolucionales bidimensionales y perfiles de matriz de puntuación de posición específica podría identificar proteínas SNARE con una sensibilidad alcanzada del 76,6%, especificidad del 93,5%, precisión del 89,7% y CCM de 0,7 en el conjunto de datos de validación cruzada. También evaluamos el rendimiento de nuestro modelo a través de un conjunto de datos independiente y el resultado muestra que somos capaces de resolver el problema de sobreajuste. En comparación con otros métodos de vanguardia, este enfoque logró una mejora significativa en todas las métricas. A lo largo del estudio propuesto, proporcionamos un modelo efectivo para identificar las proteínas SNARE y una base para futuras investigaciones que pueden aplicar el aprendizaje profundo en bioinformática, especialmente en la predicción de la función de las proteínas. [C6.4-2]

6.4.5.- Nota final

Usando el modelo, en sus conclusiones, los autores señalan que las nuevas proteínas SNARE pueden identificarse con precisión y usarse para el desarrollo de fármacos. Y tratándose de enfermedades como las neurodegenerativas, mentales y el cáncer debemos interesarnos por este trabajo aplicado al campo de la bioinformática computacional, la minería de datos y el *Machine Learning*.

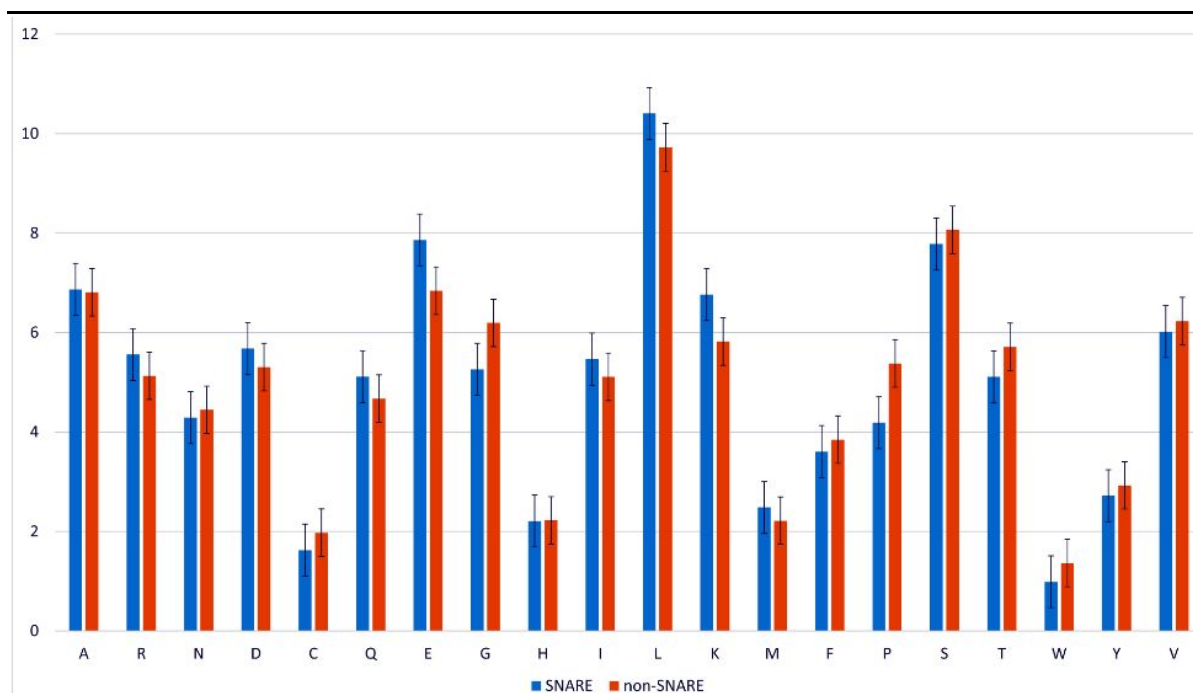


Fig. 1 A. C6.4-4- Identificando proteínas SNARE. Crédito imág (My own snapshot). URL: <https://github.com/khanhlee/snare-cnn/blob/master/figures/Figure%201.pdf>

Citar este artículo

Lloret Egea, Juan Antonio; De la Morena Vacas, Elena (2019): 6.4-El modelo SNARE-CNN, red neuronal convolucional 2D.pdf. figshare. Preprint. <https://doi.org/10.6084/m9.figshare.11472030.v1>

Bibliografía

[C6.4-1]. *SNARE (proteína)*. Wikipedia. [Recuperado (26/12/2019) de [https://es.wikipedia.org/wiki/SNARE_\(proteína\)](https://es.wikipedia.org/wiki/SNARE_(proteína))]

[C6.4.1-1] James Cowley. *Redes neuronales convolucionales Utilizar Python para implementar una red sencilla que clasifica dígitos escritos a mano.* (07-12-2018). [Recuperado (26/12/2019) de

<https://www.ibm.com/developerworks/ssa/library/cc-convolutional-neural-network-vision-recognition/cc-convolutional-neural-network-vision-recognition-pdf.pdf>]

[C6.4-2]. Le NQK, Nguyen VN. 2019. *SNARE-CNN: una arquitectura de red neuronal convolucional 2D para identificar proteínas SNARE a partir de datos de secuenciación de alto rendimiento.* PeerJ Comput. Sci. 5: e177

<http://doi.org/10.7717/peerj-cs.177>

[C6.4-3]. PeerJ Community. (19/12/2019). *PeerJ en Review 2019: los principales artículos citados del año.* PeerJBlog.

[Recuperado (26/12/2019) de

https://peerj.com/blog/post/115284882038/peerj-in-review-2019-top-cited-papers-of-the-year/?utm_source=summary_email_blog&utm_medium=email&utm_campaign=connection