



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
FACULTAD DE BIOLOGIA

Trabajo Bibliográfico - Entrega 1:
Aplicación del machine learning a la predicción del Cáncer

by

Nicolás Buzeta

José Caceres

Benjamin Barrios

Carlos Álvarez

Profesora : Alicia Nogueras

Tutor : Andrés Carreño

12 de octubre de 2020

Índice general

1	Abstract	2
2	Introducción	3

1 Abstract

El Cáncer involucra en gran medida, procesos en las células que son alterados por una falla o mutación de estos mismos. Algunos de estos procesos celulares alterados por el cáncer desde la aparición temprana hasta una etapa avanzada son en la mayoría de los casos, la señalización celular, el ciclo celular y la apoptosis, los cuales serán de vital importancia para explicar la aplicación del estudio.

La señalización celular es la comunicación entre células, que está mediada por moléculas de señalización extracelulares, esta comunicación puede ser entre células lejanas o entre células vecinas. La recepción está a cargo de proteínas receptoras que se encuentran en la superficie celular, estas activan los sistemas de señalización intracelular, los cuales distribuyen la señal hacia los efectores y son estos últimos los que median el cambio necesario en la célula según la señal recibida (Alberts y col., 2010, p. 879).

También está el ciclo celular que es el mecanismo mediante el cual todos los organismos se reproducen, este se lleva a cabo a través de una secuencia ordenada de acontecimientos ocurridos en la célula en los que duplica su contenido y luego se divide en dos. Todo este proceso está regulado por proteínas que controlan de cierta forma el realizar la división celular para sí solo llevarlo a cabo cuando sea necesario (Alberts y col., 2010, p. 1053).

Por último, la muerte celular programada se lleva a cabo en la mayoría de los casos mediante un proceso denominado apoptosis, el cual cuando una célula se daña o se infecta procede a su muerte. Este proceso se lleva a cabo mediante una secuencia de procesos moleculares, en los que la célula se destruye así misma desde dentro de forma sistemática siendo luego digerida por otras células, sin dejar rastro (Alberts y col., 2010, p. 1115).

Todos estos procesos recién nombrados, al ser alterados por la existencia del cáncer en una persona, cumplirán un rol importante al momento de determinar su existencia desde sus inicios, ya que al examinarlos, podremos darnos cuenta si hay rastro alguno de un posible cáncer a nivel intercelular.

En esta investigación nosotros ocuparemos una tecnología relativamente nueva, llamada machine learning. Esta tecnología es capaz de aprender, por experiencia, de forma autónoma e independiente («Machine Learning | Cambridge English Dictionary», s.f.). De esta forma, nosotros nos centraremos en buscar y analizar las características de los núcleos celulares. De este modo podremos detectar cambios chicos en su estructura y forma, y con estos datos, hacer una predicción sobre el estado de esa célula.

2 Introducción

La organización espacial dentro del núcleo se debe, mayoritariamente, a dos estructuras biológicas, la matriz nuclear y la lámina nuclear. La matriz nuclear es una red de fibras que se encuentra por todo el interior de los núcleos celulares y, aparte de dar rigidez al núcleo, mantiene el orden dentro del núcleo. Por otro lado, la lámina nuclear es una red que, aparte de también dar rigidez, regula eventos como la replicación de ADN y la división celular (Alberts y col., 2010; Zink y col., 2004, p. 200). En células cancerígenas, ambas de estas estructuras se pueden ver afectadas. Esto comúnmente resulta en el núcleo de la célula cambiando su forma, de una elipsoide a irregular.

La estructura del nucleus final en estas células malignas, terminan con diferencias características en la arquitectura nuclear (Rynearson & Sussman, 2011). Esto significa que no es solo posible identificar células cancerígenas, sino también es posible identificar el tipo de cáncer. Esto significa que poder detectar estos cambios no solo ayudaría a encontrar cáncer, sino también es posible detectar el tipo cuando todavía no se ha formado el tumor. Un ejemplo de esto sería la detección de células en la orina («Urine Cytology - Mayo Clinic», s.f.). Como estas células pueden venir de un tumor no visible, podríamos detectar y tratar el cáncer sin necesidad de tomar una biopsia directa.

Dada toda esta información, es muy difícil detectar estos cambios con métodos convencionales, ya que debemos detectar cambios minúsculos en la estructura nuclear. Dado esto, machine learning es la mejor tecnología para esta aplicación, ya que es capaz de aprender relaciones de manera autónoma. Esto significa que es posible darle información sobre millones de células y el programa aprenderá a identificar las células malignas sin intervención humana. Esta tecnología está viendo mucho uso actualmente, y sigue siendo investigada activamente (Carleton y col., 2018; Kourou y col., 2015).

También es posible automatizar, hasta cierto punto, el aprendizaje y uso del machine learning. En este caso, podríamos aplicar técnicas automatizadas para clasificar los núcleos dentro de las células. De esta manera solo sería necesaria extraer la célula y podríamos automáticamente encontrar el núcleo (Sirinukunwattana y col., 2016). Por lo tanto, el proceso solo necesitaría intervención humana durante la extracción, y tras eso sería completamente automatizado.

Referencias

- Alberts, B., Wilson, J. & Hunt, T. (2010). *Biología molecular de la célula*. Omega OCLC: 708275267.
- Carleton, N. M., Lee, G., Madabhushi, A. & Veltri, R. W. (2018). Advances in the Computational and Molecular Understanding of the Prostate Cancer Cell Nucleus. *Journal of cellular biochemistry*, 119(9), 7127-7142. <https://doi.org/10.1002/jcb.27156>
- Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V. & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8-17. <https://doi.org/10.1016/j.csbj.2014.11.005>
- Machine Learning | Cambridge English Dictionary*. (s.f.). Consultado el 6 de septiembre de 2020, desde <https://dictionary.cambridge.org/dictionary/english/machine-learning>
- Rynearson, A. L. & Sussman, C. R. (2011). Nuclear structure, organization, and oncogenesis. *Journal of Gastrointestinal Cancer*, 42(2), 112-117. <https://doi.org/10.1007/s12029-011-9253-5>
- Sirinukunwattana, K., Raza, S. E. A., Tsang, Y.-W., Snead, D. R. J., Cree, I. A. & Rajpoot, N. M. (2016). Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. *IEEE Transactions on Medical Imaging*, 35(5), 1196-1206. <https://doi.org/10.1109/TMI.2016.2525803>
- Urine Cytology - Mayo Clinic*. (s.f.). Consultado el 6 de septiembre de 2020, desde <https://www.mayoclinic.org/tests-procedures/urine-cytology/about/pac-20385279>
- Zink, D., Fischer, A. H. & Nickerson, J. A. (2004). Nuclear structure in cancer cells. *Nature Reviews Cancer*, 4(9), 677-687. <https://doi.org/10.1038/nrc1430>