Summary

This paper turns patient encounters into a matrix (image) to which they use a CNN to learn the differences of hospital stays between patients with nonmelanoma skin cancer (NMSC) and those without. They used a cohort from Tiwan to learn these differences and trained a discriminant classifier (AUC: 0.89) despite their small training size (n ~= 7,500).

My reaction

I thought this was a very interesting idea. However, I thought the paper was lacking in major areas. First, their methods section lacked any justification for (seemingly arbitrary) decisions they made. Second, I thought that their explanation of how they did things was confusion and I doubt I could recreate their model if I had the same data. Finally, and perhaps most importantly, they say, “This … prediction tool may facilitate determination of which patients are likely to develop NMSC, potentially allowing clinicians to intervene …”, however, if I understand correctly, a positive sample in their training data already has been diagnosed with the disease, making this model a secondary prevention model. For example, this model does not predict if a patient will get NMSC in the next year, but predicts if a patient’s labs and comorbidities signal that they have this disease, which goes against what they state they want to use the model for.

Questions

Did anyone else get confused reading the methods, or did I not get what I was supposed to out of it?

How did the filtering process work? It seems like a reach to go from 2 million samples to less than 10,000 just due to the inclusion and exclusion criteria they listed.

Why did they need to match positive and negative samples?

* “For the control group, the index date was either matched with the cancer index date or the last day available in the database.”