CS 766: PROJECT PROPOSAL

Kendall Park (kpark229@wisc.edu)

Ying Ji Chuang (ychuang6@wisc.edu)

Lanston Chu (hchu34@wisc.edu)

Mehmet F. Demirel (demirel@wisc.edu)

---

1. **Briefly explain what problem you are trying to solve.**

   We would like to reconstruct a 3D volume of radiodensities from a single 2D chest X-ray.

2. **Why is this problem important? Why are you interested in it?**

X-rays are one of the most common imaging modalities used in modern medicine. An X-ray film contains a two-dimensional transposition of three-dimensional radiodensity between the X-ray source and the film. The two-dimensional nature of X-rays creates certain limitations on interpretation. Computed tomography (CT) imaging, on the other hand, computes three-dimensional radiodensity volumes from a series of X-rays performed around the patient. CT imaging is used in cases where a plain X-ray film is insufficient. CT imaging is more expensive and exposes patients to much greater levels of radiation than standard X-ray imaging (around two orders of magnitude greater). In the case of CT imaging, the benefits that the extra dimensionality provided must be weighed against the risk of increasing a patient's radiation exposure and the additional cost of the procedure. We would not expect a 3D reconstruction of a 2D X-ray image to match the fidelity of its corresponding CT volume. It may, however, be possible that a 3D reconstruction could be "good enough" to increase diagnostic accuracy for particular diseases, allowing the patient to avoid unnecessary exposure to CT-sourced radiation and the additional expense of that modality.

It is also worth noting that there are many situations where X-ray imaging is the *only* radiological imaging modality readily available. This can often be the case in field hospitals and areas of the developing world. A 3D reconstruction of an X-ray where no CT imaging is available may prove helpful in certain diagnostic situations (eg, trauma surgery).

Finally, current machine learning (ML) vision tasks (recognition, segmentation, etc.) on X-ray images treat the X-ray films as a 2D picture—even though it contains 3D data. When radiologists interpret X-rays, they use their internal knowledge of 3D human anatomy to guide their interpretation. They are in essence combining image inputs and priori knowledge of human anatomy to achieve their diagnosis. Intuition would suggest that providing some kind of 3D understanding of a 2D X-ray film (via the embeddings from a 3D reconstruction process of that X-ray film) could improve accuracy on standard machine learning image tasks like recognition and segmentation. While testing this hypothesis is outside the scope of our project, it is a potential avenue for future research.

3. What is the current state-of-the-art?

No one has tried to solve this problem in the context of chest X-rays and chest CTs. However, there exist several works that can perform 2D-to-3D reconstruction in different but similar domains.

Henzler et al. (2018) has performed 2D-to-3D volume reconstruction of mammalian crania using a Convolutional Neural Network (CNN) with an encoder-decoder structure. However, they focused on optimizing surface structure and the generalizability of the algorithm (reconstruction any angle, wide variety of species). In constructing 3D volume from 2D images, Jackson et al. (2017) proposed a Volumetric Regression CNN (VRN) built upon paired 2D images and 3D facial models or scans. Their model is able to reconstruct 3D human faces from a single photograph taken at any angle. Karade and Ravi (2015) used a "bone template reconfiguration" algorithm involving Kohonen self-organizing maps to simulate 3D surface geometry of femur from biplane X-ray images.

Since we are interested in 2D-to-3D reconstruction, our training dataset, consisting of 3D CT scans, would also need an equivalent of the input X-ray images. We plan to generate synthetic X-rays for each of the 3D CT scans. Moturu and Chang (2019) proposed a method to create synthetic frontal chest X-rays using ray-tracing and Beer's Law from chest CT scans. Moturu and Chang's research focused on designing a neural network that can detect lung cancer, thus their methods also involved randomized nodule generation. Teixeira et al. (2018) also presented a framework that generates synthetic X-ray images from the surface geometry of the thorax and abdomen. The resulting X-ray images were only intended to serve as an approximation of the true internal anatomy. Other than that, we have not found any other papers that addressed the generation of synthetic X-ray images from CT scans specifically in the chest region. Henzler et al. (2018) also generated synthetic x-rays, but their algorithm simply flattened the CT scans at various angles and did not consider the point-source aspect of X-ray imaging. The only paper we found that included the point-source aspect of X-ray imaging was that of Moturu and Chang.

4. **Are you planning on re-implementing an existing solution, or propose a new approach?**

Several different existing ideas would be combined to solve a new problem in a different domain space, namely the reconstruction of 3D volume of human chest by a single X-ray image. While the work of Henzler et al. (2018) is focusing on mammalian crania, we believe their method may be applicable to medical imaging. While Jackson et al.'s (2017) work on human face reconstruction also provided inspiration for this project, the inputs and outputs of their work focus on visible light and surface geometry.

We plan on using Henzler et al. model architecture as a stepping-off point. After which we will iteratively modify the neural network structure to improve its efficiency in the human chest X-ray setting. We expect our final model architecture to be novel.

We will be using Moturu and Chang's (2019) approach in utilizing ray-tracing and Beer's Law to synthesize X-ray images from CT dataset. We may augment the synthetic X-rays with distortions and noise to better simulate real X-rays.

5. **If you are proposing your own approach, why do you think existing approaches cannot adequately solve this problem? Why do you think you solution will work better?**

As mentioned above, it is the first attempt to use a single X-ray image for the reconstruction of a 3D radiodensity volume of the human chest. Previous works of the 2D-to-3D problem before the 2010's did not involve deep learning, while new approaches since 2017 either apply to different fields of interest (e.g. animal anatomy, face recognition), work on different types of input (e.g. image of visible light), or output (e.g. surface geometry).

Among the relevant papers mentioned above, only Teixeira et al. (2018) and Moturu and Chang's (2019) approaches used CT scans as the training dataset. Neither of those author groups used more than 100 CT studies. In this project, we will be using over 700 chest CT scans as our training data, which could potentially increase the accuracy of the CNN to produce 3D CT of adequate accuracy from a single 2D X-ray.

6. **How will you evaluate the performance of your solution? What results and comparisons are you eventually planning to show? Include a timeline that you would like to follow.**

Similar to Henzler et al. (2018), a hold-out test set of CT will be used to generated synthetic X-rays. For those hold-out, the Euclidean loss (i.e. L2 metric) will be computed between the model's 3D reconstruction of those synthetic X-rays and the original CT volumes. It is important to note that most voxel-level quantitative similarity metrics are limited in their ability to express clinically-relevant image similarity, as one purpose of the 3D reconstruction is to assist radiologists to interpret the 2D X-ray. For instance, accurate reconstruction of certain regions (organs structure, vasculature, etc) is more important than the reconstruction of other areas (clothing, surrounding air, etc). A reconstructed volume with high L2 metric may be inferior to a lower-scoring volume if the latter can be more accurately interpreted by radiologists.

Given these considerations, we would access the results of our model by using clinically-relevant measures of performance. For example, given volumes with lung nodules, we might compare the distance of the centroid of the nodule between the predicted and the original volume. Some other ideas include measurements of heart size, carina placement, and spinal column alignment. We will explore and finalize clinically relevant features for evaluation upon reviewing our results (ideally, with a radiologist).

REFERENCES

Chang, A., & Moturu, A. (2019). *Detecting Early Stage Lung Cancer using a Neural Network Trained with Patches from Synthetically Generated X-Rays*.

Henzler, P., Rasche, V., Ropinski, T., & Ritschel, T. (2018). Single-image tomography: 3D volumes from 2D cranial X-Rays. *Computer Graphics Forum*. https://doi.org/10.1111/cgf.13369

Jackson, A. S., Bulat, A., Argyriou, V., & Tzimiropoulos, G. (2017). Large Pose 3D Face Reconstruction from a Single Image via Direct Volumetric CNN Regression. In *Proceedings of the IEEE International Conference on Computer Vision*. https://doi.org/10.1109/ICCV.2017.117

Karade, V., & Ravi, B. (2015). 3D femur model reconstruction from biplane X-ray images: a novel method based on Laplacian surface deformation. *International Journal of Computer Assisted Radiology and Surgery*. https://doi.org/10.1007/s11548-014-1097-6

Teixeira, B., Singh, V., Chen, T., Ma, K., Tamersoy, B., Wu, Y., Balashova, E., & Comaniciu, D. (2018). Generating Synthetic X-ray Images of a Person from the Surface Geometry.