Multimodal Deep Learning for Disease Trajectory Forecasting

Author(s): Surya Teja Devarakonda, Yeahuay (Joie) Wu

Problem Statement: Medical datasets typically contain data from multiple modalities including longitudinal studies, imaging data (i.e X-ray and MRI), and time series data from activity monitoring. Multimodal Deep Learning (MDL) deals with models that can leverage data from multiple modalities to make inferences of diagnostic significance.

An example of a condition that might benefit from MDL is Knee Osteoarthritis (KOA). KOA diagnosis involves inference from modalities including time series data like accelerometer data, structured data like demographics, and image data.

One of the most recent works in MDL is a technique called ShortFuse [5], which boosted the accuracy of deep learning models for time-series data by fusing them with structured covariate data and obtained state-of-the-art performance for two applications: predicting the progression of KOA, and predicting surgical outcomes for Cerebral Palsy. However, ShortFuse doesn't take advantage of medical scans which hold valuable diagnostic information. The objective of our project is to extend ShortFuse by incorporating images into its prediction model.

Methodology: For our project, we will be building deep learning models involving LSTM and CNN architectures. The models will leverage the regularized risk minimization (RRM) framework. We will use the Pytorch library and Google Cloud Computer Clusters to train our deep learning models.

Related work: Previous methods of predicting disease progression involve leveraging hand-engineered features such as histograms [2] and spectral data [4], or techniques like Principal Component Analysis (PCA) [3] and Dynamic Time Warping [1]. Furthermore, they do not exploit inter-modal dependencies since they model each modality independently.

ShortFuse was created to address these issues. In ShortFuse, covariate and time series data are integrated into the network in the initial layers to allow the model to learn relevant intermodal dependencies. ShortFuse used hybrid layers called hybrid convolutions and hybrid LSTMs to obtain state-of-the-art results in a couple of disease progression forecasting tasks. Another line of research utilizes deep neural networks with only radiography images to forecast disease progression [6][7].

To the best of our knowledge, there is no prior work on using all three modalities to forecast disease progression. In fact, there is no work combining these modalities towards any general ML problem.

Data Sets: We will start our work by using two publicly available datasets for knee osteoarthritis:

1. MOST, OAI: Contain clinical data, radiograph images, accelerometry, MR Images for KOA.

In the event that we face unforeseen difficulties with these datasets, we have listed a couple more medical datasets that contain multi-modal data which could also be used for our project.

2. ADNI: Contains clinical, biomarker, genome, and MRI scans of 819 subjects.
3. OASIS-3: Contains clinical data and MRI scans taken from a total of 1098 subjects.

Experiments:

1. To benchmark our experiment, we will implement ShortFuse, LateFuse, and two other models which only utilize longitudinally collected medical images.
2. We will first incorporate images into the ShortFuse framework using the 'LateFuse' Approach. This involves the following: 1) extracting features from the image data using a CNN 2) extracting features from time series and covariate data using the ShortFuse model 3) concatenating the features of the three modalities and passing them to a fully connected MLP for classification.
3. We will then incorporate images by using the ShortFuse approach, which involves developing hybrid-CNN layers to fuse the images in the initial layers of the network itself with the time series and covariate modalities. One idea to achieve this is to extract feature vectors from the images using another trained CNN model (or a pre-trained model trained on Imagenet), and feed it to the hybrid CNN/LSTM cells, in a similar manner as the covariate data.

Overlap Statement:

This project is done under the guidance of Prof. Madalina Fiterau. Our project will not build on past work performed by any of the team members. For Joie, who is a Ph.D. student, this project is not associated with any R.A-ship, and the work done will not be used for any other courses, including independent study.

Collaboration Plan:

1. For our baseline, we intend to implement at least 4 state-of-the-art techniques to provide a comprehensive evaluation of the models.
    a. Joie will implement [6] and [7], which use only medical radiography images in their prediction models.
    b. Surya will implement the models described in the ShortFuse paper.
2. While extending the ShortFuse model to incorporate images. Joie will implement the CNN to extract the features from medical images. Then, using the trained CNN, Joie will work on the ‘late-fuse’ approach, while Surya will work on the ‘short-fuse’ approach, as described in the Experiments section.
3. We will split work for writing the report 50/50 and write for the sections that we each worked on.

References:

1. A. Kale, N. Cuntoor, B. Yegnanarayana, A. Rajagopalan, and R. Chellappa. Gait analysis for human identification. In International Conference on Audio-and Video-Based Biometric Person Authentication, pages 706–714. Springer, 2003.

2. D. D. Dunlop, J. Song, P. A. Semanik, R. W. Chang, L. Sharma, J. M. Bathon, C. B. Eaton, M. C. Hochberg, R. D. Jackson, C. K. Kwoh, et al. Objective physical activity measurement in the osteoarthritis initiative: are guidelines being met? Arthritis & Rheumatology, 63(11):3372–3382, 2011.
3. P. Federolf, K. Boyer, and T. Andriacchi. Application of principal component analysis in clinical gait research: identification of systematic differences between healthy and medial knee-osteoarthritic gait. Journal of Biomechanics, 46(13):2173–2178, 2013.

4. J. Lee, R. W. Chang, L. Ehrlich-Jones, C. K. Kwoh, M. Nevitt, P. A. Semanik, L. Sharma, M.-W. Sohn, J. Song, and D. D. Dunlop. Sedentary behavior and physical function: objective evidence from the osteoarthritis initiative. Arthritis Care & Research, 67(3): 366–373, 2015.

5. Fiterau, Madalina & Bhooshan, Suvrat & Fries, Jason & Bournhonesque, Charles & Hicks, Jennifer & Halilaj, Eni & Ré, Christopher & Delp, Scott. (2017). ShortFuse: Biomedical Time Series Representations in the Presence of Structured Information.

6. Tiulpin, A., Thevenot, J., Rahtu, E., Lehenkari, P. and Saarakkala, S. (2018). Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach.

7. Norman, Berk & Pedoia, Valentina & Noworolski, Adam & Link, Thomas & Majumdar, Sharmila. (2018). Applying Densely Connected Convolutional Neural Networks for Staging Osteoarthritis Severity from Plain Radiographs. Journal of Digital Imaging. 10.1007/s10278-018-0098-3.