Time-Distanced Gates in Long Short-Term Memory Networks -- Supplementary Materials

Abstract. In this material, we describe how the demonstration dataset, Tumor-CIFAR, is created. The is the supplementary document of the DLSTM paper. The references and mathematic computing are provided. In addition, a comparison about our method and linear interpolation of registration is discussed. The code and datasets are publicly available at https://github.com/MASILab/tumor-cifar.

1 Simulation: Tumor-CIFAR

1.1 Introduction

Medical images are a precious and rare resources. It is hard to share large amounts of lung CT scans, especially for longitudinal data. CIFAR10 [1] dataset consists of 60K color natural images with the size of 32×32, which is convenient to verify the effectiveness of algorithms without excessive computing cost. Compared with MNIST [2], CIFAR10 consists more diverse image patterns. We manually add synthetic "nodules" to CIFAR10 with approximate clinical knowledge of lung nodules. Motivated by [3], we model the growth rate of malignant nodules at three times than the benignant one. This information guides us to simulate the Tumor-CIFAR dataset.

1.2 The Implementing of Tumor-CIFAR

Step 1. Create 5 time-intervals $X = [x_0, x_1, x_2, x_3, x_4]$ with absolute of Gaussian distribution:

$$X \sim |N(\mu, \sigma^2)|$$

where $\mu = 1.67$ and $\sigma^2 = 1$ for version 1, $\mu = 5$ and $\sigma^2 = 3$ for version 2, and $N(\mu, \sigma^2)$ represents is the Gaussian distribution with μ as mean and σ^2as variance. The operation $|\cdot|$ is the absolute operation.

Step 2. Create 5 time-points $T = [t_0, t_1, t_2, t_3, t_4]$ by boosting the 5 time-intervals, where

$$t_i = \sum_{k=0}^{i} x_k$$

Step 3. Create the nodule growth rate g, where $g \sim |N(1, 0.2)|$.

Step 4. Randomly choose threshold *th,* where *th* belongs to uniform distribution $th \sim U(0,1)$. If *th* < 0.5 we regard this subject as BENIGN, otherwise, this subject is MALIGNANT.

We create two different versions in step 5 according the rule learned from [1] that the growth rate of malignant nodules is about 3 times than the benign one (we call this NODULE RULE in the following).

Step 5. For version 1, the growth rate *g* of MALIGNANT subjects would increase to $3 \times g$ to match the NODULE RULE. For version 2, the time points *T* of MALIGNANT subjects would reduce to *T/3* to match the NODULE RULE.

Step 6. Randomly choose two nodule locations, and generate the nodules. The nodule location (x_l, y_l) is chosen by

$$(x_l, y_l) \sim \big(U(7,25), U(7,25)\big)$$

Note that the image size is 32×32. Then the nodule size s_i is computed by

$$s_i = t_i \times g$$

And the intensity of the nodule pixel p_i^{nodule} is calculated by

$$p_i^{nodule} = (1 + \alpha \cdot s_i) \cdot p_i^{origin}$$

where *i* is the image index, and p_i^{origin} is the pixel value before adding nodule, $\alpha = 0.1$.

Step 7. Finally, we add 10% salt-and-pepper noise among all pixels.

1.3 Example images of Tumor-CIFAR

Fig. 1 presents image examples. We can see both salt-and-pepper noise and "nodules" on the images. In version 1 Tumor-CIFAR, images are modeled with same time interval distribution, but with different nodule sizes between benign and malignant. While in version 2, the data with same nodule size distribution, different time intervals between benign and malignant.
The source code of generating these datasets, the raw datasets (some examples), and related raw files (including time intervals and nodule sizes) can be found at the link: https://github.com/MASILab/tumor-cifar.

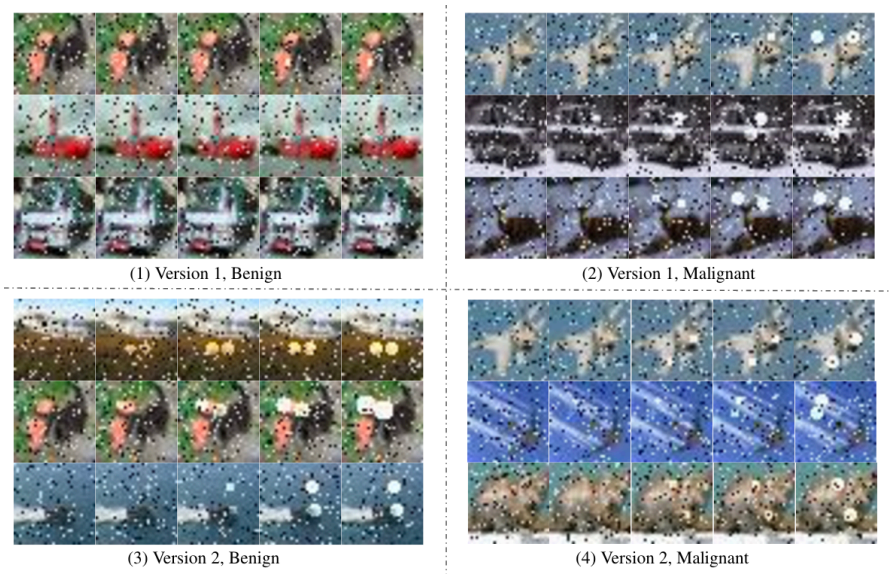

Fig. 1. Example images from two version Tumor-CIFAR datasets. In version 1, the nodule size of malignant image is apparently bigger than the benign ones because the time interval distributions are the same between Benign and Malignant. Since nodule size distribution are the same in version 2, we barely can see the nodule differences according only image data.

2 Linear Interpolation of registered lungs in longitudinal imaging

We introduce a LSTM variant in the manuscript, termed as Distanced LSTM (DLSTM), to address the lung cancer detection with longitudinal imaging. We introduce a discussion of the comparison of our method and another straightforward baseline: linear interpolation of registered lungs in longitudinal imaging under the context of LSTM.

2.1 Step to Prepare the Data

Step 1. Do lung registration for each neighbor images.

For each longitudinal pair, we register the prior image I_0 to the last image I_1 using NiftyReg [4] for each longitudinal subject using affine registration.

Step 2. Evenly space images with linear interpolation.

After obtaining the registered image I_0', we get interpolated $\widehat{I_0'}$ using:

$$\widehat{I_0'} = \frac{t}{T_1 - T_0}(I_0' - I_1) + I_1$$

where T_1, T_0 are the time stamp of images I_1, I_0. t is the self-defined time step. In our implementation, we follow the national lung screening trial guideline (https://www.cancer.gov/types/lung/research/nlst) that the intended screening interval is one year. Then, the $[\widehat{I_0'}, I_1]$ are used for the input of model.

2.2 Results and Discussion

Table 1. Experimental results on clinical datasets (%, average (std) of cross-validation)

Method	Accuracy	AUC	F1	Recall	Precision
LSTM	86.27(1.29)	90.27(1.15)	74.17(2.47)	69.73(2.62)	79.56(5.69)
Inter-LSTM	*86.06(1.66)*	*90.52(0.61)*	*73.38(2.42)*	*67.19(4.45)*	*81.45(5.61)*
DLSTM1	**86.97(1.45)**	91.17(1.53)	**76.11(2.68)**	**72.71(2.38)**	80.04(5.18)
DLSTM2	86.98(1.20)	**91.41(1.51)**	75.54(1.67)	71.24(5.01)	81.22(6.11)
DLSTM3	85.99(1.13)	91.10(1.69)	74.68(2.89)	70.51(6.07)	80.23(5.55)
DLSTM4	86.91(1.37)	91.07(1.28)	75.85(1.94)	72.39(3.65)	80.21(6.34)

Table 2. Experimental results on cross-dataset test (%, external-validation)

Method	Accuracy	AUC	F1	Recall	Precision
Train and Test both on longitudinal subjects					
LSTM	85.89	83.80	57.32	57.69	56.92
Inter-LSTM	*85.90*	*85.43*	*55.62*	*52.5*	*59.15*
DLSTM1	88.63	**89.05**	68.24	74.36	63.04
DLSTM2	**89.47**	88.62	69.14	70.00	68.29
DLSTM3	**89.47**	87.39	67.11	63.75	**70.83**
DLSTM4	89.26	88.42	**69.46**	72.50	66.67

As shown in Table 1 and Table 2, the linear interpolation after lung registration for longitudinal images (termed as Inter-LSTM in the tables) not necessarily improves the baseline LSTM even though the time stamp information is included in the linear interpolation step. To our knowledge, compare with our method, the Inter-LSTM may introduce the following obstacles: (1) the registration might distort the nodule shape and intensity, which may influence the classification, (2) The nodule changes (e.g., nodule size, nodule density) are not necessarily linear growing overtime, the mandatory linear interpolation might lose original information, (3) The effects of linear interpolation highly depends on the lung registration, noises can be easily introduced if the registration is not perfect.

3 References

1. Krizhevsky, A. and G. Hinton, Learning multiple layers of features from tiny images. 2009, Citeseer.
2. LeCun, Y., C. Cortes, and C.J.J.U.h.y.l.c.e.m. Burges, The MNIST database of handwritten digits, 1998. 1998. 10: p. 34.
3. Duhaylongsod, F.G., et al., Lung tumor growth correlates with glucose metabolism measured by fluoride-18 fluorodeoxyglucose positron emission tomography. 1995. 60(5): p. 1348-1352.
4. Modat, et al., Fast free-form deformation using graphics processing units. Computer Methods And Programs In Biomedicine,98(3), 278–284. (2010).