Forecasting 0-2 Hours Short-duration Heavy Rainfall based on Deep Learning

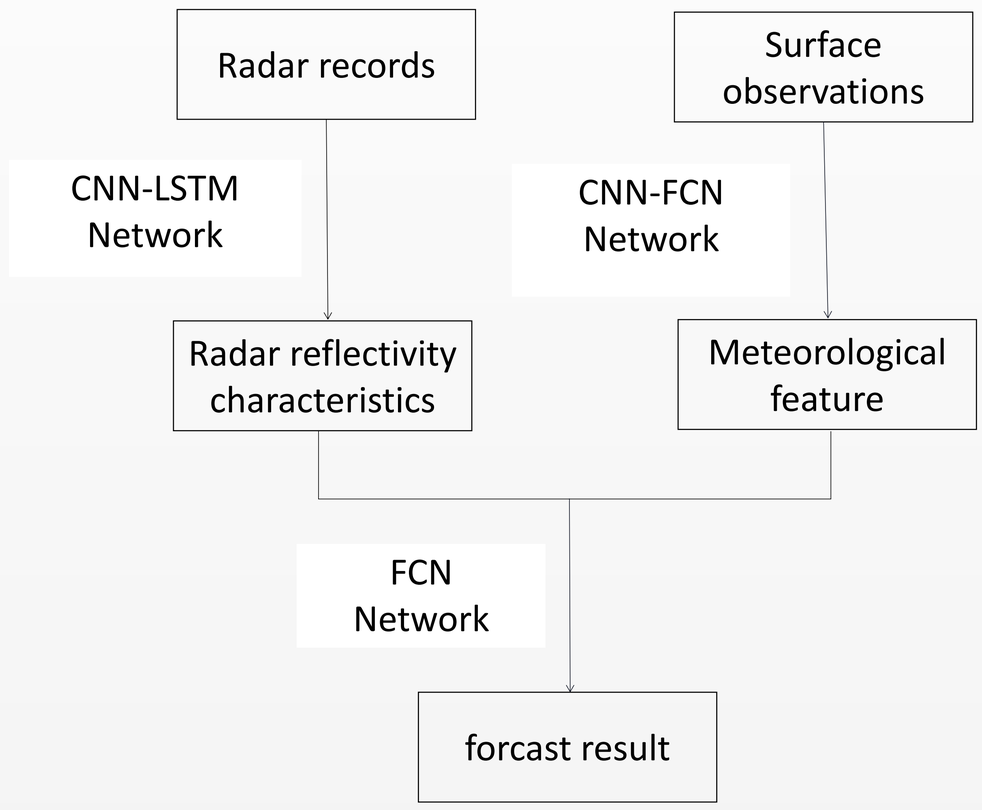
Jie XIAHOU，An XIAO , Daoyang NIE , Fulei ZHU
(Jiangxi Meteorological Observatory)
Email: xiahoulove@126.com

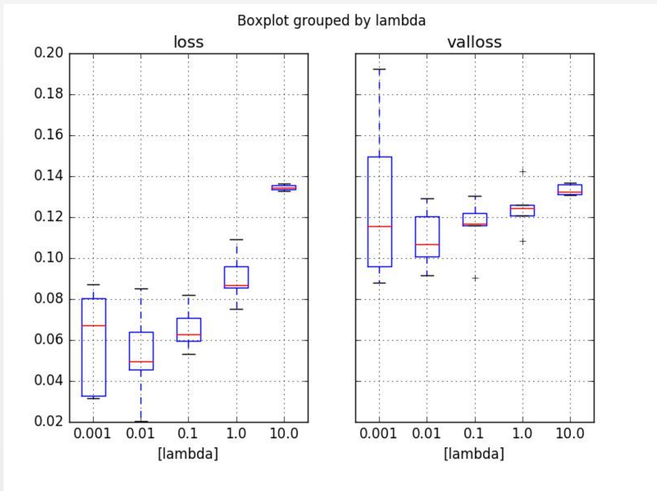
Introduction

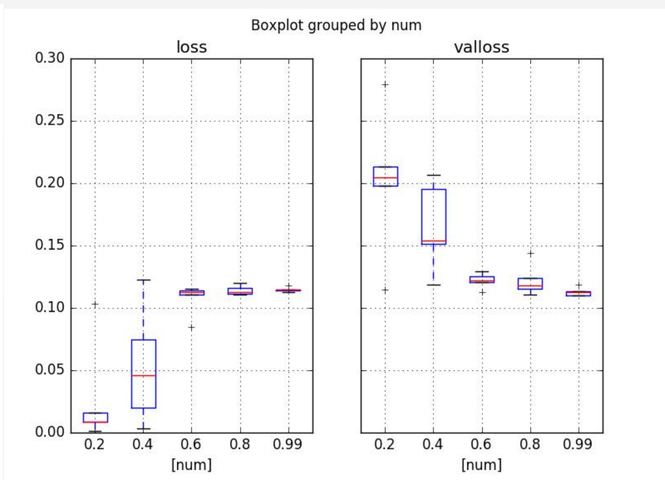
Short-duration heavy rainfall weather is a convective weather with strong disaster causing. A large amount of precipitation in a short time can easily lead to a variety of meteorological disasters.However, the forecasting of short-duration heavy rainfall weather is faced with great challenges.In fact, the forecast time of short-duration heavy rain is actually short, mainly 0-2 hours near forecast, and in longer time potential forecast can be made.Radar echo can be used to capture the convective weather system, and the atmospheric environment factors are in affect with the development of the convective weather system at its all stages.Based on the above considerations, this paper combines the two observation datasets when constructing the model.In the model training stage, radar data and surface meteorological data are trained separately to find their own different information, then they are jointly trained and enter into the final layer for obtaining the forcast results of short-duration heavy rainfall.The result is compared with the traditional radar echo only model.

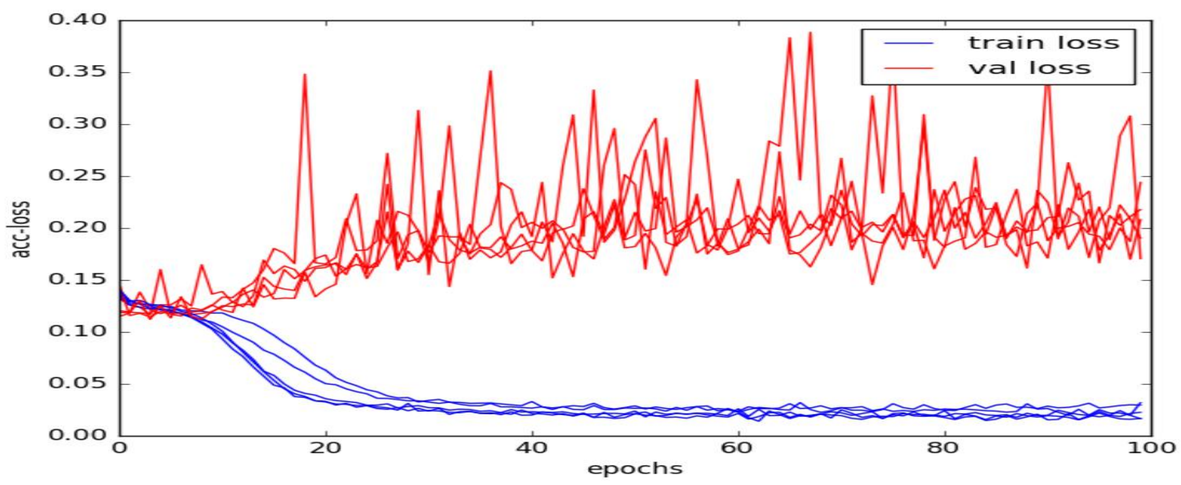
Model Structure

Convolution processing is carried out for continuous observation of radar data. On this basis, the output features of convolution are taken as the input of the short-time memory model, which are used to output the radar forecast features and serve as the input of subsequent models.At the same time, feature extraction is performed on surface observation data sets using Convolution Neural network-fully Connected Network (CNN-FCN) model. The output can also be used as the input of subsequent models.At last,The combined output method uses Fully Connected neural Network (FCN) to give the two appropriate weights to obtain the short-duration heavy precipitation forecast results.









Build Model

The deep learning model based on multi-layer neural network can achieve perfect fitting on training data through a large number of violent iterations, but the prediction of test data is not good as the train data. This is the key problem of model building in deep learning. Overfitting can be caused by many factors.Different deep learning models have different reasons for overfitting, depending on the data and the model builder. In order to improve the prediction skills of the model, detailed diagnosis to the causes of overfitting must be made. The common diagnosis method of overfitting is drawing learning curves. In this paper, the learning curves are used to make diagnosis for the newly established deep learning model to find out the cause of overfitting.The model parameters were determined by repeated experiments.The curve graph and boxplot on the left show the relationship between the regularization coefficient, data volume, iteration times and the loss function in five experiments.

model	Forcast hour	POD	FAR	hit	miss	False alarm	Correct nagetive	total
CNN_LSTM (radar+surface)	1h	0.677	0.634	107	51	185	2896	3239
	2h	0.673	0.662	111	54	217	2857	3239
CNN_LSTM (radar)	1h	0.707	0.614	118	49	188	2884	3239
	2h	0.452	0.681	76	92	162	2909	3239
NWP	1h	0.583	0.704	88	63	209	2723	3083
	2h	0.411	0.695	62	89	141	2791	3083

Score

The score during the main rainy season from April to August in 2020 show that within 0-2 hours, the POD and FAR of the deep learning model based on radar data performance better than the regional NWP, but the improvement is limited for the second hour. However, with the surface meteorological elements added to the deep learning model, the POD of the model has been greatly increased.