Chen, Li; Fei, Hongxiao; Li, Haifeng; He, Jiabao; Tan, Fengyun
Analysis on information geometric measurement of internal transfer of deep neural network.
(Chinese. English summary) [Zbl 1438.68151]

Summary: When deep learning is used to deal with the computer vision tasks, under little number of new task data, the pre-trained model weight based on very large data is trained as an initial weight to get better generalization ability. At this point, former explanations are based on the intuitive analysis and lack of reasonable mathematical methods. In this paper, deep neural network, which trains on internal layers with fixed structure, changed into internal transfer ability in deep neural network. The changes of learning process are formalized into a mathematical expression. Considering the influence of the data set on the training process, the information geometric analysis method is used to determine the metrics and connections over manifolds of different data sets, which can realize the embedding mapping between different data sets. At the same time, the change of parameter space is also put into a manifold space to explore its common influence on learning process. Finally, a mathematical explanation is provided for the internal transfer phenomenon. Meanwhile, after the analysis and experiments, the process of internal transfer is identified as a change which can make the network search for optimal search in a wider space. Therefore, the model can obtain a relative better solution in learning process.

MSC:
68T07 Artificial neural networks and deep learning
62B11 Information geometry (statistical aspects)
68T45 Machine vision and scene understanding

Keywords:
deep learning; transfer learning; information geometry

Full Text: DOI