Mengxi Wang

Personal Information

Status:	**MS Student**
Program:	**Computer Science and Engineering**
School:	**Tandon School of Engineering, New York University**
Website:	**https://www.linkedin.com/in/mengxi-w-190b55116/**
RA Period:	**From 2016-01 to 2017-01**

Biography

I'm a senior software engineer at Comcast. Before that, I was a research assistant in NYU Multimedia and Visual Computing Lab, advised by Professor Yi Fang. I am broadly interested in 3D Computer Vision, Pattern Recognition and Deep Learning.

M.S. Thesis: Shape Correspondence on Riemannian Manifolds by Multiple Shape Descriptors

1 Description

This thesis provides a novel method to perform the shape correspondence, which employs multiple shape descriptors as the input data and redesigns the intrinsic neural network architecture based on the geodesic convolution neural network. The new method concatenates multiple shape descriptors based on their characteristics to ensure the comprehensiveness of the input data. In addition, the new method redefines the network architecture, which removes unnecessary layers and introduces the dropout layer, batch normalization layer and concatenating layer to achieve dense predictions. The dropout layer prevents model from overfitting by randomly disconnecting two neurons. The batch normalization layer accelerates the training process by reducing internal covariate shift. The concatenating layer connects the outputs from three geodesic convolution layers to decrease the loss of information during the training process. This experiment shows the new method has state-of-the-art performance in shape correspondence and provides much stronger predictions than other methods in multiple tasks. Even though the new method has issues in finding correspondence for shapes from different categories, it still offers valuable references for future work. and a state-of-the-art convolutional-neural-network object detector. Experiments show that our framework significantly improves the accuracy of GPS localization and is capable of providing semantic labels in the 3-D domain at real-time.

2 Method

In this thesis, we provides a novel method to perform the shape correspondence, which employs multiple shape descriptors as the input data and redesigns the intrinsic neural network architecture based on the geodesic convolution neural network. The new architecture removes the FTM layer and adds the batch normalization layer and dropout layer. Additionally, it incorporates the concatenating layer into the model. Through the recombination of these layers, this

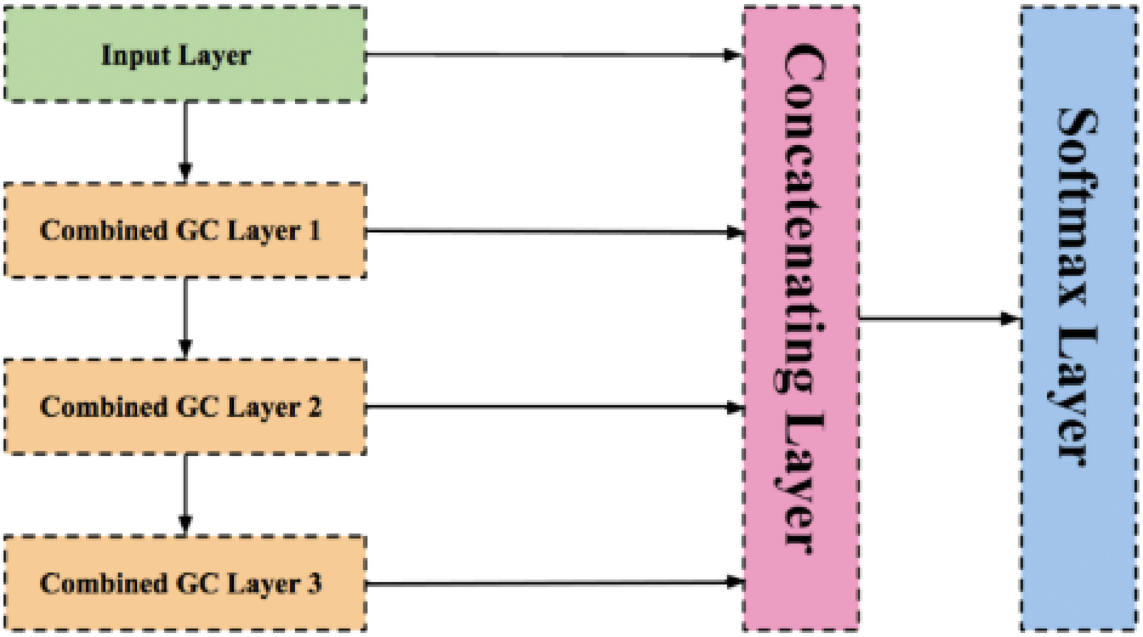

Figure 1: Our complete intrinsic neural network architecture.

paper proposes the new structure shown in Figure.1. An input layer is followed by three combined GC layers. The output of these four lay- ers feeds into the concatenating layer and then goes to the softmax layer. The final result is the prediction of the shape correspondence. Figure.2 shows the combined GC layer. Each GC layer is connected to a dropout layer. At the same time, an angular max-pooling layer and a batch normalization layer follow the output of the GC layer. Because these layers are always combined, these four layers can be referred to as the combined GC layer. In the new model, the combined GC layer is used to replace the continuous four layers, which are the GC layer, dropout layer, angular max-pooling layer and batch normalization layer. Three combined GC layers follow the input layer, and the outputs of these three layers feed into the concatenating layer. Finally, the output of the concatenating layer is processed by the softmax function to obtain the final probability distribution vector. The main point is the appearance of the concatenating layer. The reason is that in the training process, the feature map of each convolution layer becomes abstract, which implies a loss of information. By adhering to a coarse output, a dense feature map can be obtained, which leads to a dense prediction.

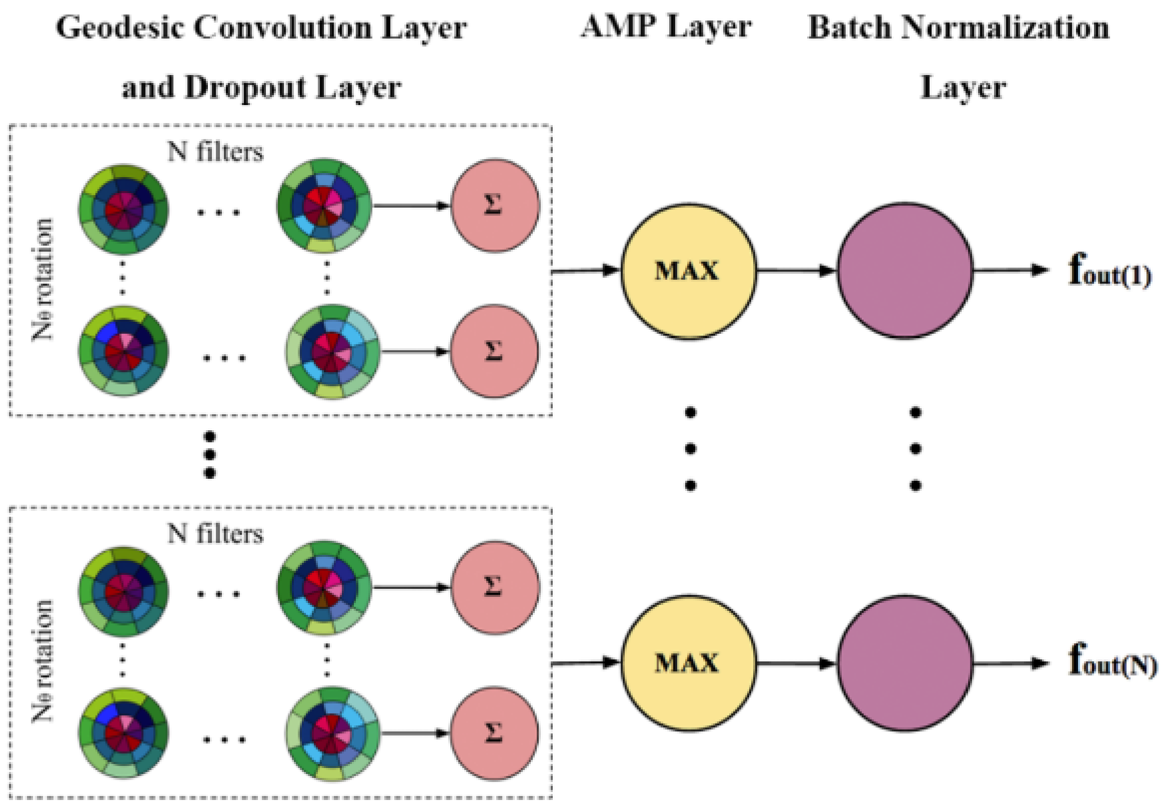

Figure 2: Combined GC layer.

Method	Complete Model	No Multiple Descriptors	No Concatenating Layer	No Concatenating Layer & Multiple Descriptors
Precision	73.61 %	69.56 %	67.48 %	60.89 %

Table 1: Experimental results of the method comparison in the identical class.

3 Results

In this section, we present the final experiment results. Two datasets are used in the experiment in shape correspondence. One is the TOSCA dataset, which contains high-resolution three-dimensional non-rigid shapes. The TOSCA dataset contains a total of eighty objects, including eleven cats, nine dogs, three wolves, eight horses, six centaurs, four gorillas, twelve female figures, and two different male figures, one of which containing seven poses, and the other containing twenty poses. Typically, the number of vertices is approximately 50,000. In the same class, the models have the same triangulation and an equivalent number of vertices numbered in a compatible way. The other dataset is the FAUST dataset,

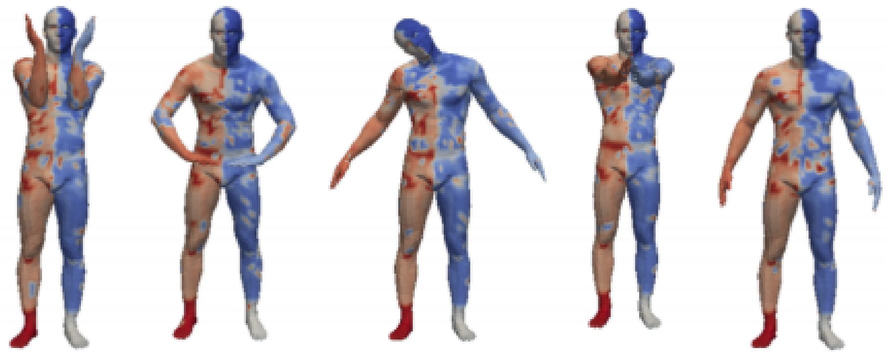

Figure 3: Illustration of the complete neural network architecture.

which contains 300 real, highresolution human scans. These human models have dense ground-truth correspondences, and each human model has 6890 vertices. A previous study resampled human models in TOSCA to ensure that those human models have the same number of vertices with the shapes in FAUST. In Table 1, our method clearly has a higher accuracy than the other methods. This finding proves that our architecture has the sufficient ability to accom- plish shape correspondence. Relative to GCNN3, the new structure improves the result by 13.34%. Figure.3 shows the complete process and the neural network architecture. Fig- ure 4.2a, 4.2b, 4.2c and 4.2d illustrate three shape descriptors (the geometry vector, SHOT descriptor, HKS) and the final output result. The geometry vector can indicate the coarse correspondence of the whole body, but it cannot clearly distinguish each part, such as the left and right hand, shank, and thigh. In contrast, the SHOT descriptor can accommodate the detailed sections. For example, the model of the SHOT descriptor has the asymmetric distribution of spots with a distinct color depth in the body. In addition, the stripe on the left leg is distinct from the stripe on the right leg. Relative to the geometry vector and SHOT descriptor, the HKS is better at determining main parts, such as hands and feet, which helps refine the classification. The model of final output result shows that the new method can distinguish symmetric parts and make good correspondence of shapes.