

# **Prayer (Salat) Posture Classification and Analysis Using Convolutional Neural Network, AlexNet and GoogleNet Deep Learning Neural Networks**

**Course: Artificial Intelligence**

**Presented To:**

**Dr. Muhammad Farrukh Shahid**

**By:**

**Salman Ahmed Khan (19K-1043)**

**Sadeem Sattar (19K-1102)**

**Abdullah Tilal Khan (19K-1103)**

**Dated: 16 May 2022**

# Abstract:

Prayer is the essential part in every religion existing on the planet. In the religion Islam, prayer that is Salat is the fundamental part. It is the second pillar of Islam that the believers have to obey. It is the essential worshipping activity that the believers have to perform five times a day. Salat consists of multiple human gesture perform by a worshipper. It contains a predefined set of posture activity that must be followed precisely. However, for several people, these postures are not accurately performed due to new to the religion or even learned in an incorrect manner. Furthermore, the duration of each posture has to be balanced. Therefore, to cater this issue, we propose to develop an artificial intelligence model that aims to identify the posture through an image. We classify three prayer posture that includes Ruku, Sujud and Qiam. The object is accomplished through Deep Learning Models including Convolutional Neural Network, AlexNet and GoogleNet. Furthermore, we analyze the accuracy obtained by these deep neural networks. Experiment results demonstrate that the precision attains for a training dataset of 457 images and 196 test images of different postures by CNN is 90.8%,

82.6% by AlexNet and 90.3% by GoogleNet.

# Introduction

Human Activity Recognition has been extensively investigated using different techniques including sensing technologies, computer vision, and more recently using deep learning. It consists of the classification of the activity of a person from the data collected from sensors (e.g., accelerometer, camera, and laser scanner). The advances in human activity recognition have enabled different applications in different areas such as healthcare, sports activities, violence detection, older people monitoring, postures' recognition, to name a few. In this paper, we consider a particular human activity application with special interest to the Muslim community around the world, namely the recognition of postures of the Islamic prayer, also known as Salat. Salat is the second pillar in Islam and is the most important worshipping activity that is repeated five times a day by any Muslim. Besides its spiritual value, it consists of a series of postures that must be executed in a predefined sequence. The motivation of this work is that several people (e.g., kids, beginners) are likely not to perform the postures

correctly. There are different reasons for this to happen like, for example, the lack of proper learning of prayer movements, or being careless. Besides, another requirement in Salat is the tranquillity; that is, every posture should be executed in a proper and sufficient amount of time while making invocations and reading the Quran.

In this paper, we propose a novel solution to the posture recognition of Salat using convolutional neural networks (CNN), AlexNet and GoogleNet which aims at identifying the basic postures of Salat namely, standing (Qiam), Bowing (Ruku), Prostration (Sujud), using state-of-the-art CNN algorithms.

## Methodology

In order to detect the prayer posture and classify it, we have used CNN, AlexNet and GoogleNet. The train image dataset is transmitted into the neural network to detect and classify the posture. The images are resized to 224x224x3 resolution. The dataset is then shuffled and then contains 457 train images and 196 test images.

Convolutional Neural Network (CNN) consist of two convolutional layers with 128 and 64 filter each of consist 3x3 tensors. The relu activation function is applied. Each convolutional layer is passed to the

MaxPooling layer consist of 2x2 tensors. The result is then flattened and then passed to the Dense Layer with 512, 10 and 3 neurons. The two dense layer is carried out by the relu activation function and the last dense layer contains softmax as an activation function. The applied learning rate is 0.0001 and the loss is calculated with categorical crossentropy.

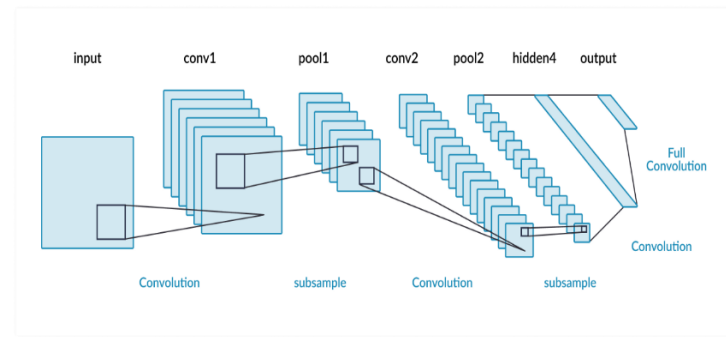


Figure 1 - Convolutional Neural Network

AlexNet consist of five convolutional layers with a 96, two 256 and three 384 filters carried out with relu activation function. After, each convolutional layer, Batch Normalization is applied. Then pooling layers are used to perform max pooling. Three Dense Layers are applied with relu activation function after the five convolutional layer, batch normalization and max pooling. After each dense layer, a dropout of 0.5 is applied and the last dense layer with 3 neurons is then passed to the softmax activation

function. The applied learning rate is 0.0001 and the loss is calculated with categorical crossentropy.

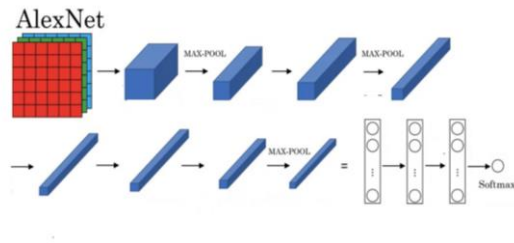


Figure 2 - ALEXNET Neural Network

GoogleNet is a 22-layer deep convolutional neural network that's a variant of the Inception Network. GoogleNet architecture consists of 22 layers (27 layers including pooling layers), and part of these layers are a total of 9 inception modules. A dropout of 0.7 and 0.4 are applied after two dense layers and global average pooling. The applied learning rate is 0.0001 and the loss is calculated with categorical crossentropy.

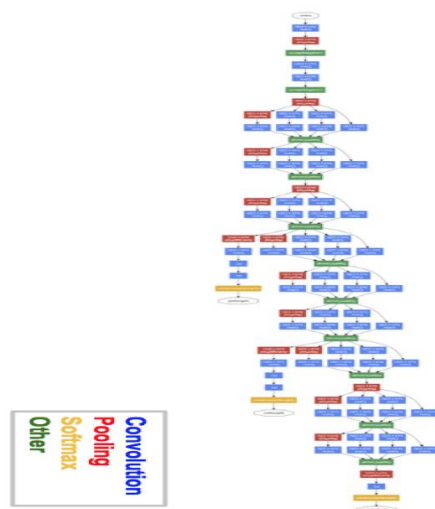


Figure 3 - GoogleNet Neural Network

## Dataset:

Prayer Postures Dataset consist from a set of images of students and laboratory members performing different Salat positions using mobile phones, from various view angles, especially captured for the purpose, complemented with images (and video frames) of people in prayer collected from the Internet.

The total number of collected images is 653 that we manually labelled into three classes which include Ruku, Sujud and Qiam. The total images are then shuffled and 70-30 ratio is performed through which training dataset of 457 images and 196 test images of different postures are split.



Figure 4 - Prayer Dataset

## Result:

The Experiment shows that the accuracy obtained by 457 trained images and giving 196 test images to CNN is 90.8%, AlexNet 82.6% and GoogleNet 90.3% with 50 iterations.

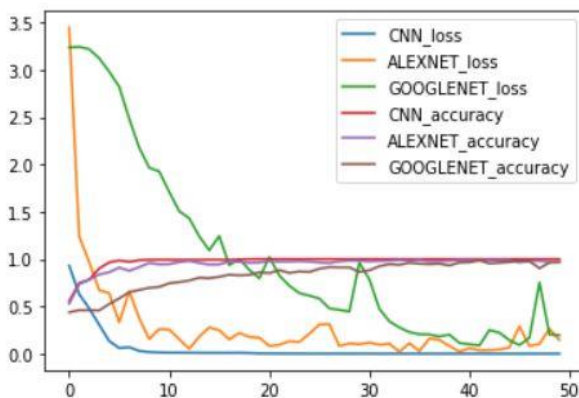


Figure 5 - Graph Showing Loss and Accuracy Analysis

## Conclusion:

We evaluate the study of classifying and detecting Islamic prayer postures. The classification was performed by training the Convolutional Neural Network (CNN) model, AlexNet and GoogleNet Models. The loss and accuracy from these three models are then analyzed by modeling a graph that shows the decrease of loss and increase of accuracy from these models. The training dataset consist of 457 images and test dataset contains 196 images through which we evaluate our models. Each image was resized to 224x224. The prayer postures are then classified into

Ruku, Sujud and Qiam. We achieve 90.8% accuracy from CNN, 82.6% accuracy from AlexNet and 90.3% accuracy from GoogleNet.

This study is a first step that will be further developed in future works, by enlarging the training dataset, investigating other network architectures, conducting thorough hyperparameter optimization, and integrating the object detection model into a broader artificial intelligence assistive framework that aims to guide Muslim worshippers to evaluate the correctness of their postures during prayer.

## Reference:

- [1] Koubaa, Anis & Ammar, Adel & Benjdira, Bilel & Al-Hadid, Abdullatif & Kawaf, Belal & Al-Yahri, Saleh Ali & Babiker, Abdelrahman & Assaf, Koutaiba & Ba Ras, Mohannad. (2020). Activity Monitoring of Islamic Prayer (Salat) Postures using Deep Learning. 106-111. 10.1109/CDMA47397.2020.00024.