ABSTRACT

Stochastic Gradient Decent (SGD) is one among the core techniques behind the success of deep neural networks. The gradient provides information on the direction within which a function has the steepest rate of change. the most problem with basic SGD is to vary by equal sized steps for all parameters, no matter gradient behavior. within the proposed diffGrad optimization technique, the step size is adjusted for every parameter in such how that it should have a bigger step size for faster gradient changing parameters and a lower step size for lower gradient changing parameters.Various architectures of CNN are available. the most advantage of CNN is its image classification capability. The three different popular CNN models are implemented for image classification on easily available dataset.The small begginer model of CNN was also implemented for mnist dataset so more results are available for proving the optimization capability of praposed method.The three models were trained using various keras inbuilt optimizers like Adam and SGD. Finally for enhancing the validation accuracy of model the models were trained using DiffGrad optimizer. The results showed that the DiffGrad optimizer outperforms all other existing optimizers. As accuracy using praposed method increased, the DiffGrad was proven to be efficient optimization technique for Convolutional Neural Network.

|  |  |
| --- | --- |
| List of Acronyms & Symbols | |
| ML | Machine Learning |
| CNN | Convolutional Neural Network |
| SGD | Stochastic Gradient Descent |
| Adam | Adaptive Moment Estimation |
| MNIST | Modified National Institute of Standards and Technology |
| MLP | Multilayer Perceptron |
| DFC | DiffGrad Friction Coefficient |
| FFT | Fast Fourier Transform |
| GPU | Graphics Processing Unit |
| FPGA | Field Programmable Gate Array |

|  |  |  |
| --- | --- | --- |
| List of Figures | | |
| Fig. 2.1 | AlexNet Architecture | 7 |
| Fig. 2.2 | VGG16 Architecture |  |
| Fig.2.3 | GoogLeNet Architecture |  |
| Fig.2.4 | A Flowchart of CNN |  |
| Fig.3.1 | ReLu Function |  |
| Fig.3.2 | Optimization Landscape |  |
| Fig.3.3 | Local minima scenario |  |
| Fig.3.4 | Saddle point scenario |  |

|  |  |  |
| --- | --- | --- |
| List of Tables | | |
| 3.1 | Comparison of CNN models | 4 |
|  |  |  |
|  |  |  |
|  |  |  |