(12) PATENT APPLICATION PUBLICATION
(19) INDIA

(22) Date of filing of Application :09/12/2023

(21) Application No.202311084046 A

(43) Publication Date : 12/01/2024

(54) Title of the invention : METHOD FOR IMPROVING THE ACCURACY AND EFFICIENCY OF DEEP LEARNING MODELS

(51) International classification	:G06N0003040000, G06N0003080000, G06N0020000000, G06K0009620000, G10L0015080000
(86) International Application No Filing Date	:NA :NA
(87) International Publication No	: NA
(61) Patent of Addition to Application Number Filing Date	:NA :NA
(62) Divisional to Application Number Filing Date	:NA :NA

(71)**Name of Applicant :**
   1)**Jugnesh Kumar**
      Address of Applicant :Echelon Institute of Technology, Kabulpur, Kheri-Manjhawali Road, Neharpar, Faridabad – 121101, Haryana, India Faridabad --------- -----------
   2)**Manish Kumar**
   3)**Stuti Saxena**
   4)**Mushtaq Ahmad Rather**
   5)**Preetishree Patnaik**
   6)**Sunita**
Name of Applicant : NA
Address of Applicant : NA
(72)**Name of Inventor :**
   1)**Jugnesh Kumar**
Address of Applicant :Echelon Institute of Technology, Kabulpur, Kheri-Manjhawali Road, Neharpar, Faridabad – 121101, Haryana, India Faridabad --------- -----------
   2)**Manish Kumar**
Address of Applicant :Echelon Institute of Technology, Kabulpur, Kheri-Manjhawali Road, Neharpar, Faridabad – 121101, Haryana, India Faridabad --------- -----------
   3)**Stuti Saxena**
Address of Applicant :Echelon Institute of Technology, Kabulpur, Kheri-Manjhawali Road, Neharpar, Faridabad – 121101, Haryana, India Faridabad --------- -----------
   4)**Mushtaq Ahmad Rather**
Address of Applicant :Echelon Institute of Technology, Kabulpur, Kheri-Manjhawali Road, Neharpar, Faridabad – 121101, Haryana, India Faridabad --------- -----------
   5)**Preetishree Patnaik**
Address of Applicant :St Andrews Institute of Technology & Management, Haily-Mandi Road, Khera Khurrampur, Farrukhnagar 122506, Gurugram, Haryana, India. Gurugram ----------- -----------
   6)**Sunita**
Address of Applicant :Echelon Institute of Technology, Kabulpur, Kheri-Manjhawali Road, Neharpar, Faridabad – 121101, Haryana, India Faridabad --------- -----------

(57) Abstract :
The present invention relates to the field of artificial intelligence and machine learning, specifically focusing on deep learning models. More specifically, it addresses methods to enhance the accuracy and efficiency of these models. The method for improving the accuracy and efficiency of deep learning models, includes optimizing model architecture by identifying and leveraging critical components to reduce unnecessary complexity while preserving or enhancing accuracy, advancing training algorithms to improve convergence speed and accuracy during the training phase, introducing network pruning and quantization techniques for model deployment in resource-constrained environments, thereby reducing model size and computational requirements without compromising accuracy, implementing dynamic inference strategies that adaptively adjust computational resources based on input complexity during model inference to optimize resource utilization without sacrificing accuracy, and incorporating novel regularization methods to enhance model generalization and prevent overfitting, thereby improving the robustness and reliability of deep learning models.

No. of Pages : 19 No. of Claims : 10