INIT TASK 2 RESEARCH EVALUATION

Paper Title: *ImageNet Classification with Deep Convolutional Neural Networks*

Authors: Alex Krizhevsky, Ilya Sutskever & Geoffrey E. Hinton

Source: Communications of the ACM (2017); originally presented at NIPS 2012

1. Paper Review

This paper introduced *AlexNet*, a deep learning model that completely changed the direction of computer vision. The authors trained a large convolutional neural network on the ImageNet dataset, which has over a million high-resolution images across 1000 classes. The network had eight layers (five convolutional and three fully connected) and used around 60 million parameters. It achieved a top-5 error rate of 17%, a huge improvement compared to older models. The paper's key contributions included using ReLU activation functions for faster learning, dropout to reduce overfitting, and data augmentation to improve accuracy. They also trained the model on two GPUs in parallel, which was a big technical step at the time. Overall, AlexNet showed that deep neural networks could learn directly from data and perform better than hand-engineered feature methods used before it.

2. Personal Reflection

Reading this paper made me realize how much of today's AI progress started with this one experiment. The authors not only built a powerful model but also proved that data and computation could replace manual feature design. I liked how they focused on simple but effective ideas using ReLU, data augmentation, and dropout to overcome real-world challenges like slow training and overfitting.

One limitation I noticed is that the model required huge resources: training for almost a week on multiple GPUs and using a massive labeled dataset. That made it difficult for smaller labs to replicate at the time. If I had to improve it today, I'd try combining their ideas with self-supervised learning or lightweight CNN architectures like ResNet or MobileNet to make it faster and less data-dependent.

Overall, I think this paper deserves its legendary status. It bridged research and real-world application, and most of the models we use now whether for image recognition, face detection, or autonomous driving trace their roots back to AlexNet.