3911037 3

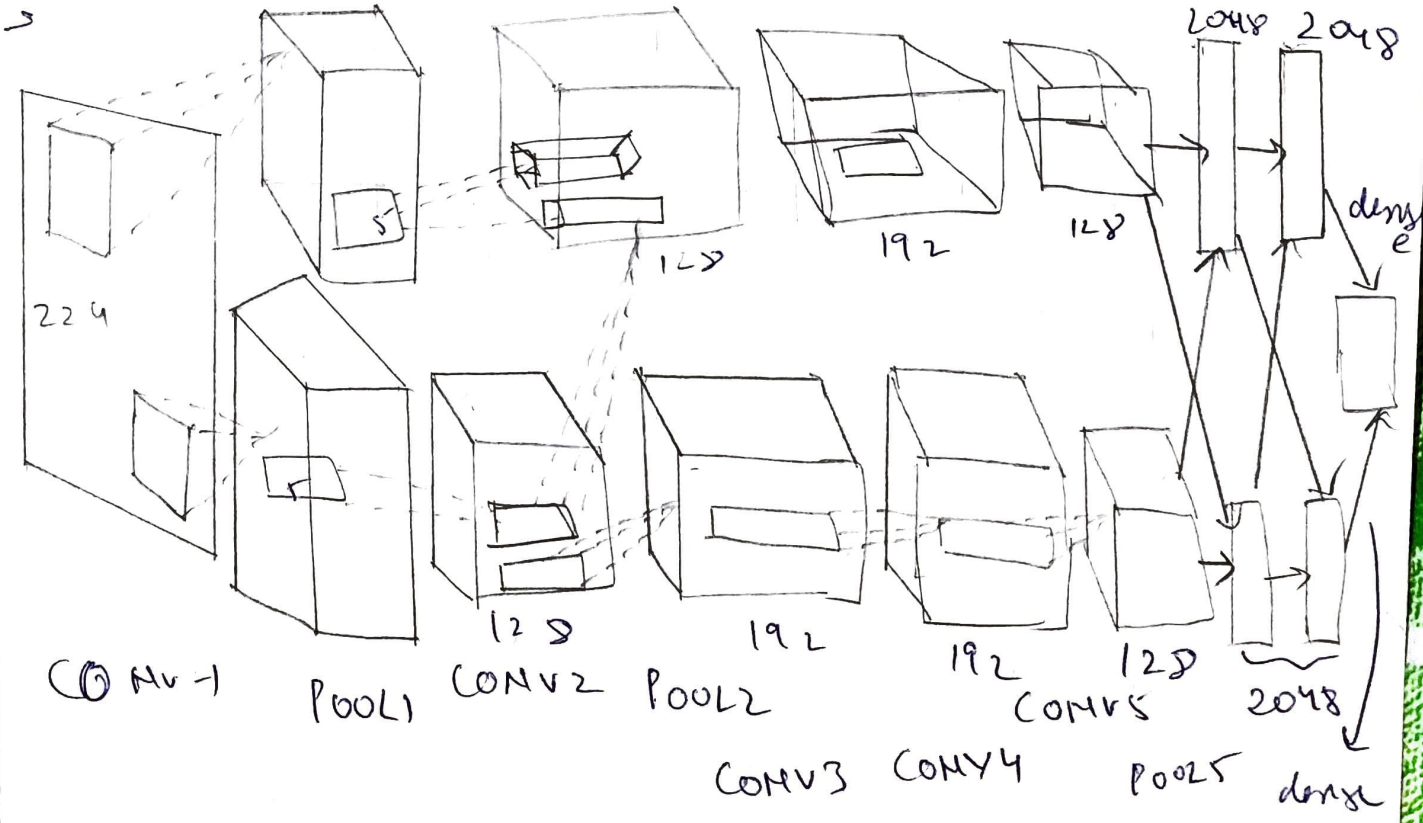
HARIHARAN BP

CASE-2 - DLNN

PART-B

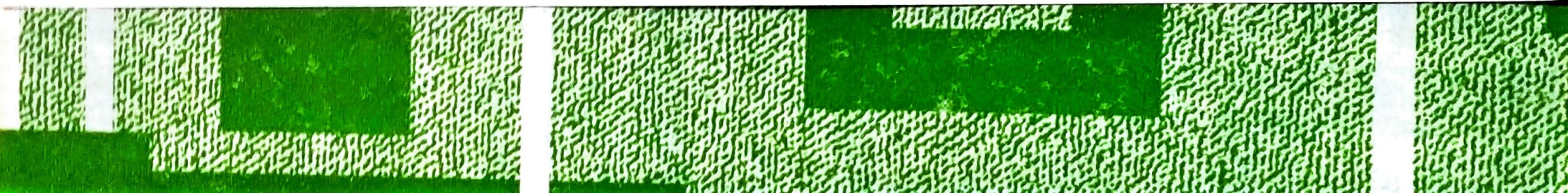
(7) → A GPU based CNN A model AlexNet can be used for the detection and identification of cancer cells from the given CT-images.

→ It was proposed in 2012 in the research paper named Imagenet Classification with deep Convolution Neural Network.



→ It has 8 eight layers with learnable parameters

→ The model has 5 layers with a combination of max pooling followed by 3 fully connected layers.



→ The fully connected layers use Relu activation except o/p layer.

→ They found out that using the Relu as an activation funct accelerated the speed of the training process.

→ Also use dropout layers which prevents the model from over fitting.

→ The model is trained on huge image dataset ~~and~~ ~~different huge~~ across huge classes.

→ The 1st conv. layer has ~~1000 filters~~ 96 filters.

→ Activation funct used is Relu.

→ Next the filter size is reduced to 5 X 5

→ * Then we have a max pooling layer of size 3X3 with stride 2

→ 3rd conv. operation ~~to~~ with 384 filters of size 3X3 stride

→ 4th conv. with 384 filters and 5th ~~with~~ and final fc layer with 256 filters.

(9) → The objective function is a non convex func^n.

→ All non-linear problems can be modeled by using convex func^n

→ It has multiple feasible regions and multiple locally optimal points.

→ There can't be a general algorithm to solve it efficiently in all cases.

→ Neural networks are universal func^n approximators. to do this they need to be able to approximate non-convex func^n.

→ The non-convex problems) optimal optimization can be solved using:-

(i) Stochastic Gradient descent

(ii) Momentum

(iii) Mini-batching

(iv) SVRG.

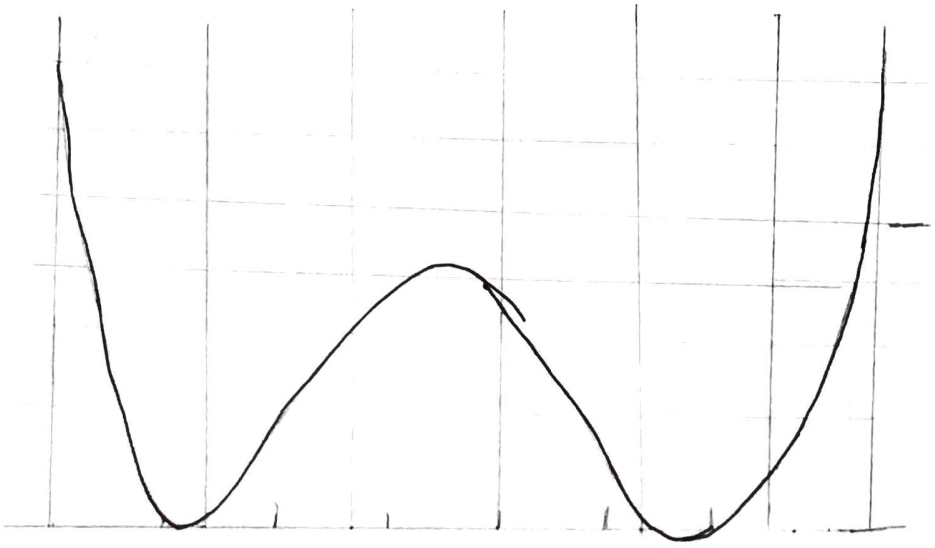
→ The following are the need for Non-Convexity:-

(i) Very Flat Regions

(ii) Presence of Saddle points.

(iii) Many Local minima are present

(iv) Has Varying Curvature



Non- Convex Region

PART-A

(1) Uses:-

(i) Skip the networks

(ii) Reduces training time

(iii) Reduces Complexity.

(12)

(2) They control ~~the~~ the training of any given ML / DL ~~produce~~ models thereby optimizing the these parameters we can reduce the training time and then we try to increase the performance

~~performance~~

(3) ~~Rely Rela~~

(5) For correctness let us assume, a case where

A = 1, B = 1. O/P = 2A + 3B.

So is, if A increases the O/P increases too.

On other side, if B = -100 then the O/P is 0.

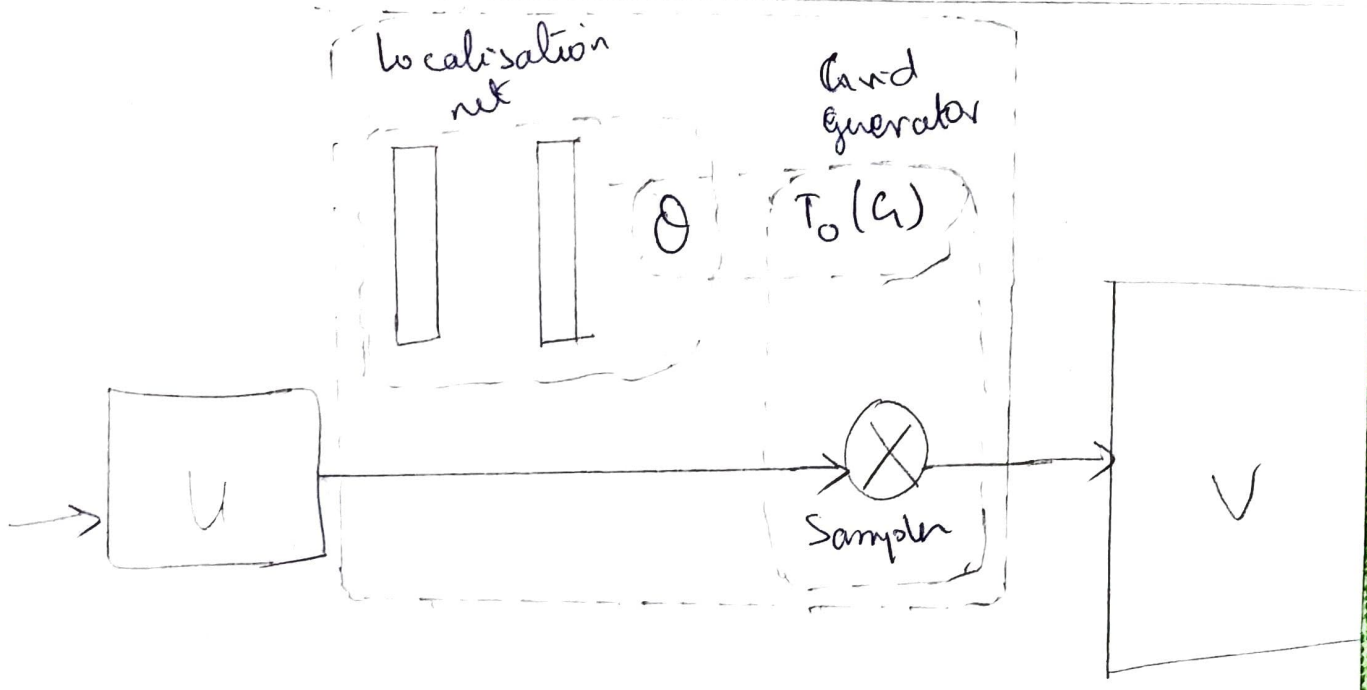
So, if A increases slowly/moderately then O/P as stays 0. So, we can conclude that 'A' might alter our O/P or it might not. so, it ~~whole~~ d wholly depends on the value of B.

(4) ResNet (Residual Network) can be used to solve this issue. It solves the issue by skiping the ~~It consists~~ nets vanished networks.

ResNet is a deep network with over ~~100~~ 100 layers. It is based on the idea of skips connection which came from highway network. ~~wh~~ ~~were~~ where gated shortcuts connection are was used.

(2)

(5)



→ Helps to crop out and scale normalizes the appropriate region.

→ which can simplify the subsequent classification task

→ Consists of 3 parts. Localization, Grid, Sampler,

→ Used for Performing transformation such as Cropping, Rotation etc.