ES 665
Project Presentation

JustRaigs Challenge

(Justified Referral in AI Glaucoma Screening)

- Referral performance
- Justification performance

Team Members:

Atharva Tiwari

Kunj Jasoria

Mihir Agarwal

What is Glaucoma ?

- **Glaucoma** is a group of eye diseases that can cause vision loss and blindness by damaging a nerve in the back of eye called the optic nerve.
- This damage is often caused by **abnormally high pressure** in the eye.
- There are many different types of glaucoma, but the most common type is called **open-angle glaucoma** —that's what most people mean when they talk about glaucoma.

healthy eyes glaucoma

How big this disease is?

- It is estimated that **75 million** people worldwide are affected by primary open-angle glaucoma (POAG).
- This number is estimated to reach **111.8 million** by **2040**.
- In India alone around **12 million** people are affected by it.
- People **over 60** years of age, family members of those already diagnosed with **diabetes**, **steroid users**, as well as those with **high myopia**, **hypertension** and **eye injury** are at an increased risk of glaucoma.

Problems faced in this field

- At first, glaucoma **doesn't** usually have **any symptoms**.
- Hence, it often goes **unnoticed in its early stages** but can lead to **significant vision impairment** in advanced stages.
- There's **no cure** for glaucoma.
- All we can do is to try to **detect it early** and start treatment to stop the damage before hand.

Role of AI in Glaucoma Screening

- As glaucoma often presents no early symptoms, **regular screening** is vital for prevention of irreversible vision loss.
- Utilizing AI algorithms to analyze **color fundus photographs (CFPs)** can lead to earlier and more accurate detection of subtle glaucomatous changes than current methods.
- AI-driven glaucoma screening enables scalable, **economical**, and **precise diagnostics**, making advanced eye care accessible in resource-constrained environments.

Fundus images are photographs taken of the interior surface of the eye, including the retina, optic disc, macula, and posterior pole (the back part of the eye).

Tasks and Objectives

- **Task 1 (Binary Classification):** Differentiate between 'Referable Glaucoma' (RG) and 'No Referable Glaucoma' (NRG) based on fundus images.

- **Task 2 (Multi-label Classification):** Identify specific glaucomatous features in images marked as RG, such as disc hemorrhages and nerve fiber layer defects.

Some Important Terms

- **Referable glaucoma:** It refers to eyes where the fundus image exhibits signs or features indicative of glaucoma that require further examination or referral to a specialist.

- **No referable glaucoma:** It refers to cases where the fundus image does not show significant indications of glaucoma and does not require immediate referral.

JustRaigs Challenge

- The Justified Referral in AI Glaucoma Screening (JustRAIGS) challenge offers a large dataset of 110k+ annotated fundus photographs for developing AI algorithms.
- Dataset includes 101,442 training images and 9,741 test images, labeled as "referable glaucoma" or "no referable glaucoma."
- Additional annotations provided for referable glaucoma images, detailing up to ten glaucomatous features.

Tasks

Task 1: Referral performance

Binary classification of referable glaucoma and no referable glaucoma

- No referable glaucoma (NRG)
- Referable glaucoma (RG)

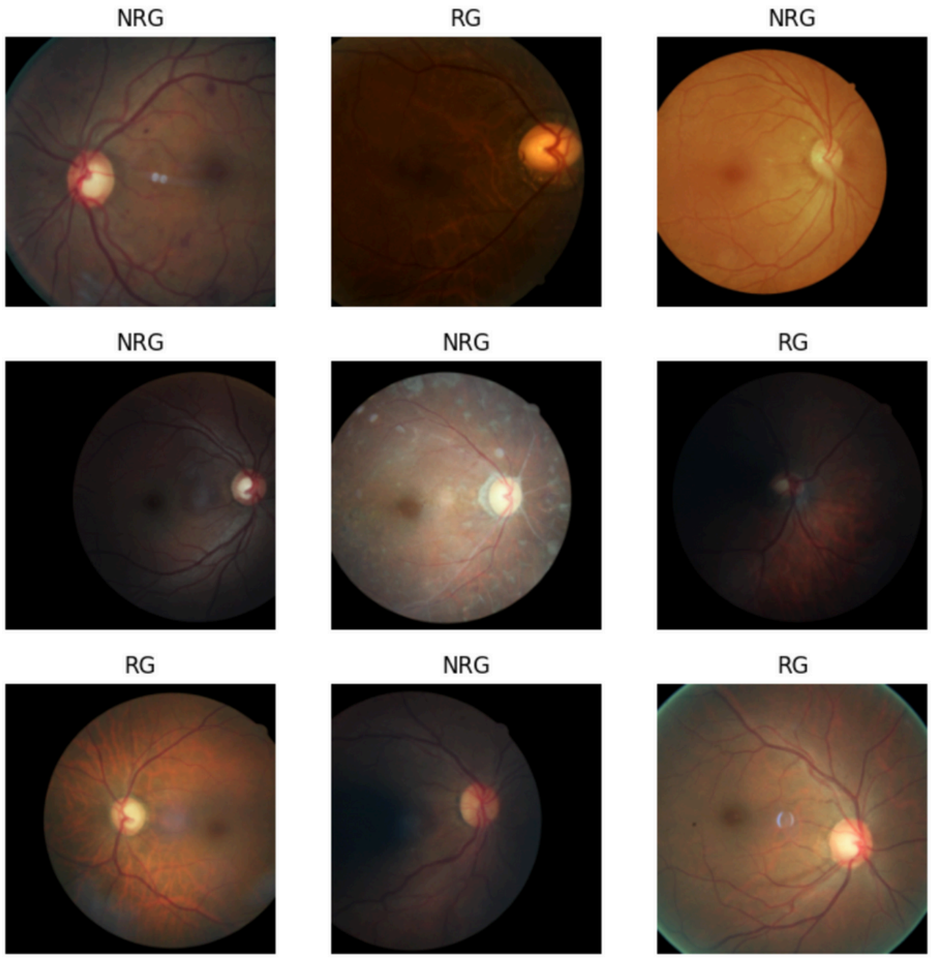
Task 2: Justification performance

Multi-label classification of ten additional features

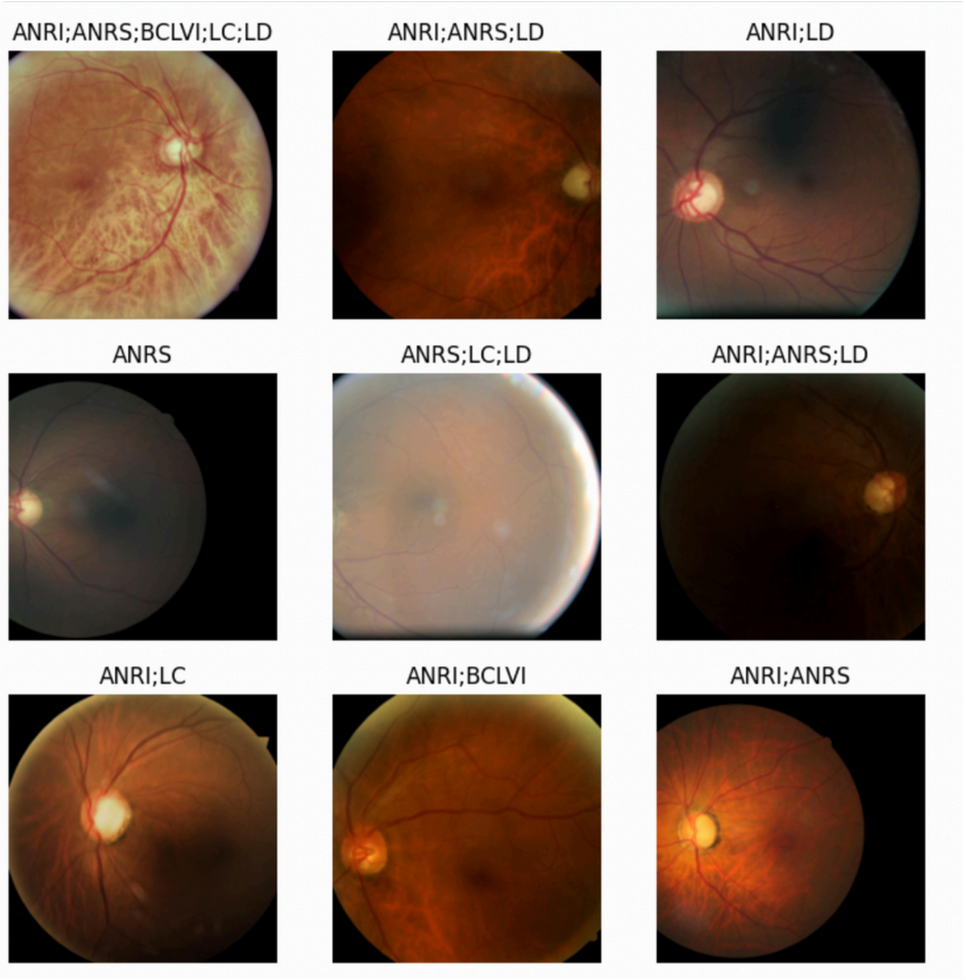
- Appearance neuroretinal rim superiorly (ANRS)
- Appearance neuroretinal rim inferiorly (ANRI)
- Retinal nerve fiber layer defect superiorly (RNFLDS)
- Retinal nerve fiber layer defect inferiorly (RNFLDI)
- Baring circumlinear vessel superiorly (BCLVS)
- Baring circumlinear vessel inferiorly (BCLVI)
- Nasalisation of vessel trunk (NVT)
- Disc hemorrhages (DH)
- Laminar dots (LD)
- Large cup (LC)

Tasks

Task 1: Referral performance



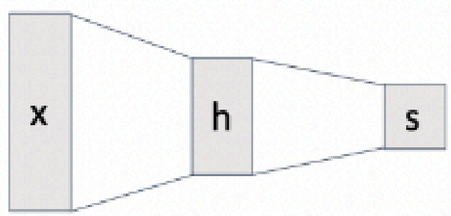
Task 2: Justification performance



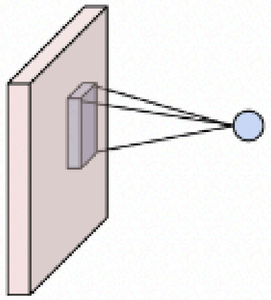
CNN – Convolutional Neural Network

Components of a Convolutional Network

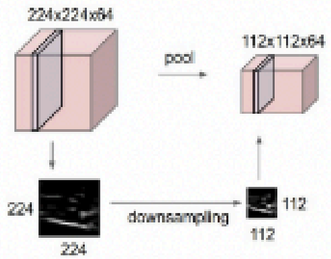
Fully-Connected Layers



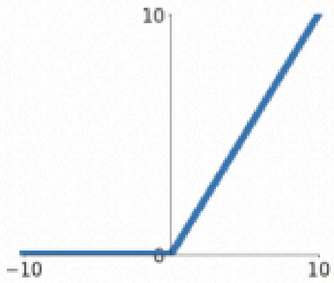
Convolution Layers



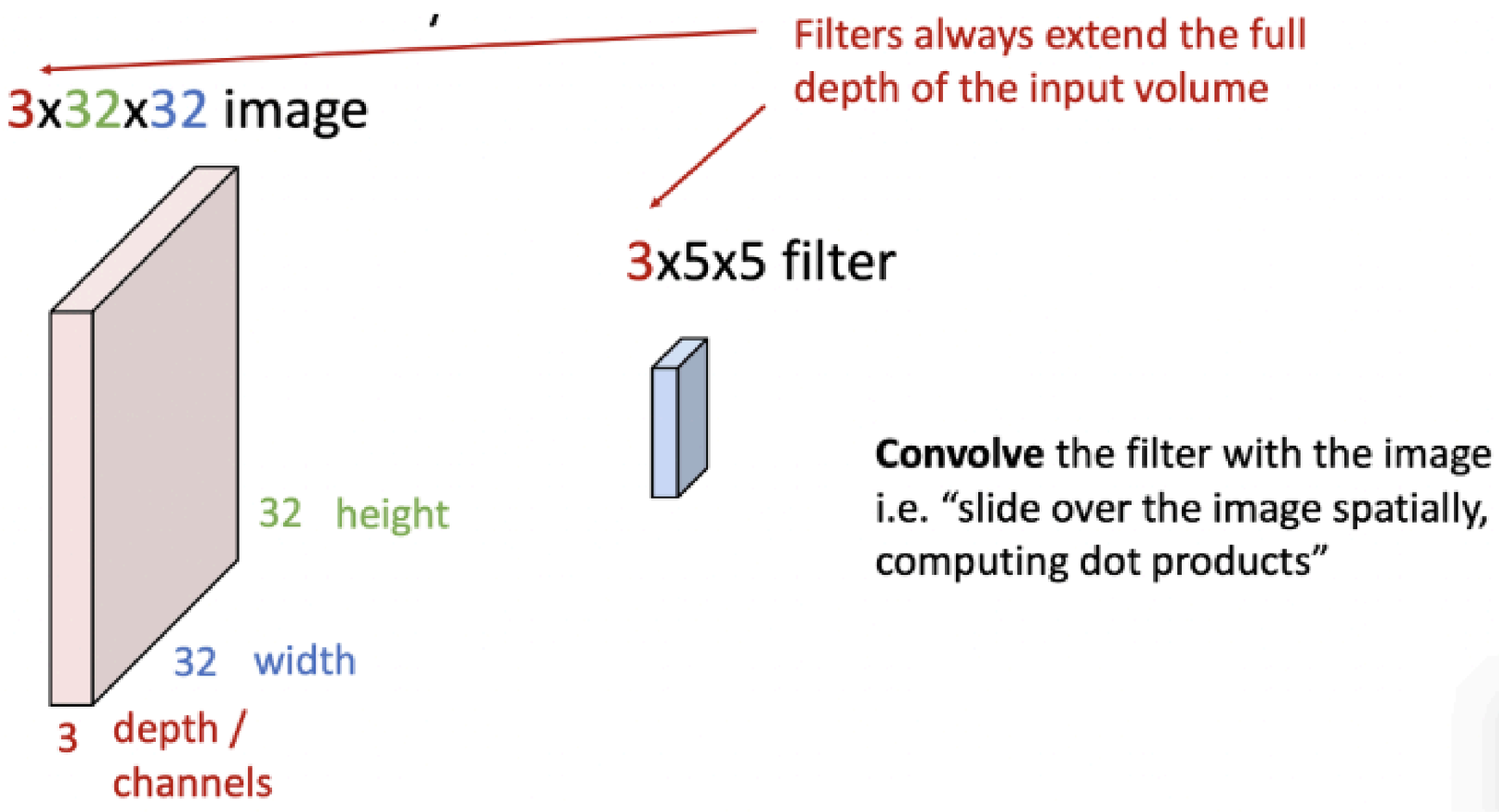
Pooling Layers



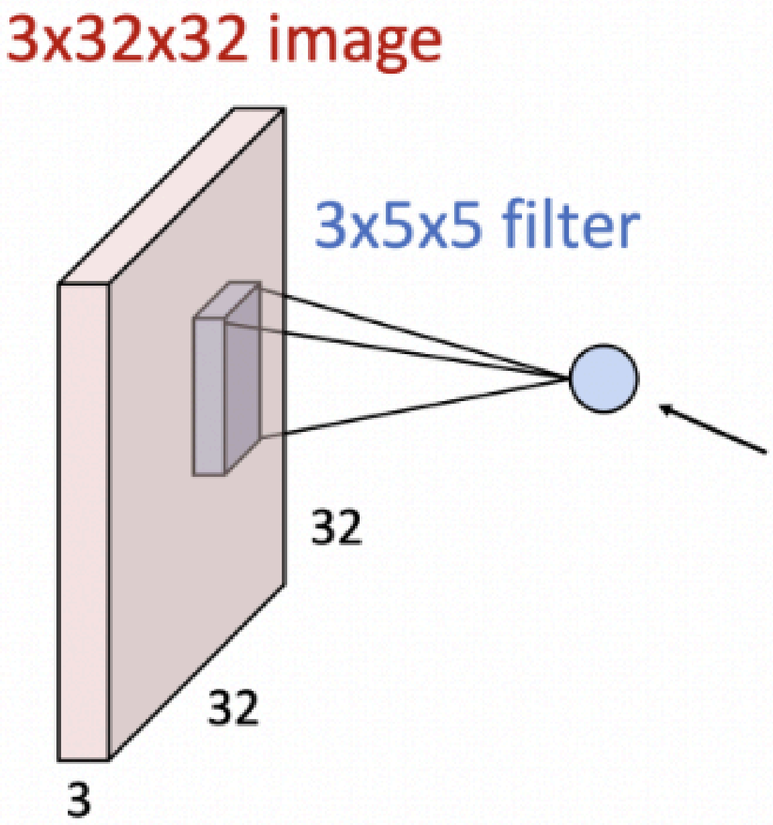
Activation Function



Convolutional Layer



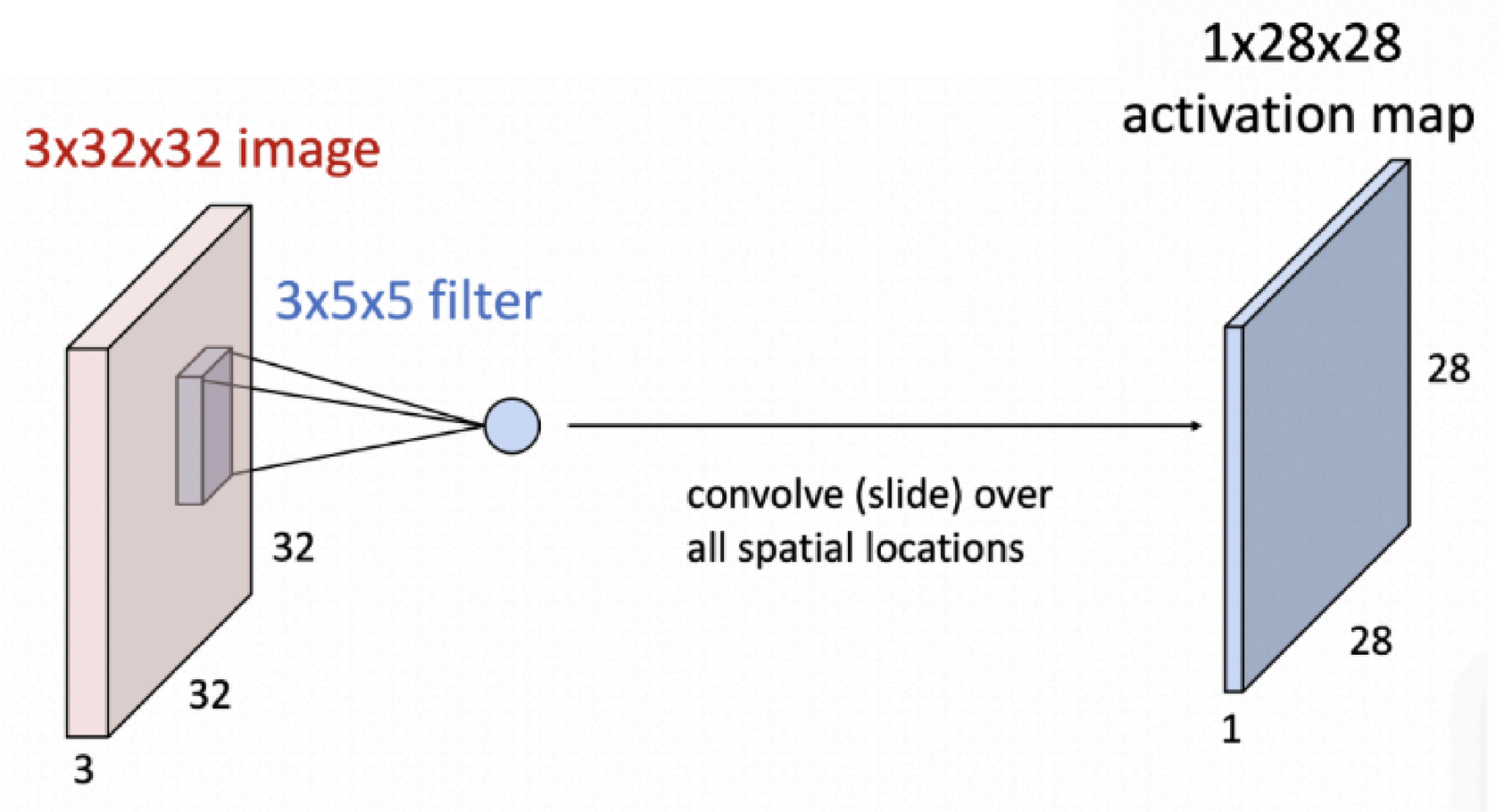
Convolutional Layer



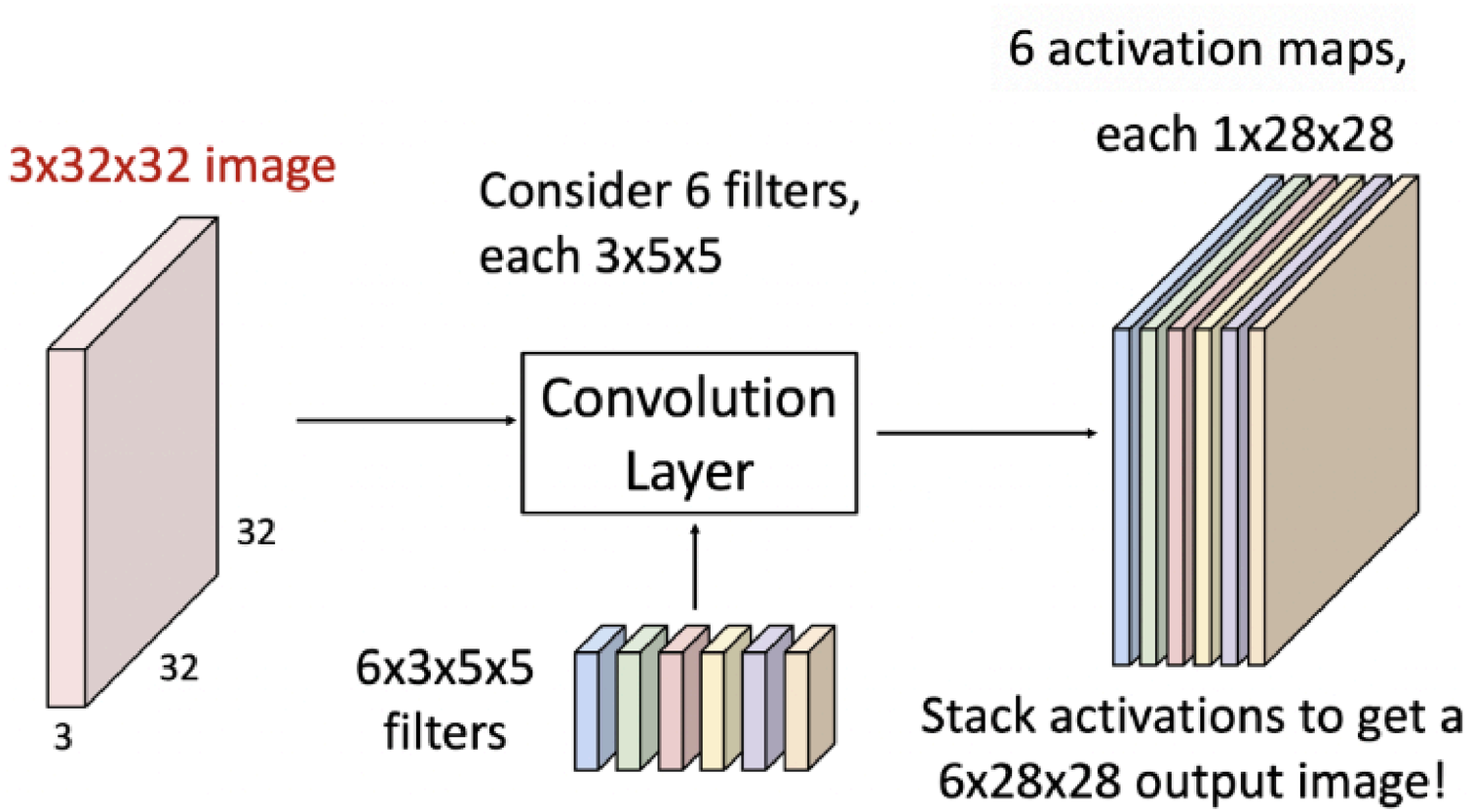
1 number:
the result of taking a dot product between the filter and a small 3x5x5 chunk of the image
(i.e. 3*5*5 = 75-dimensional dot product + bias)

$$w^T x + b$$

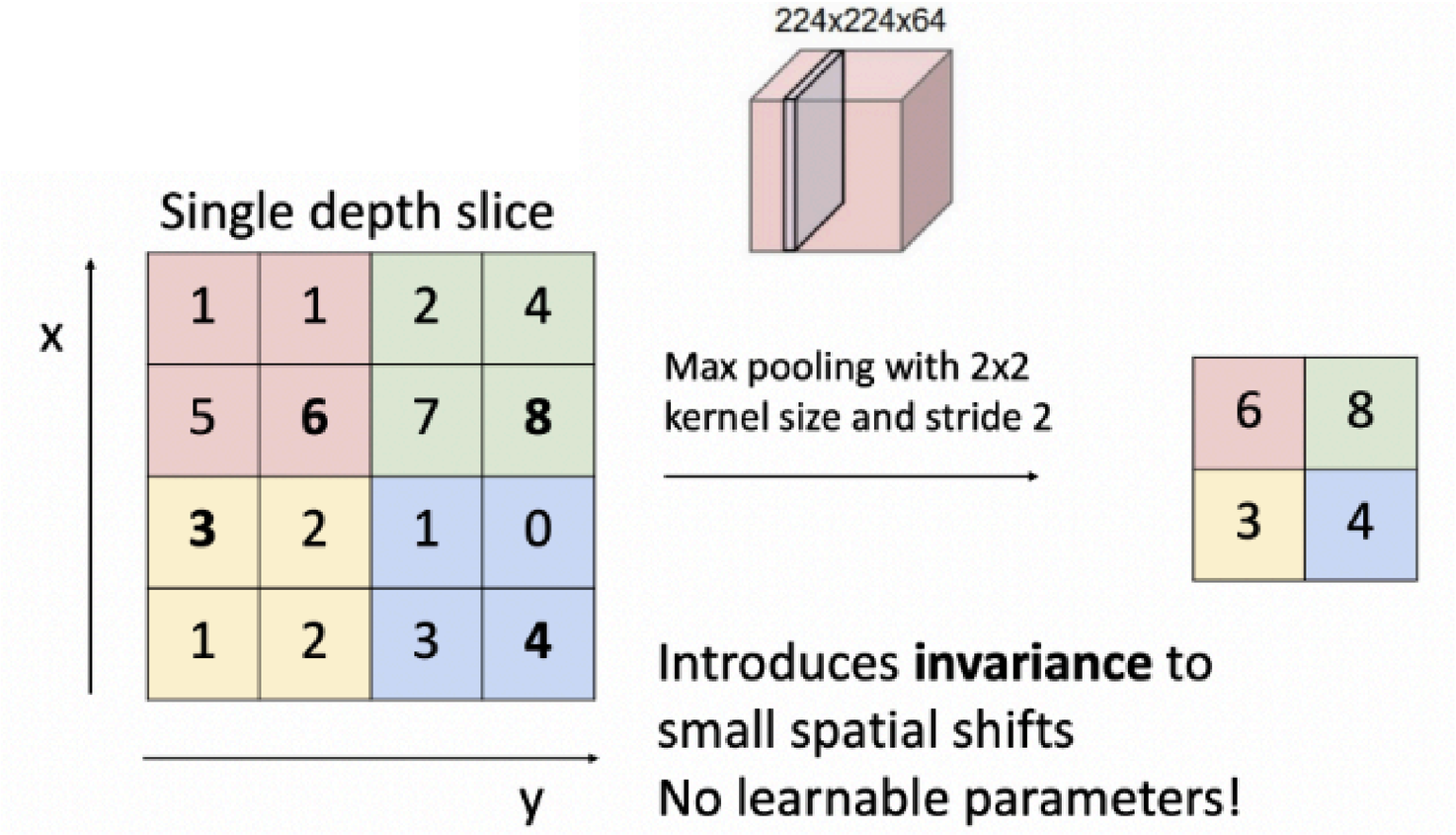
Convolutional Layer



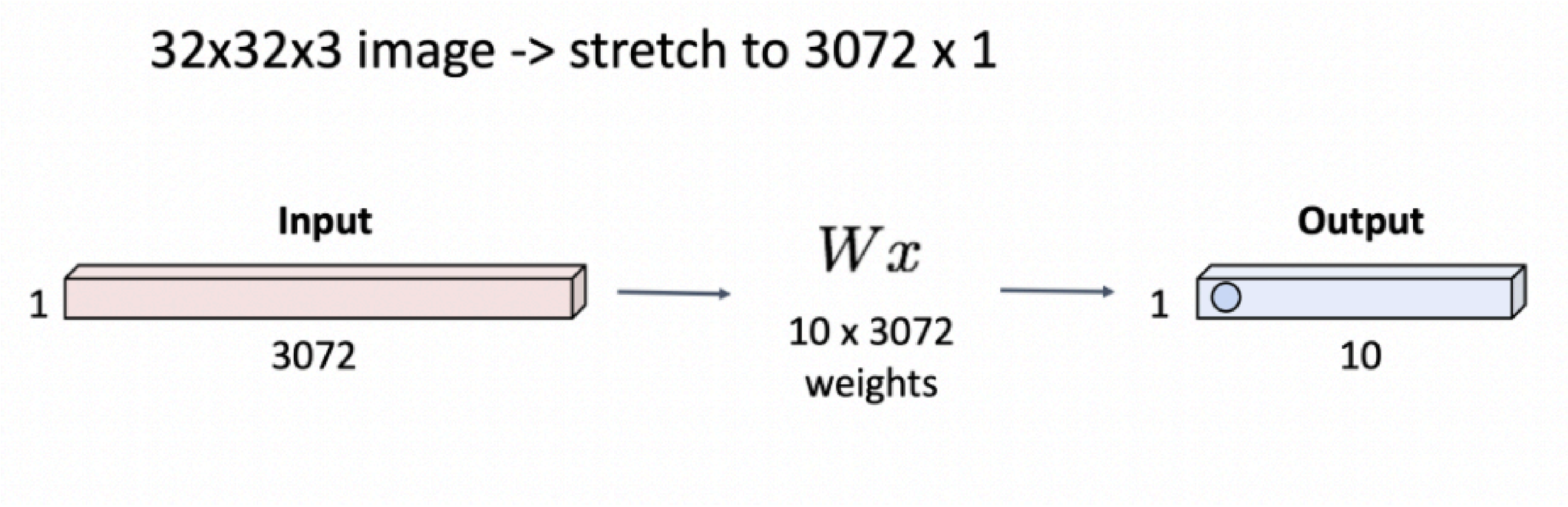
Convolutional Layer



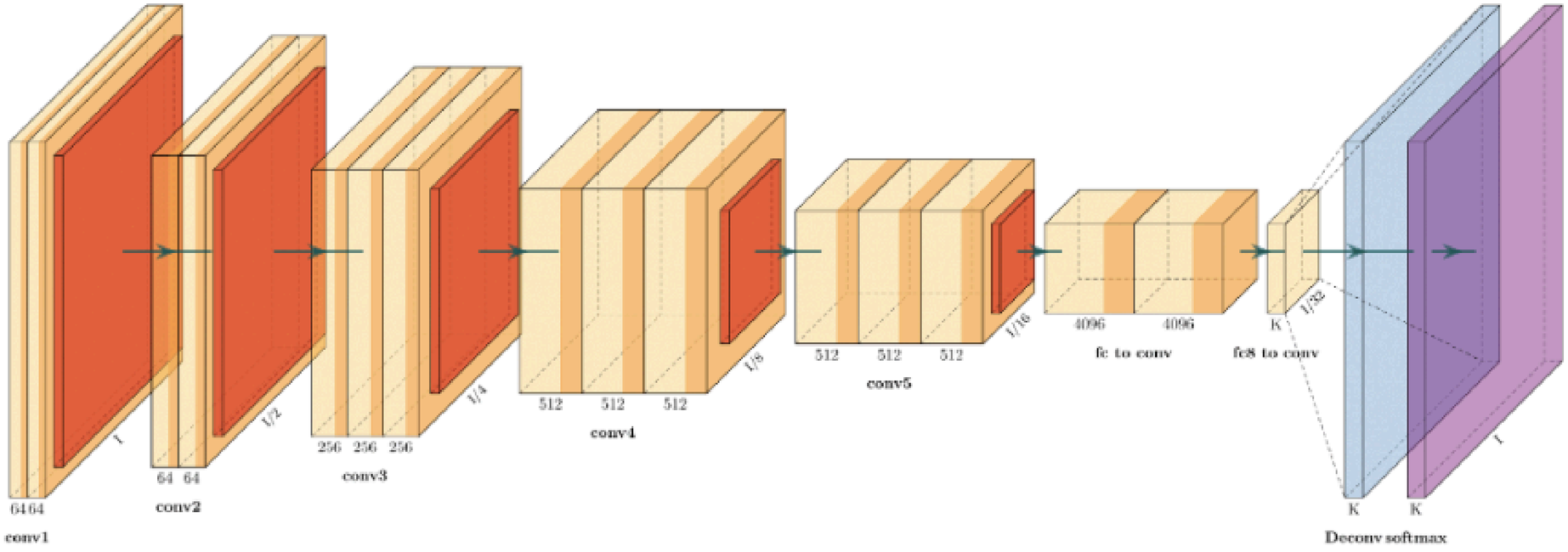
Max Pooling


224x224x64
Single depth slice
x
y
1 1 2 4
5 6 7 8
3 2 1 0
1 2 3 4
Max pooling with 2x2
kernel size and stride 2
6 8
3 4
Introduces **invariance** to
small spatial shifts
No learnable parameters!

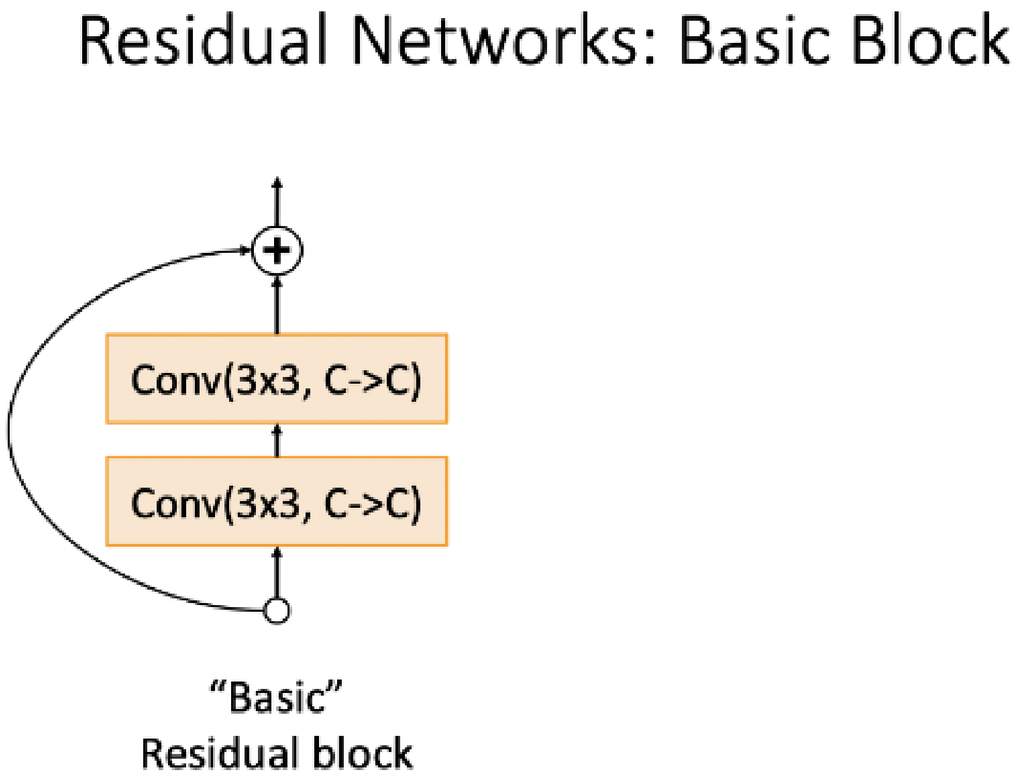
Fully Connected Layer



Resnet

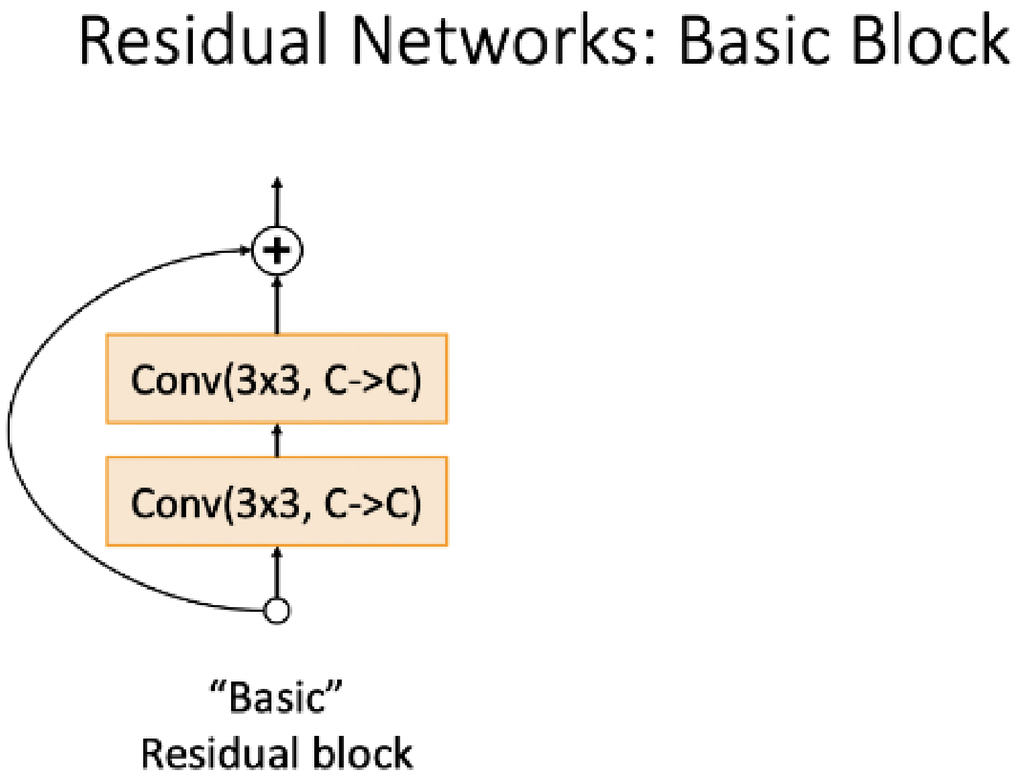

64 64
conv1
1
64 64
conv2
1/2
256 256 256
conv3
1/4
512 512 512
conv4
1/8
512 512 512
conv5
1/16
4096 4096
fc to conv
K
1/32
fc8 to conv
K K
Deconv softmax
1

Resnet



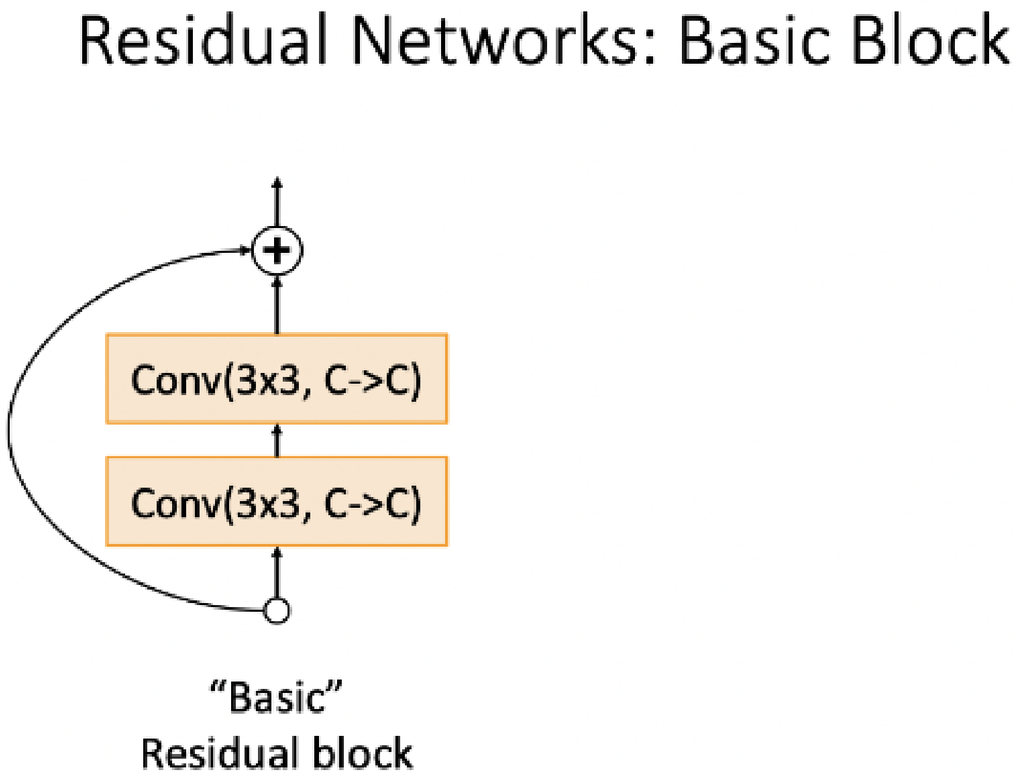
Skip connections are used in ResNet to address the problem of **vanishing gradients** during training. As the number of layers in a neural network increases, the gradients tend to become smaller and smaller as they propagate backward through the layers. This can make it difficult to update the weights of the earlier layers, and as a result, these layers may not learn meaningful features.

Resnet



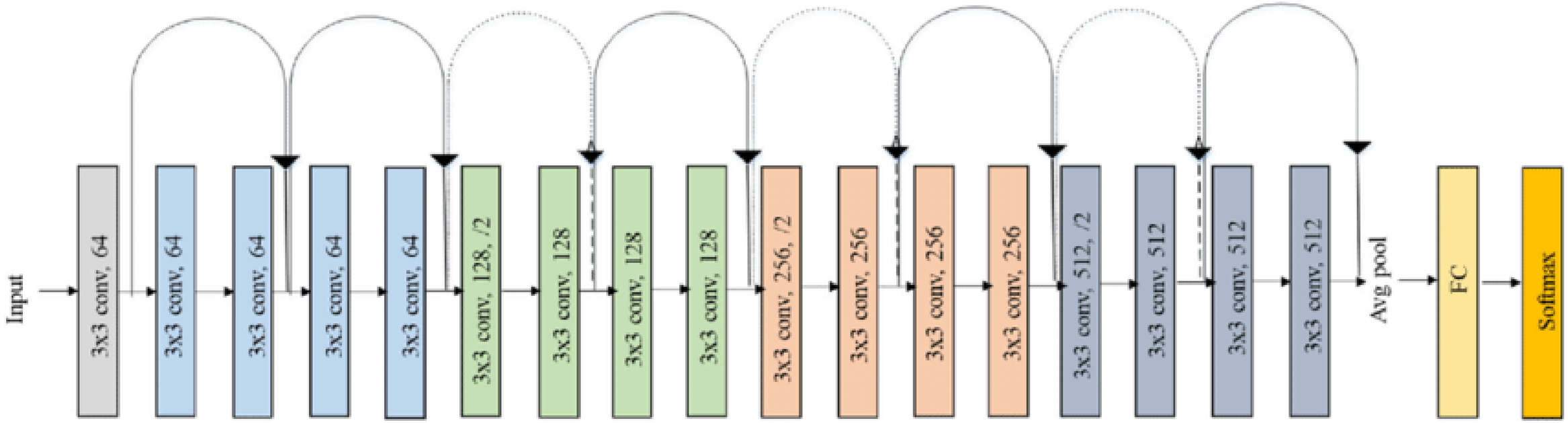
Skip connections are used in ResNet to address the problem of **vanishing gradients** during training. As the number of layers in a neural network increases, the gradients tend to become smaller and smaller as they propagate backward through the layers. This can make it difficult to update the weights of the earlier layers, and as a result, these layers may not learn meaningful features.

Resnet

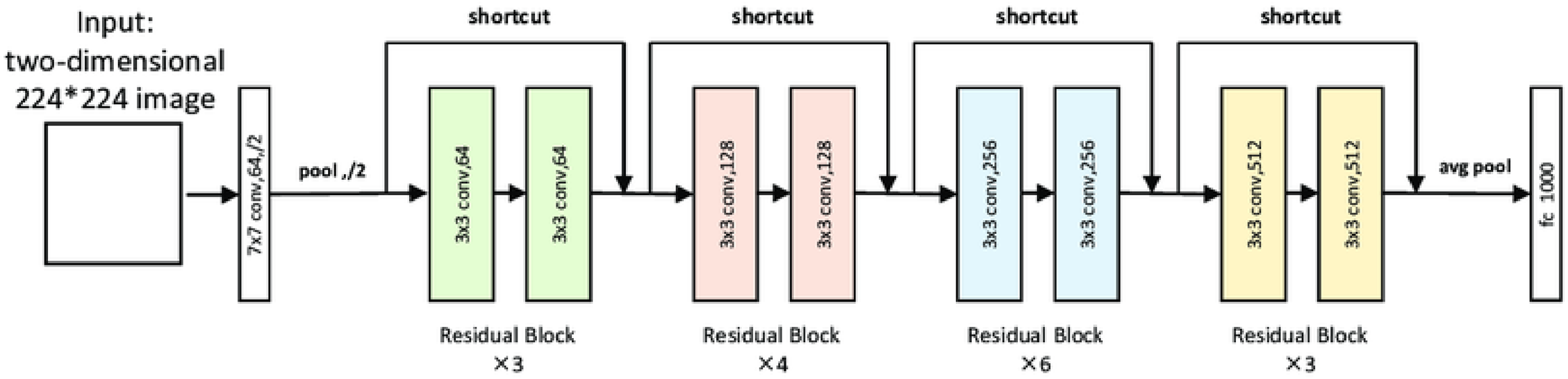


Skip connections are used in ResNet to address the problem of **vanishing gradients** during training. As the number of layers in a neural network increases, the gradients tend to become smaller and smaller as they propagate backward through the layers. This can make it difficult to update the weights of the earlier layers, and as a result, these layers may not learn meaningful features.

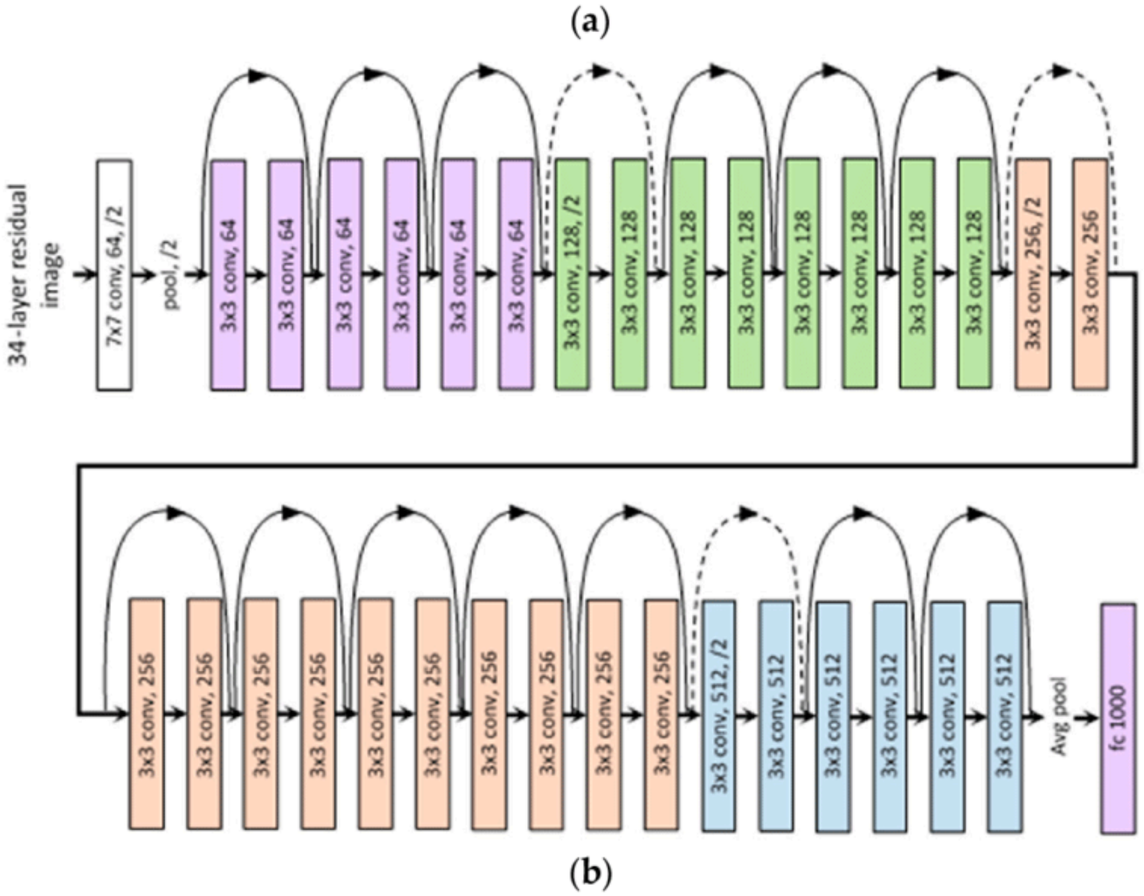
Resnet18


Input
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 256, /2
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 512, /2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
Avg pool
FC
Softmax

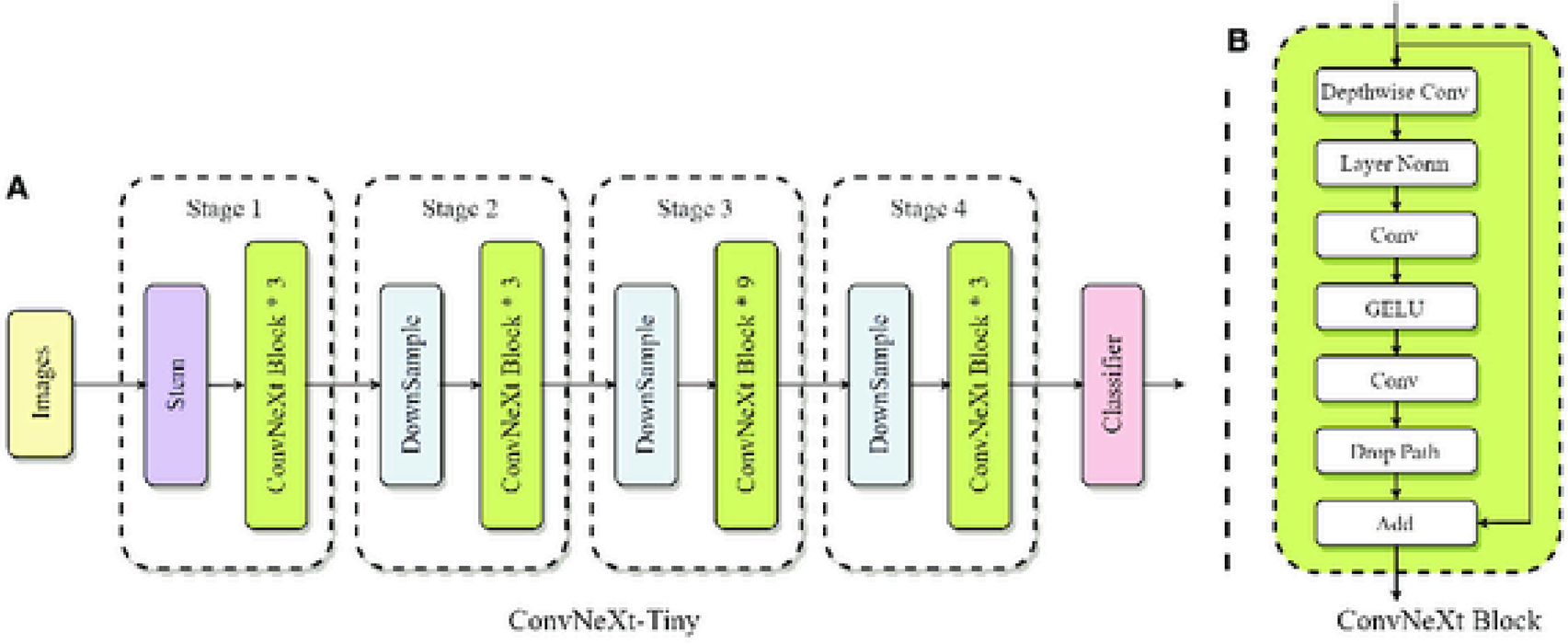
Resnet34


Input:
two-dimensional
224*224 image
7x7 conv,64,/2
pool ,/2
shortcut
3x3 conv,64
3x3 conv,64
Residual Block
×3
shortcut
3x3 conv,128
3x3 conv,128
Residual Block
×4
shortcut
3x3 conv,256
3x3 conv,256
Residual Block
×6
shortcut
3x3 conv,512
3x3 conv,512
Residual Block
×3
avg pool
fc 1000

Resnet50


(a)
34-layer residual image
7x7 conv, 64, /2
pool, /2
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 256, /2
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 512, /2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
Avg pool
fc 1000
(b)

ConvNeXt-Tiny



Ensemble Learning

- Ensemble learning involves combining predictions from multiple models to improve overall performance.
- Ensemble of ResNet18, ResNet34, and ResNet50 involves training these CNNs separately and then combining their predictions.
- Each ResNet model learns different representations and features from the data.
- Ensemble leverages the diversity of these models to make more accurate predictions.
- Combining predictions from multiple models can help mitigate individual model biases and errors.

GelU Activation

$$\mathrm{GELU}(x) = xP(X \leq x) = x\Phi(x)$$
$$\approx 0.5x\left(1 + \tanh\left[\sqrt{2/\pi}\left(x + 0.044715x^3\right)\right]\right)$$

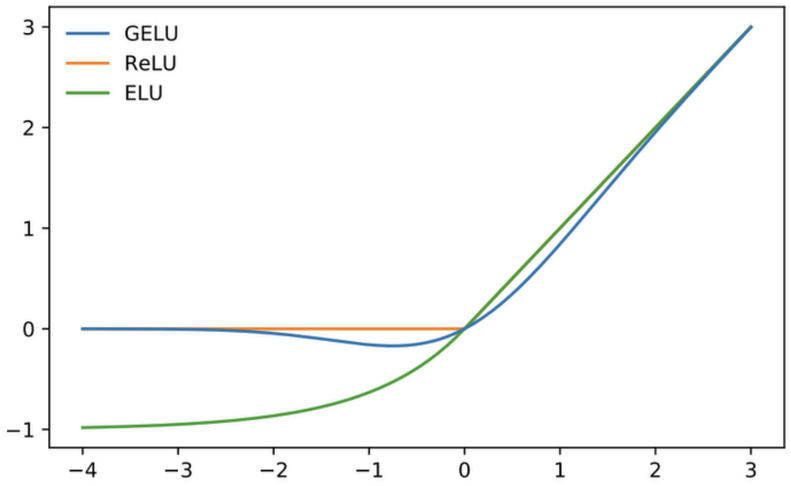
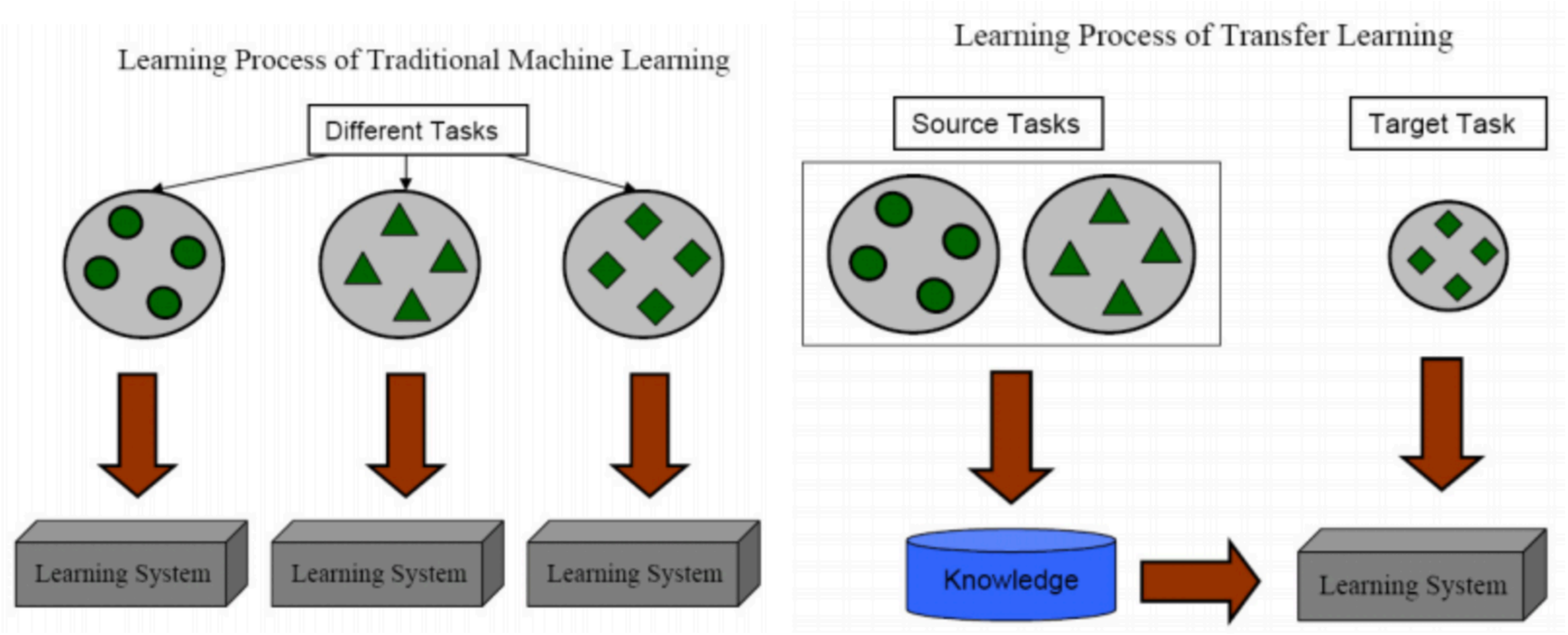


Figure 1: The GELU ($\mu = 0, \sigma = 1$), ReLU, and ELU ($\alpha = 1$).

Transfer Learning

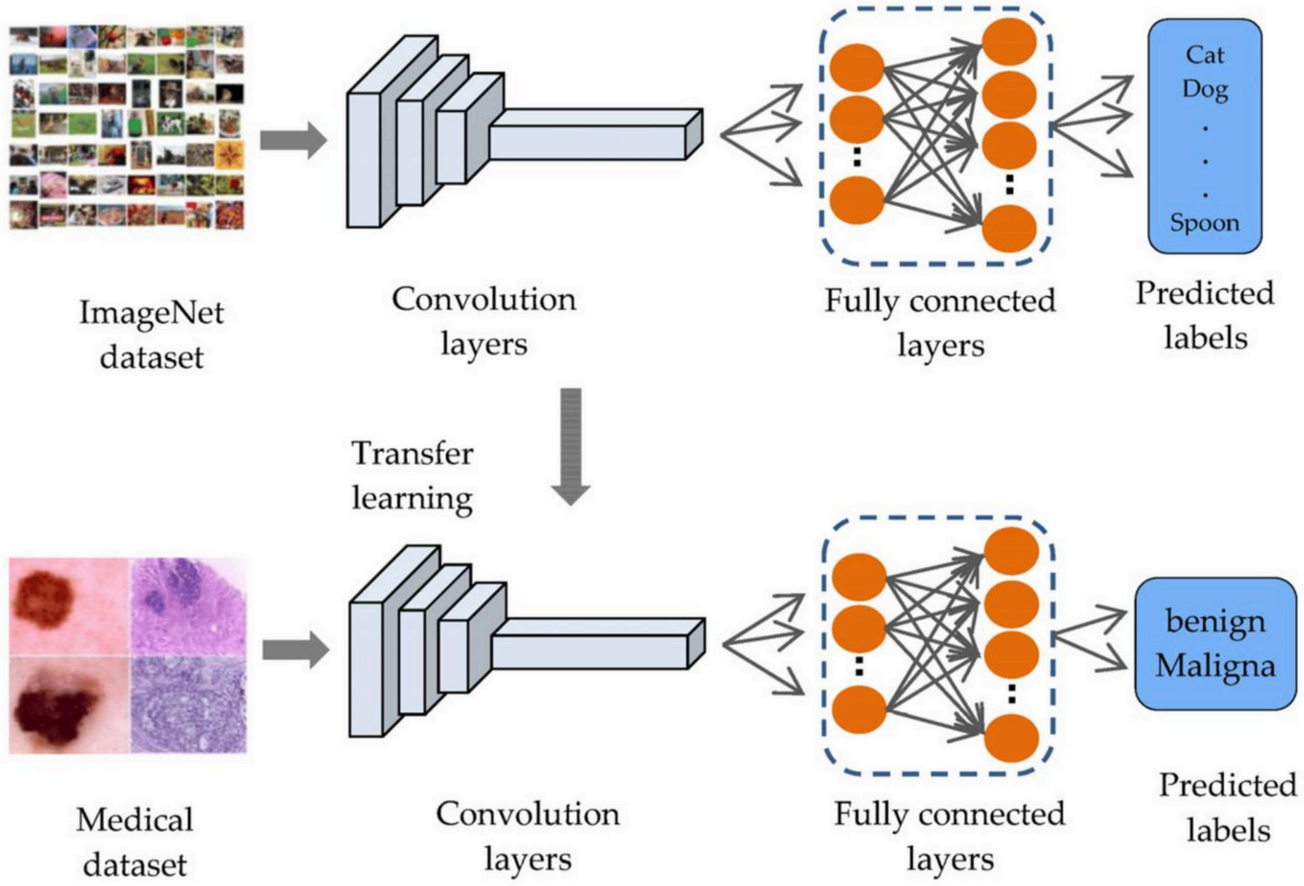
- Typically, training a model takes a **large** amount of compute resources and time.
- **Transfer learning** is a machine learning method where a model already developed for a task is **reused** in another task.
- It enables the training of deep neural networks with **less** data compared to having to create a model from scratch.

Transfer Learning



(a) Traditional Machine Learning (b) Transfer Learning

Transfer Learning



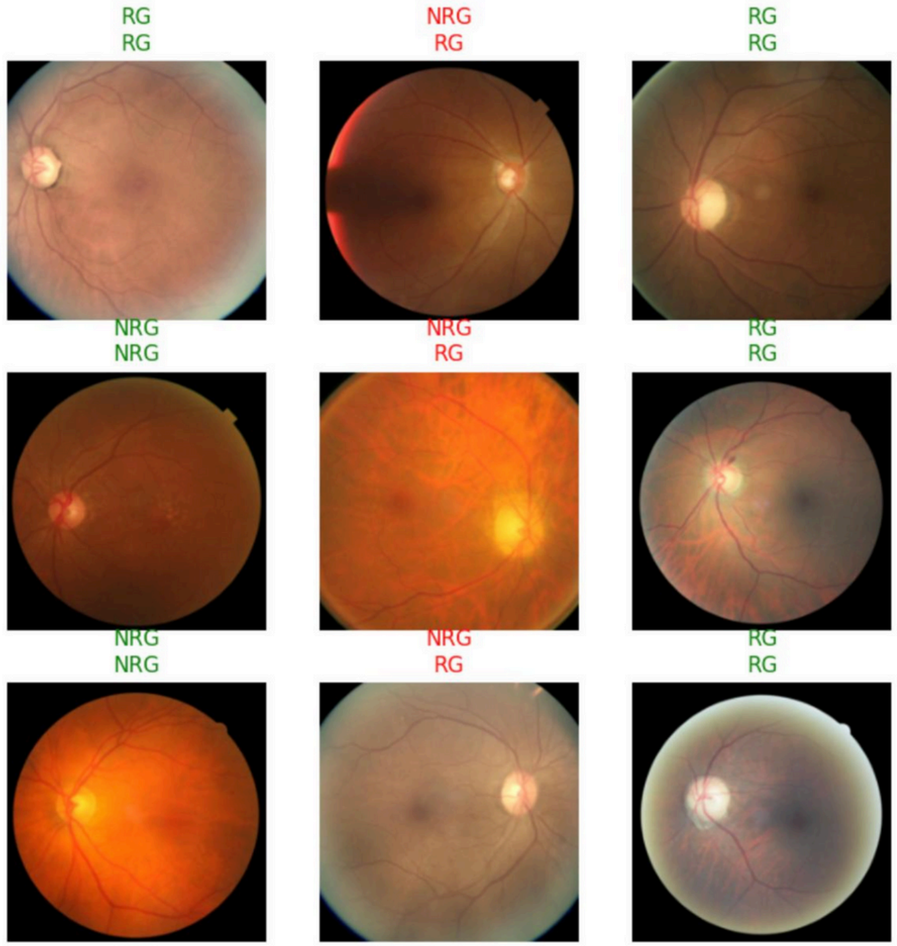
Fine Tuning in Transfer Learning

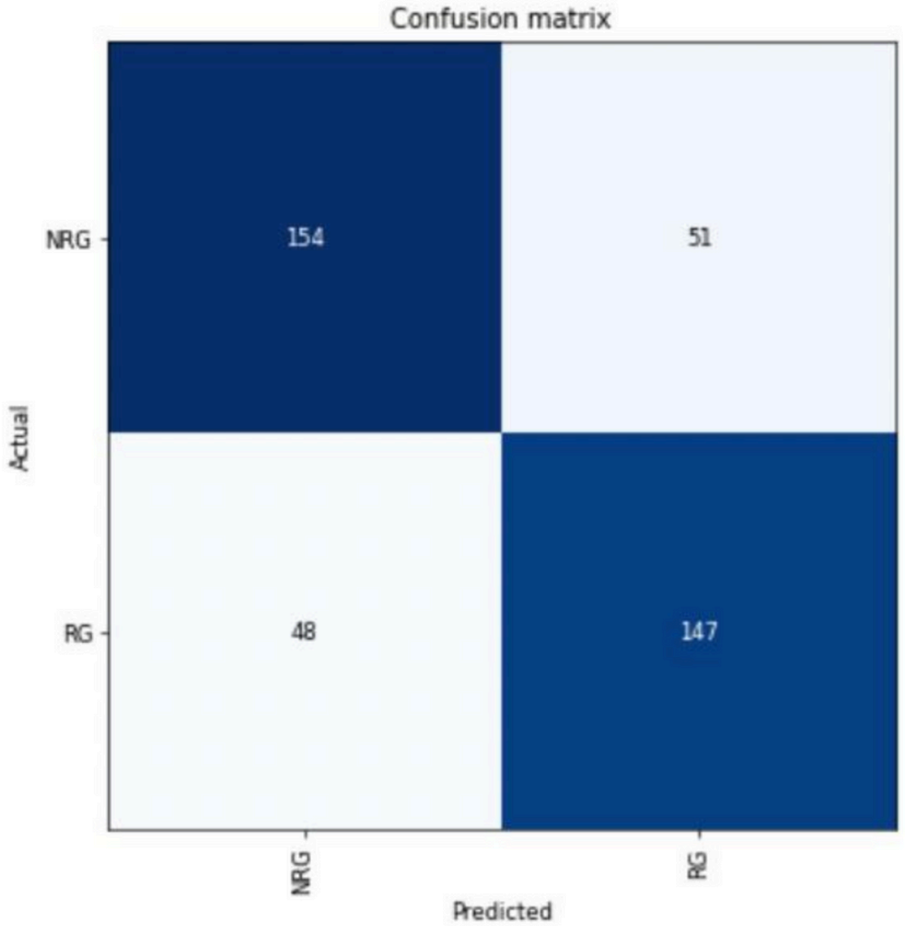
- Fine-tuning is a process that adapts a pre-trained model to a new, related task by continuing the training process.
- It involves making slight adjustments to the weights of an already trained model so that it can perform a new task with a smaller, specific dataset.

Results

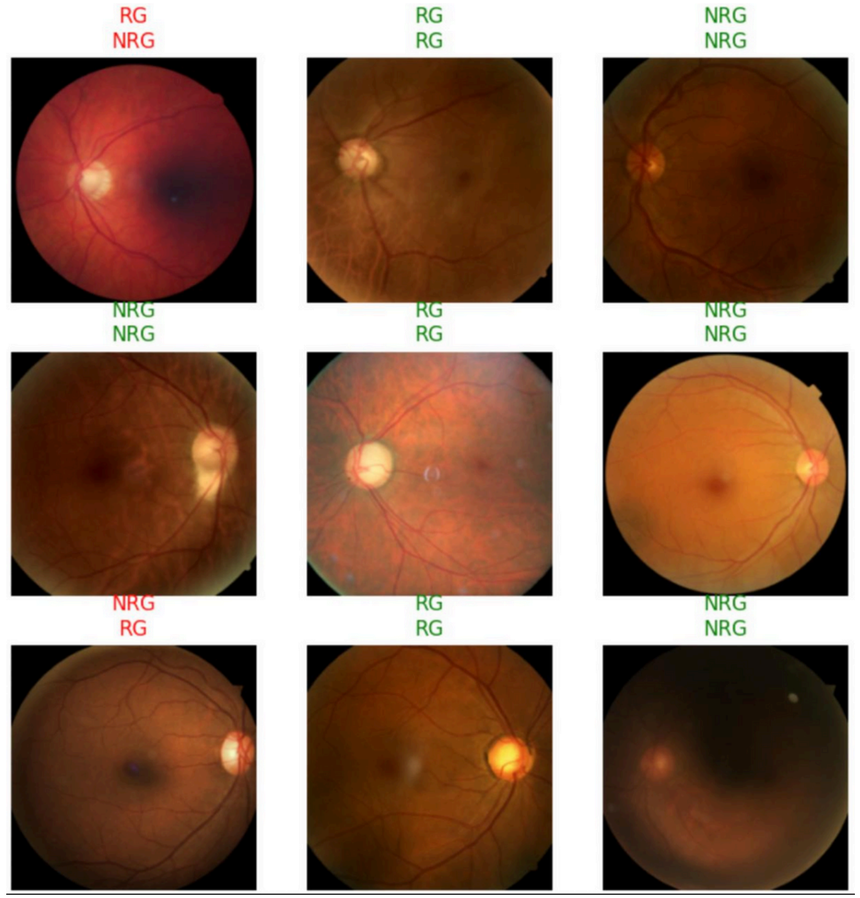
Model	Epochs	Accuracy(in %)
ConvNeXT-Tiny	1	72
ResNet 34	1	72.5
ResNet 34	2	75.5
ResNet 50	2	78
Ensemble of ResNet18, Resnet 34, Resnet 50	3	**80.5**

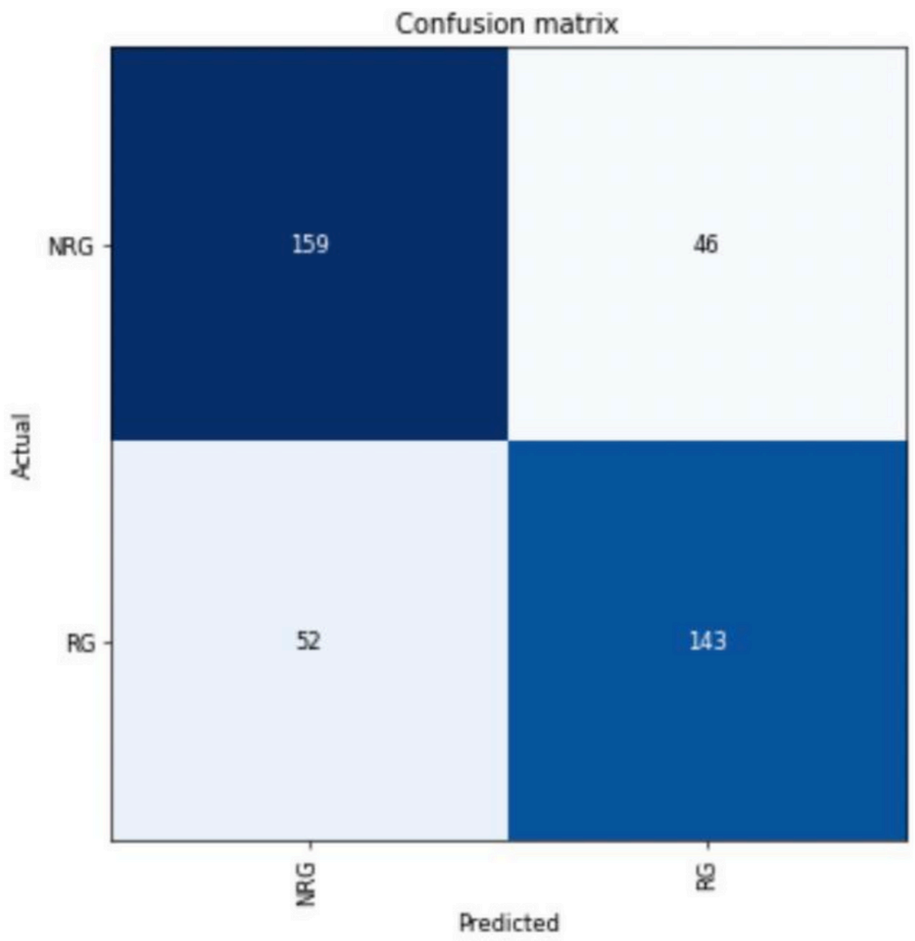
Results for ResNet34 epoch=1



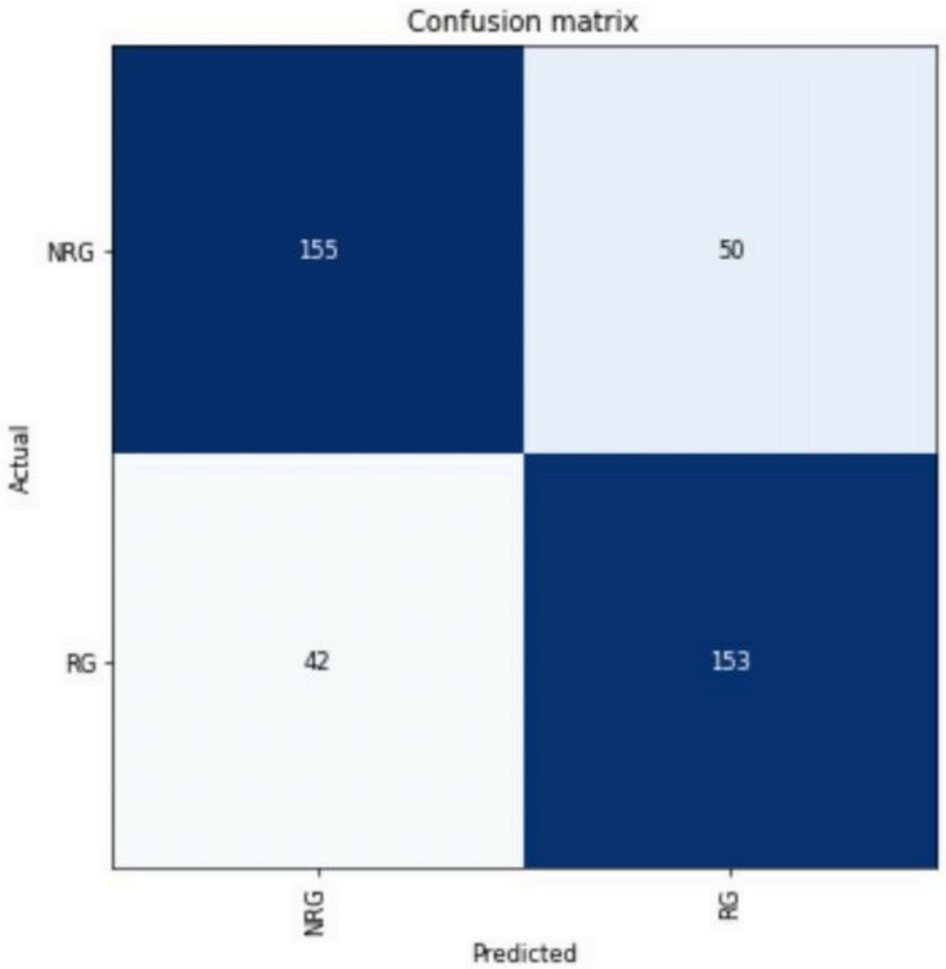


Results for ResNet34 epoch=2

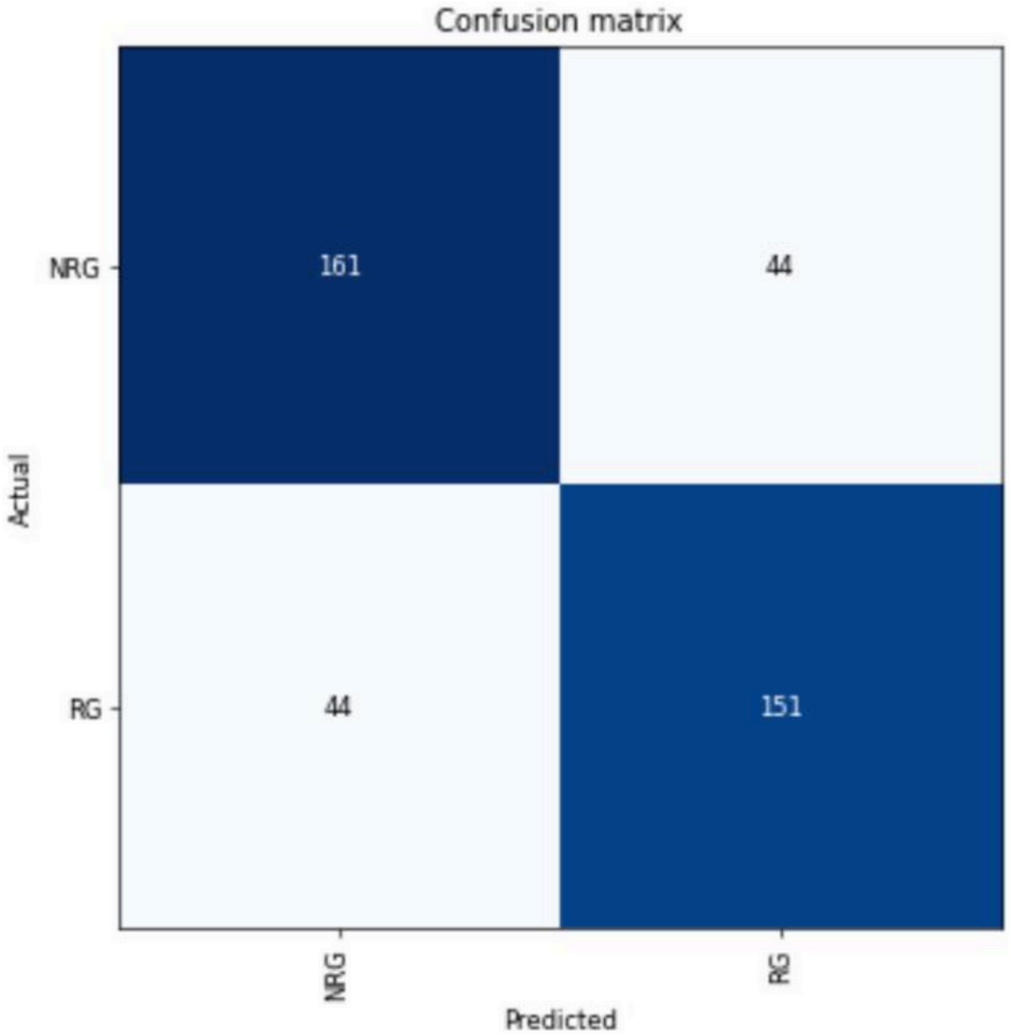




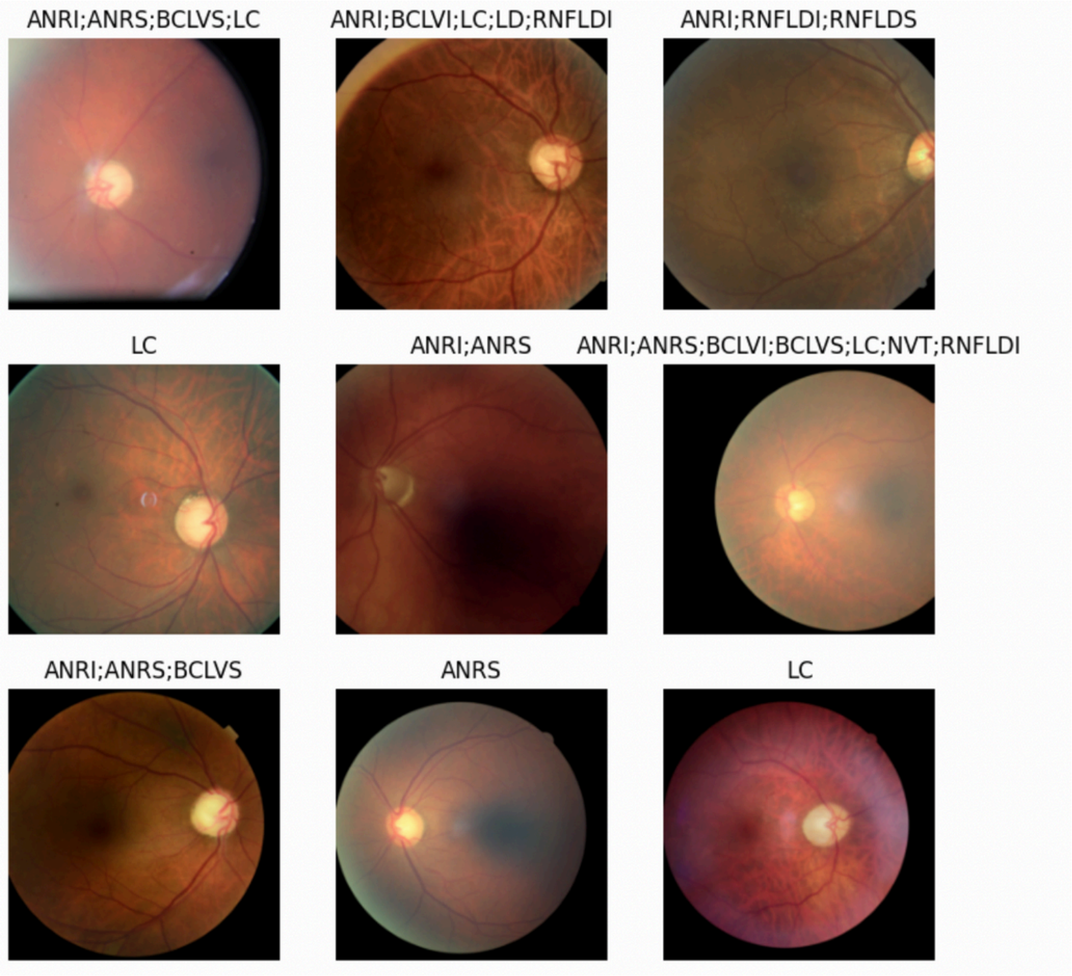
Results for ConvNeXT epoch=1



Results for RESNET 50 epoch=2



Results for Task 2 (Multi-label Classification)



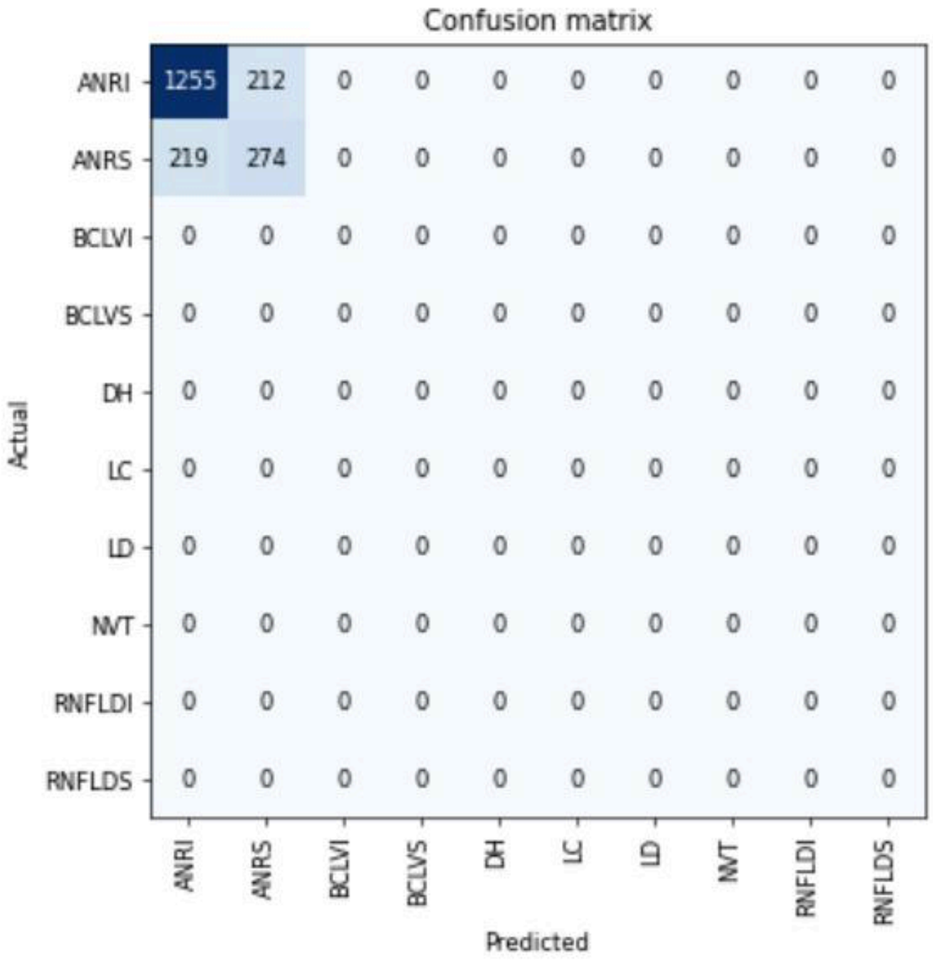
Data Set Specifications

Feature Label	Occurences (out of 982)
ANRI	686
ANRS	605
BCLVS	111
NVT	157
LC	383
LD	395
DH	26
RNFLDS	63
RNFLDI	77

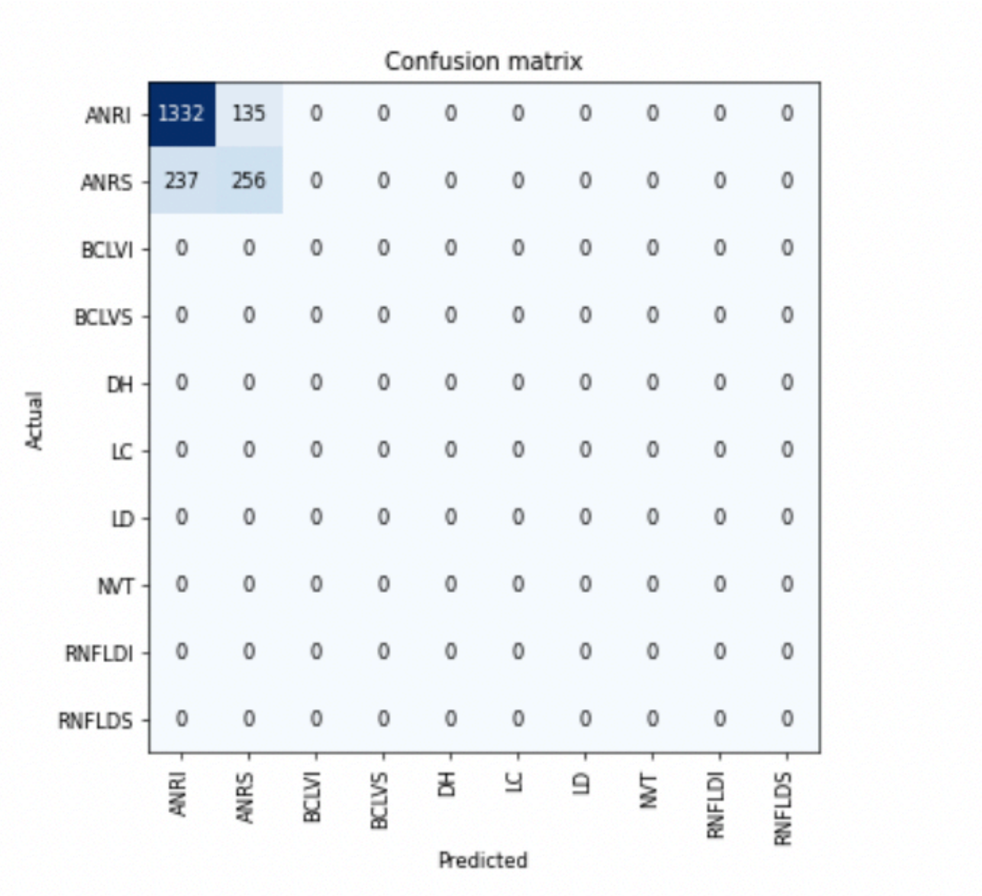
Results

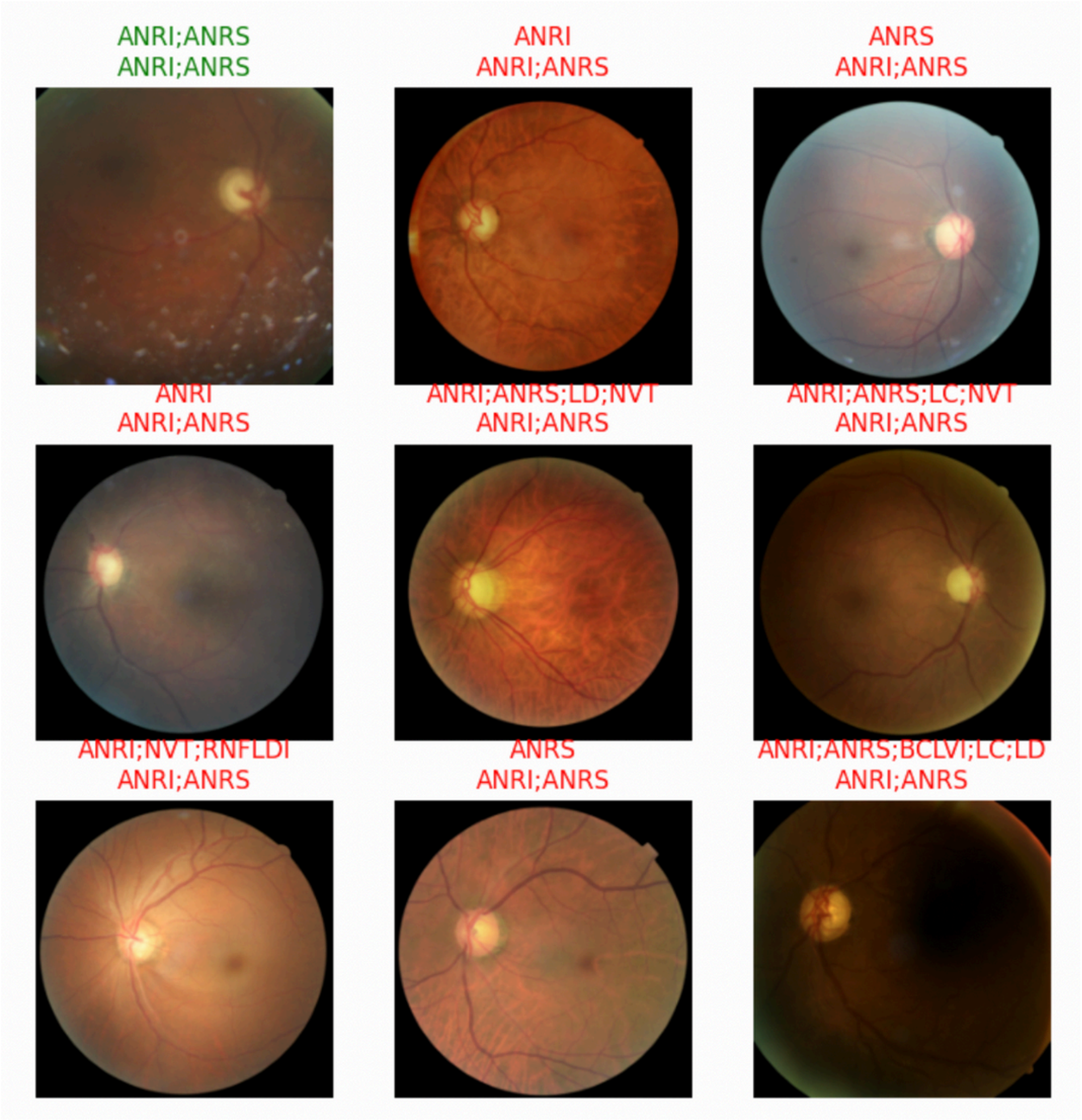
Model	Epochs	Multi-Class Accuracy(in %)
ConvNeXT-Tiny	2	81.9
ResNet 50	2	80.9

Results for RESNET 50 epoch=2



Results for ConvNeXT epoch=2




ANRI;ANRS
ANRI;ANRS
ANRI
ANRI;ANRS
ANRS
ANRI;ANRS
ANRI
ANRI;ANRS
ANRI;ANRS;LD;NVT
ANRI;ANRS
ANRI;ANRS;LC;NVT
ANRI;ANRS
ANRI;NVT;RNFLDI
ANRI;ANRS
ANRS
ANRI;ANRS
ANRI;ANRS;BCLVI;LC;LD
ANRI;ANRS

Future Work

- Fine Tune even bigger models with more computational resources and time.
- Explore boosting and ensemble techniques.
- Explore other loss functions and metrics for Multi Lable, classification.

Thank you