Day-19, Oct-21, 2024

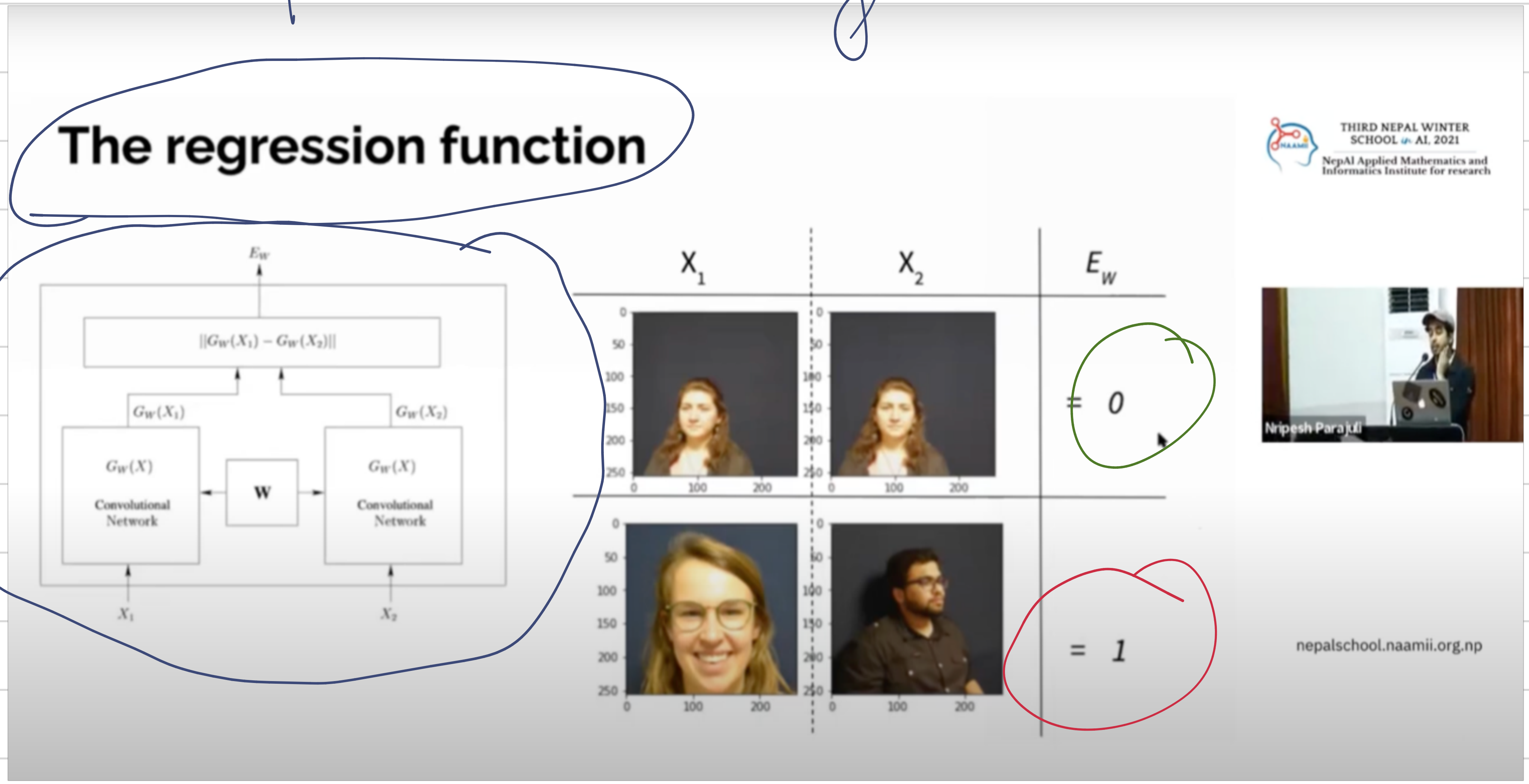
#. Continuing CNNs. (Face Detection - Theoretical Knowledge)

Regression function:

$$E_W(x_1, x_2) = \| G_W(x_1) - G_W(x_2) \| \text{ ——— eqn } ①$$

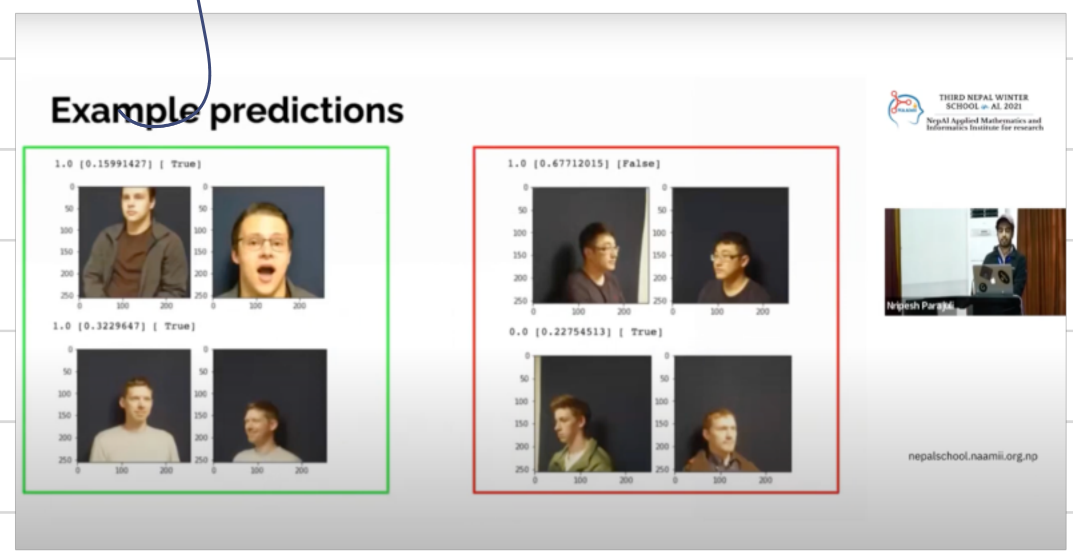
① Loss function, CNNs: → Concept Similar to SVM (Support Vector Machine) loss, within SVM margin there is no penaliaation (penalty).

Ideal Case for face Matching.



Ew Can be 0.12 or 0.15 or 1.15, 1.09

(Example predictions on Data Augmentation!)



Models And Training:

- We'll train 2 types of network with this:
  - Custom: train a model from scratch
    → 3 hidden convolutional layers, followed by 2 dense layers
    ~ 2 million parameters.

- Pretrained: a mobilenetv3 model trained on ImageNet and applied.

Transfer Learning

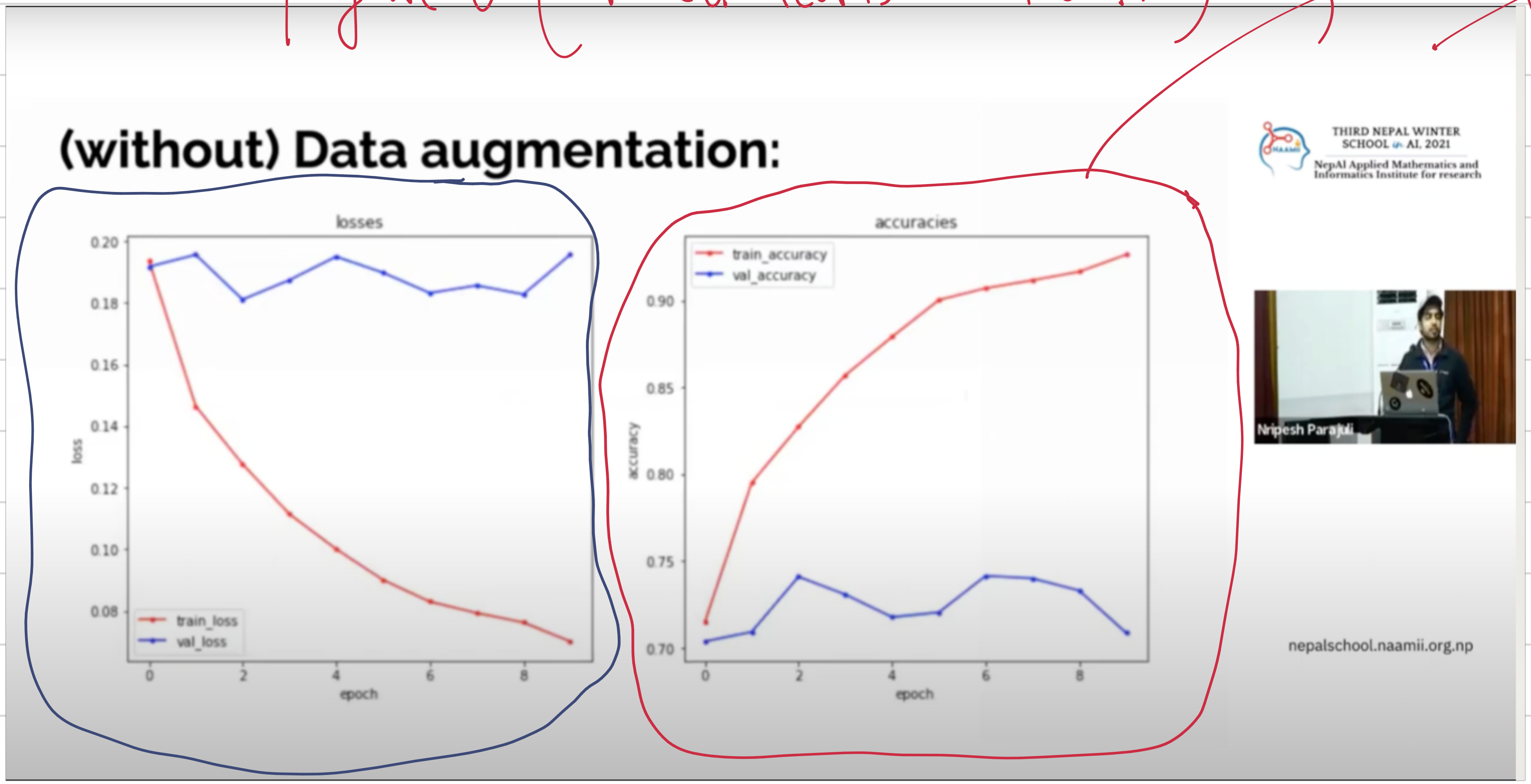
  - Many more hidden layers due to efficient organization of mobilenet.
  - 3.6 million parameters
  - 2.1 million trainable, 1.5 not trainable (frozen mobilenet params)

Few other steps applied/tested

- Data Augmentation (for better generalization)
    - Images are randomly translated, rotated, flipped horizontally
    - Since fitting the data + augmentation in memory is hard, data is generated on fly with the augmentation applied during training.
- To avoid overfitting, models with best validation data performance are used saving call backs. We can optimize on loss or other metrics.
    - Lowest loss
    - Best accuracy

Figure ① (Model Learns noise too) → Overfitting

How to Avoid Overfitting for this ?



→ val_loss
→ train_loss

→ train_accuracy
→ val_accuracy

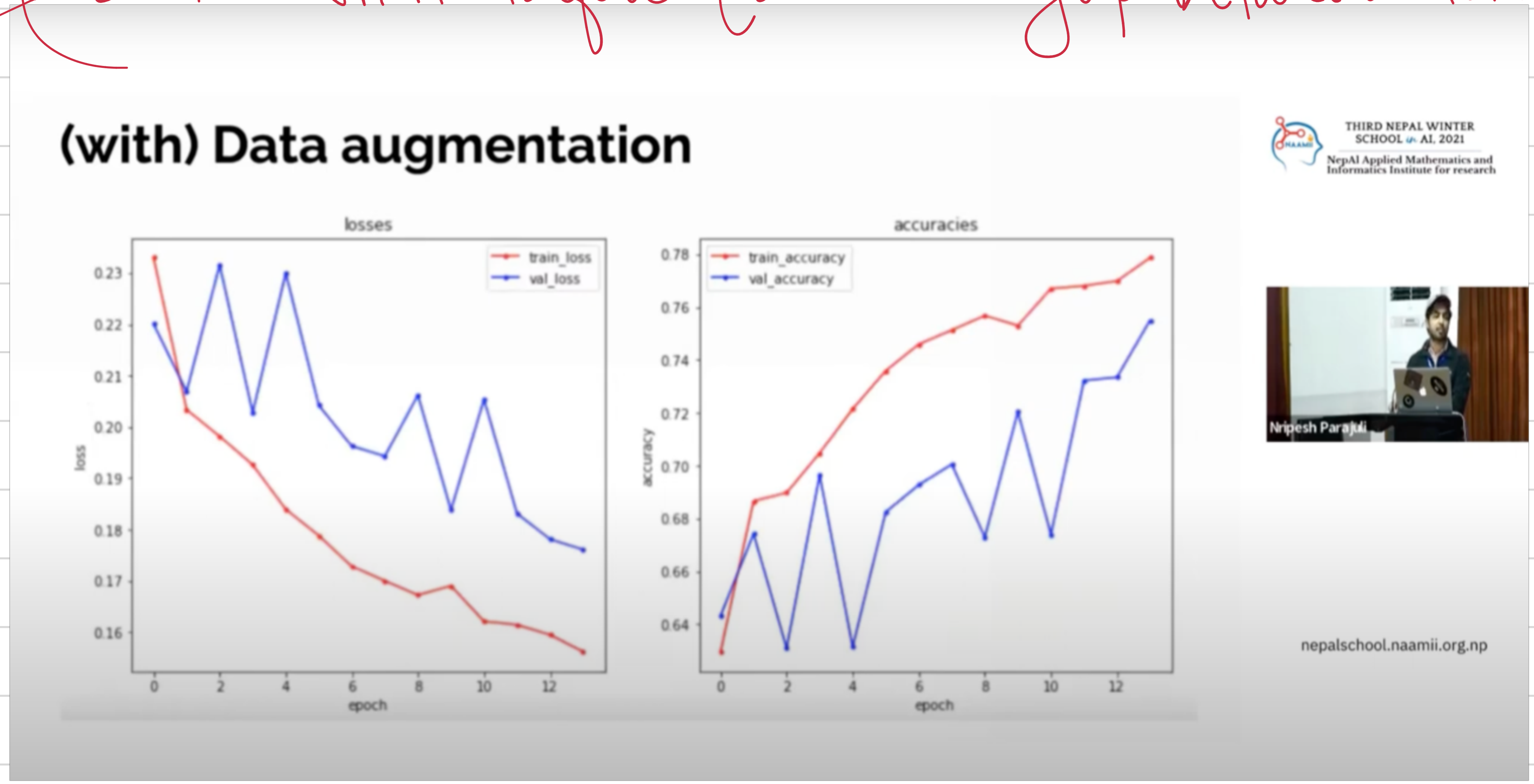
① train Simpler model

② train epoch - 2 or 3 (so the gap is not more.)

③ Decreasing the model size so no good practice (decreases the accuracy).

Of course there are regularization, Cross-Validation, test / train Splitting.

Figure 1. (Better than Before (See the gap between train and val))

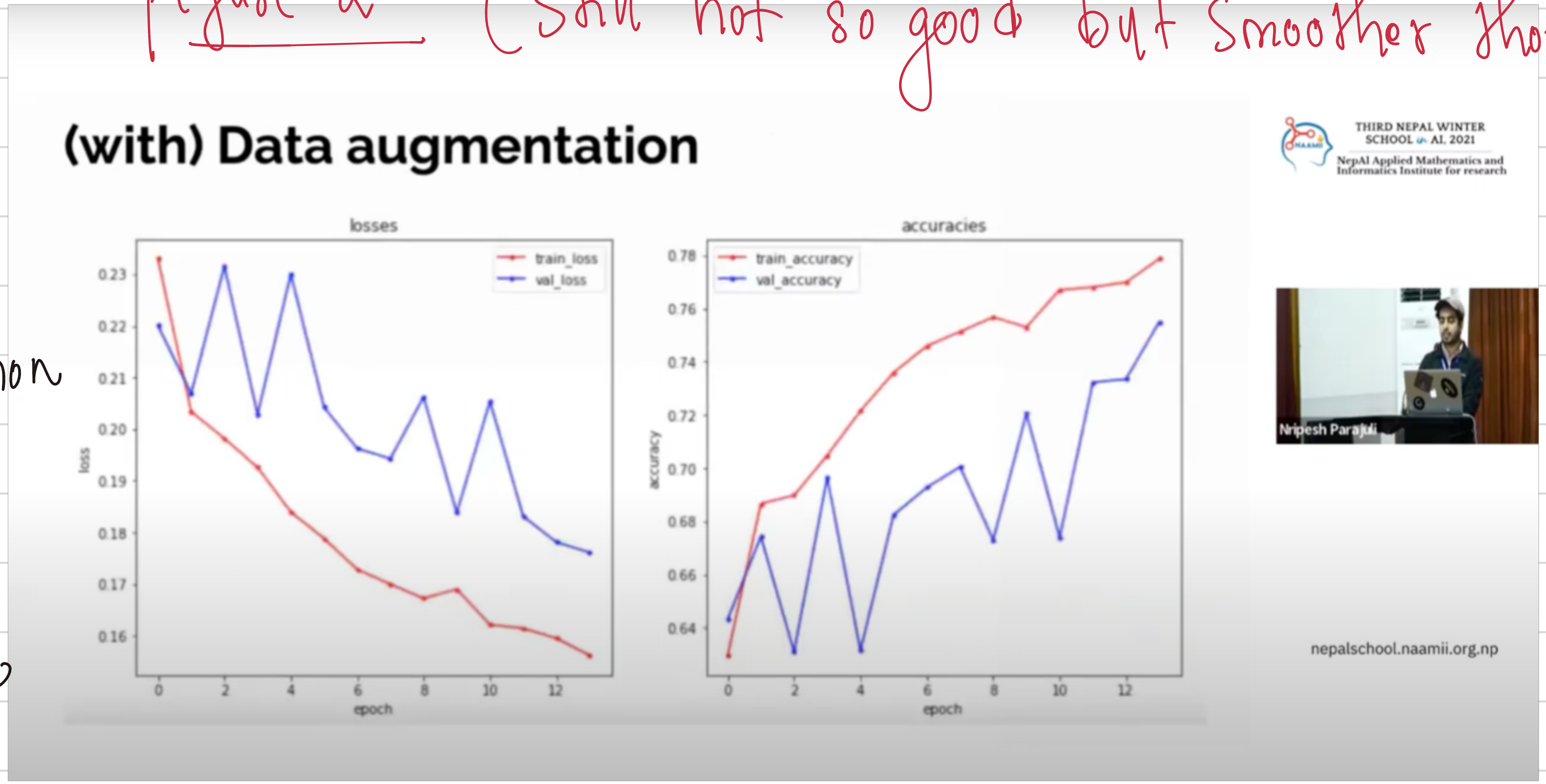


Cons:

① Noisy because looking at the Accuracies points

we are not sure our next point is going to be local minima or maxima, or global minima/maxima.

Figure 2. (Still not so good but Smoother than figure 0)



Tips:

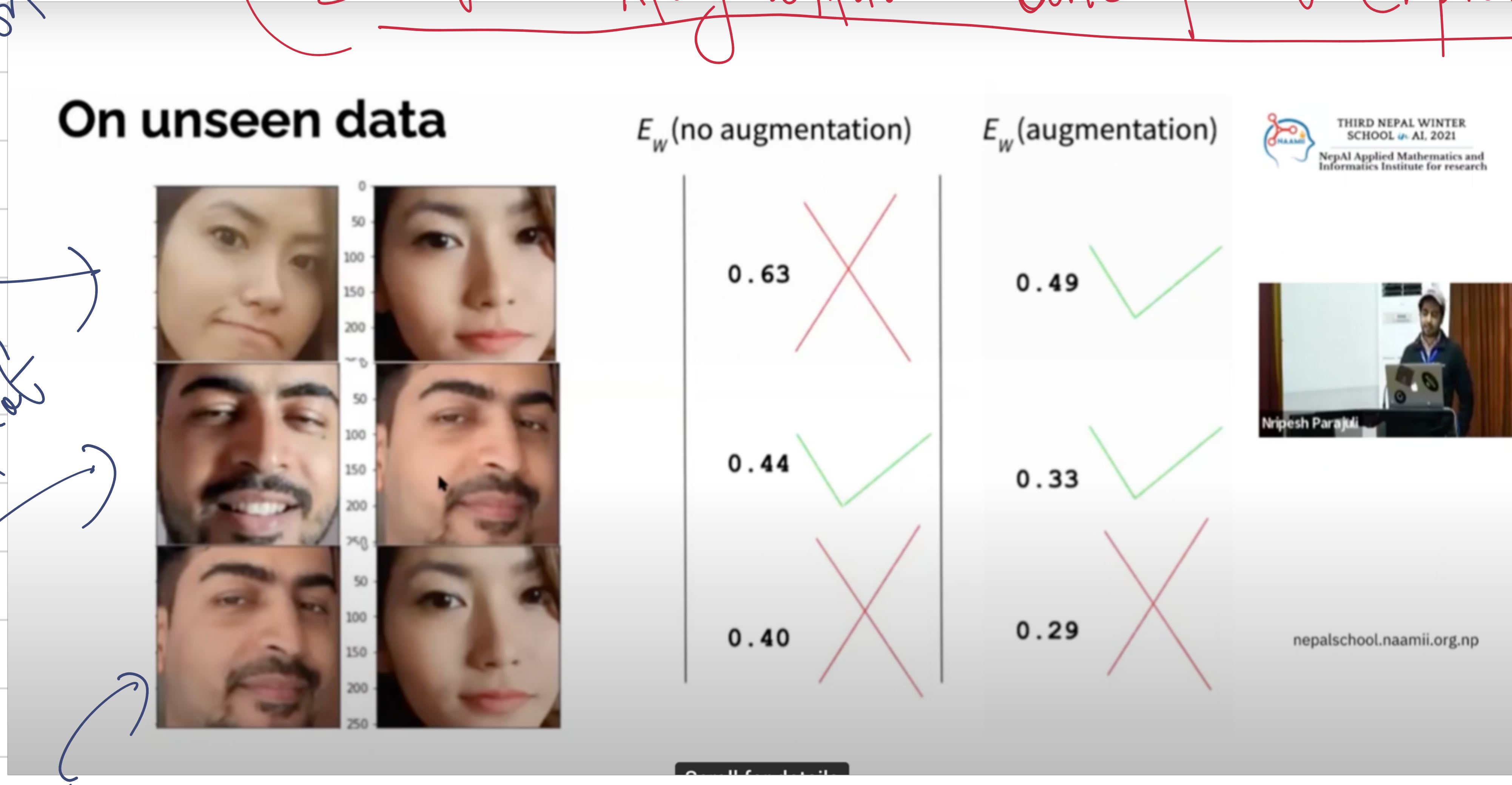
With Augmentation Model also Confront with the noise.

Although there is a noise, Model trying to minimize the loss function.

Better to Evaluate the Model with the statistically varying Input Images.

(See the Augmentation and facial Expression ?)

facial Expression ?



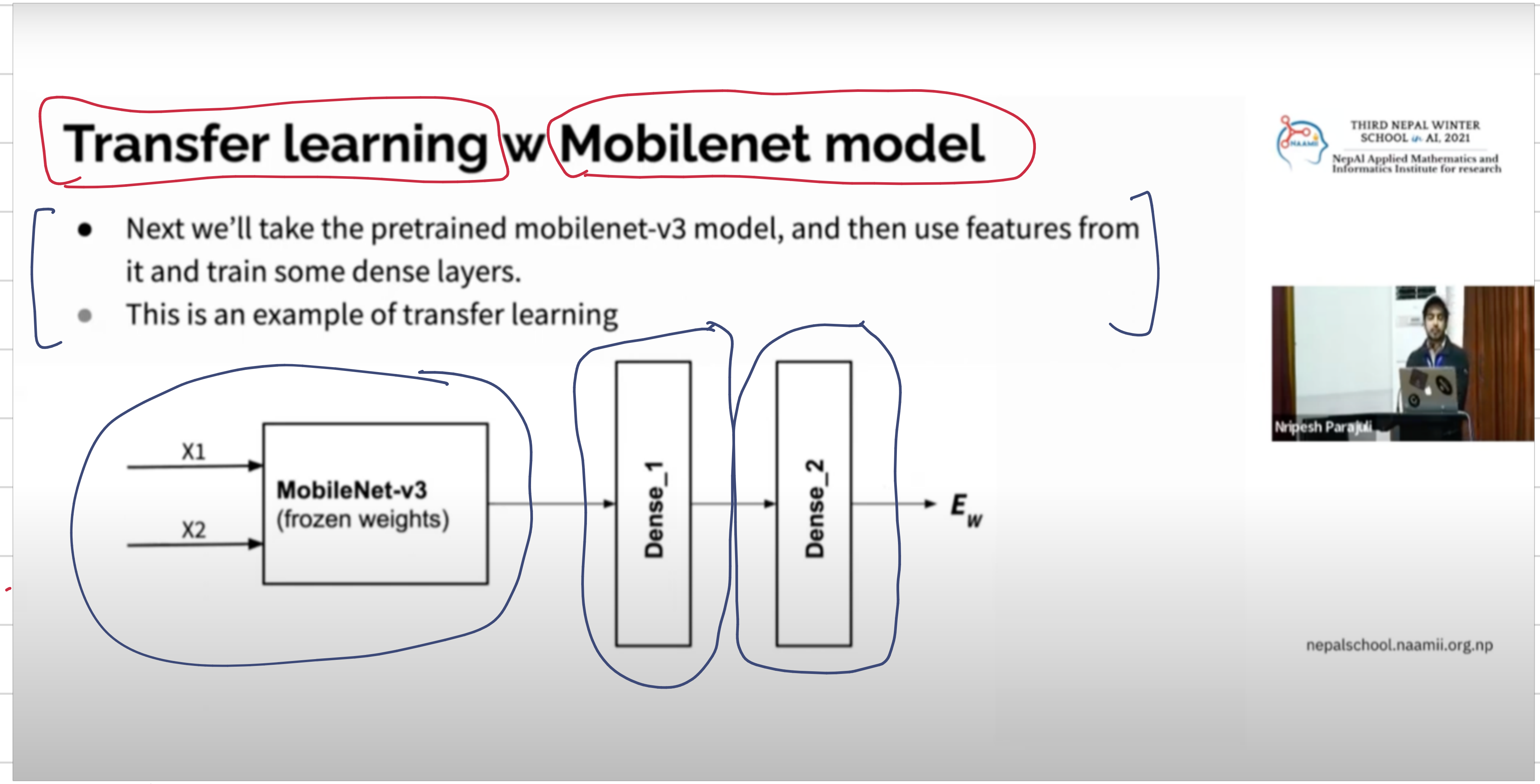
Working Great

Working Great.

Figure- 4. Showing How facial Expression & Augmentation working performance

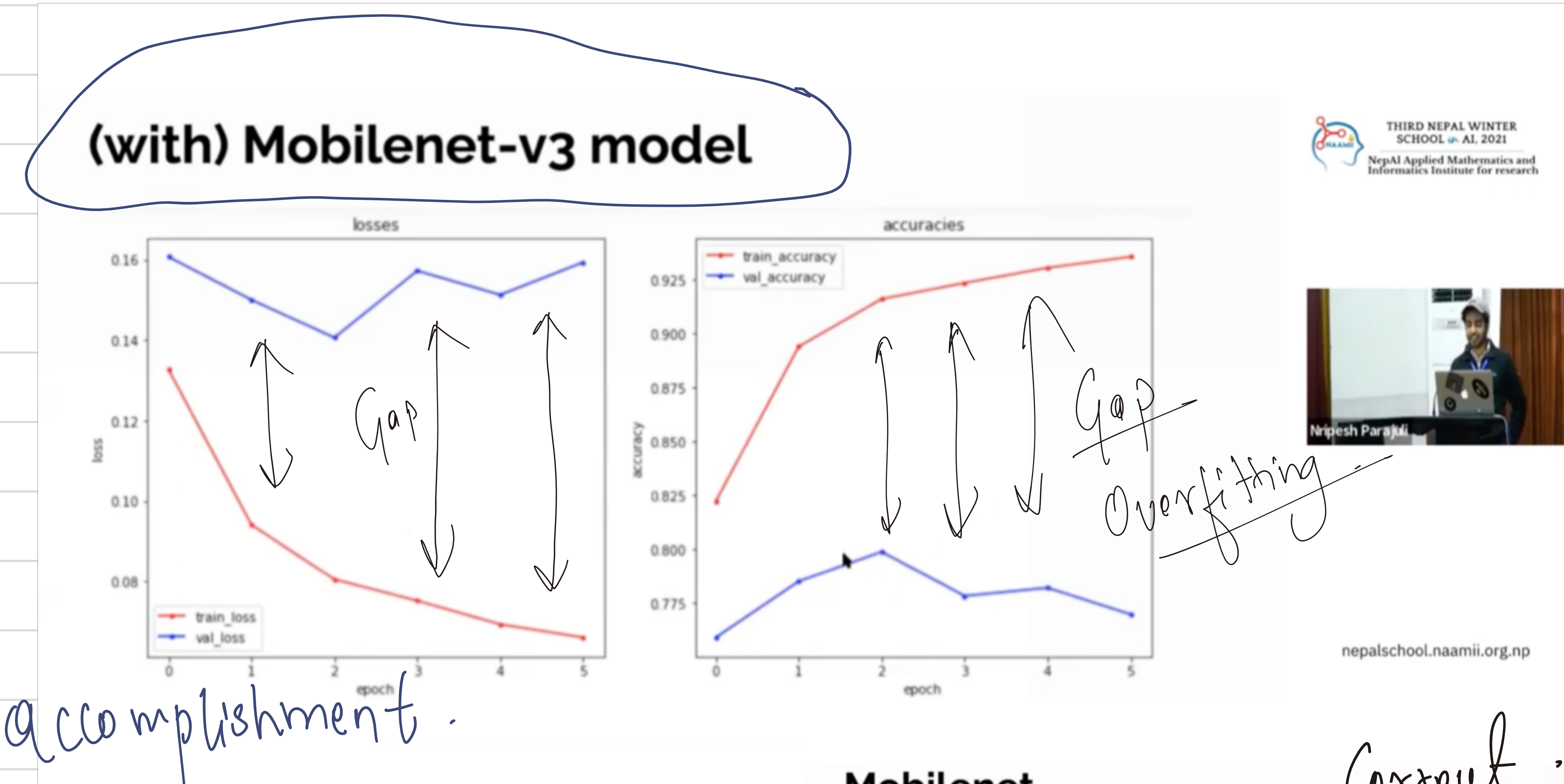
#Mobilenet model is targeted for mobile devices.

#lightweight & efficient model.

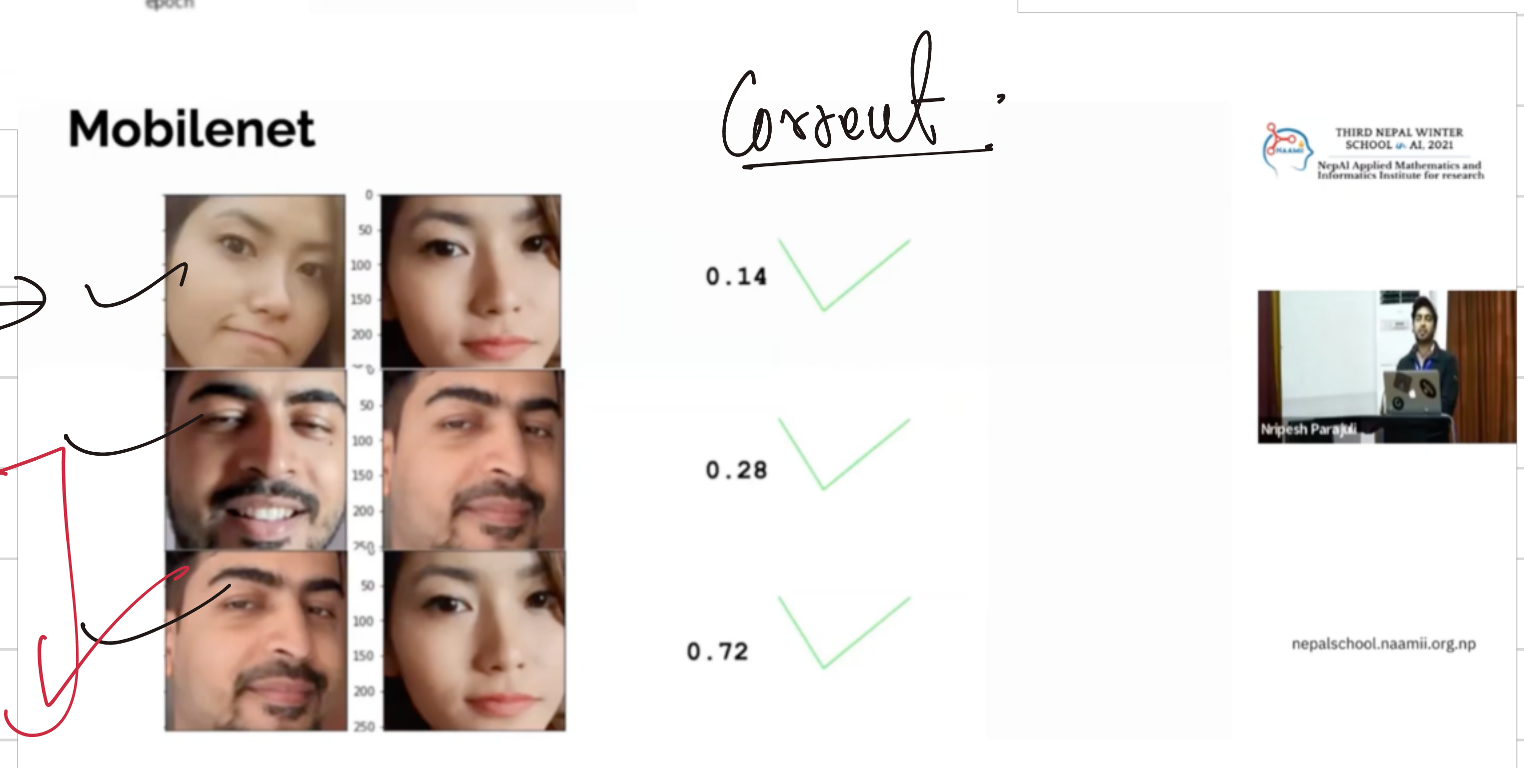


→ There are other types of pretrained models available such as:

Decreasing the model acuracy from 0.75 to 0.80 is big accomplishment.



Even 0.05 is big achievement.



ImageNet - Based Transfer Learning working fine

Summary - → ImageNet → trained on vast categories of Image Data (Everything)

First Layer

Edge Detection,
Gaussian edge,
Generic tasks

→ feature Map →

n - Hidden Layers.

Extracting the features more

→ Flatten →

Output.

Better to use custom CNN model for small dataset

WordNet:

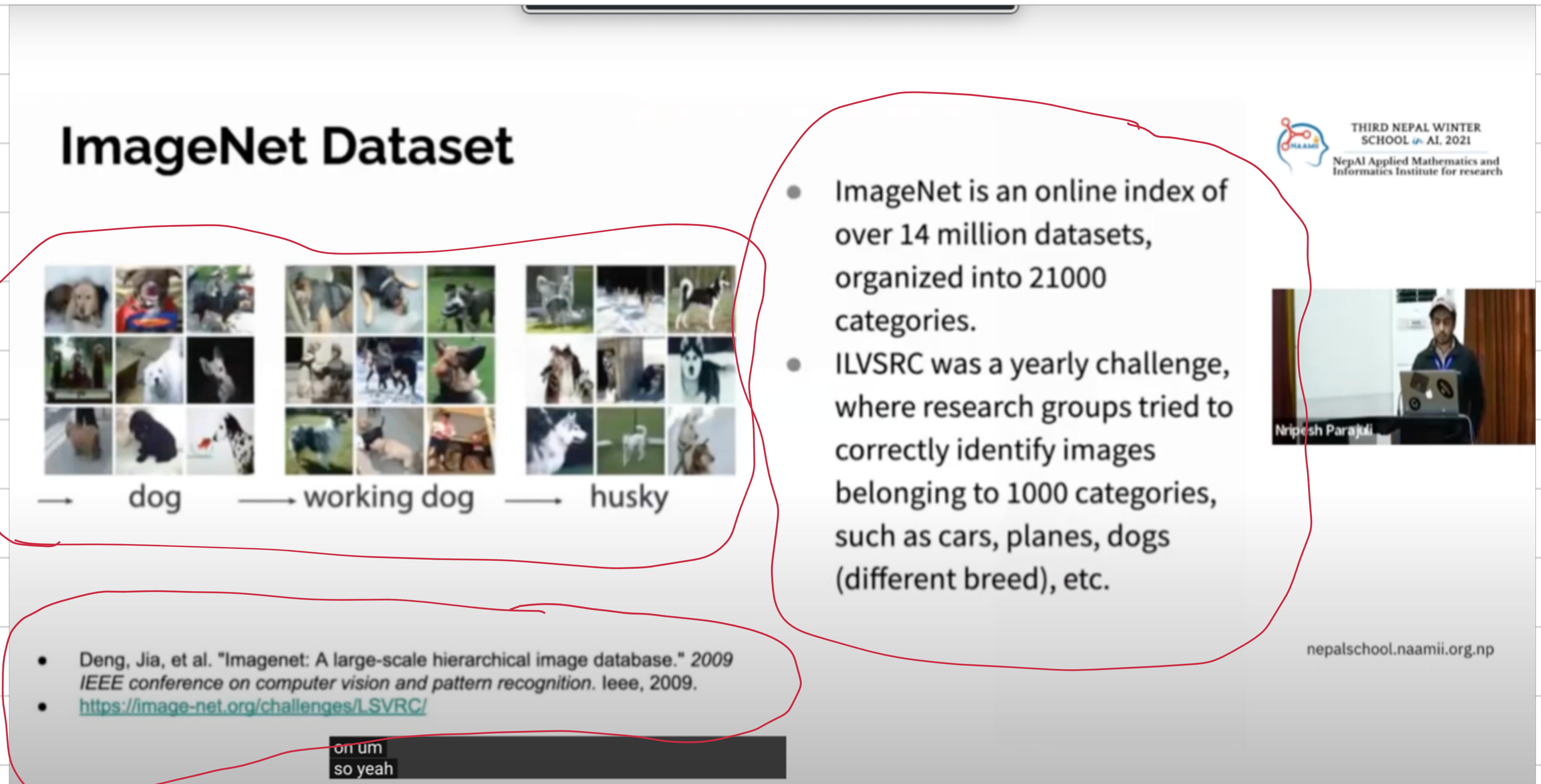
A large-Scale Lexical database of English, designed to represent the meanings of words and their relationships to each other. It is hierarchical form. Synset - dog, Hypernym - animal, Hyponym - puppy - - - -.

CNN Architecture :

Errors Categories.

① Top 1. (Max)

② Top 5



(Top 5)

Suppose 5 class of predictions done by Model. Among 5, one class is only true then the model is Accepted in ILVSRC Challenge.

From Textbook. (FOPL)

#. FOPL (Syntax and Semantics):

FOPL is an extension of propositional logic that allows for the use of quantifiers and predicates, enabling more expressive logical statements about objects and their properties.

FOPL: Syntax and Semantics.

- Syntax: Constants, Variables, Predicates, Functions, Connectives, Quantifiers, Semantics

Constants: Objects in the domain. Eg: John, Apple, Cricket.

Variables: placeholders for objects: Eg: x, y, z.

Predicates Properties or relationship between objects. Example:

loves (John, Marry) → John loves Marry

L (x, y).

Functions: Map objects to other objects. Eg. Father (John)

Could be "The father of John".

Connectives: $\wedge$, $\vee$, $\neg$, $\rightarrow$, $\leftrightarrow$ (if and only if)

Quantifiers: Symbols used in logic to express the quantity of individuals in a domain that satisfy a given condition –

Allow to make general statements about entire groups of objects.

Quantification → process of using quantifier to specify scope of a statement.

Universal Quantifier ($\forall$): statements true for all objects in the domain

$\forall x$ loves (John, x) "John loves everyone"

- $\forall_x$ Open (Bank, x) "Bank opens for everyone"

Existential Quantifier ($\forall$): Statement is true if there exists at least one object for which the statement is true.

$\exists_x$ Loves (John, x) → "John loves someone"

$\exists_x$ Drives (Peter, x) → "Peter drives someone / some car"

Semantics: Statements interprets, Constants predicates and function assigning meaning to them.

Quantification Examples

$$\forall x (P(x) \rightarrow (Q(x) \wedge (\neg J(x) \wedge Q(x)))$$

$$\forall x (P(x) \rightarrow Q(x) \wedge J(x))$$

$$\exists x (P(x) \rightarrow Q(x))$$

$\forall x \exists y \; R(x,y) \rightarrow$ for all x, there exists a y such that $R(x,y)$ is true.

Inference in FOPL involves drawing conclusions from given premises.

FOPL $\xrightarrow{\text{Convert}}$ PL $\quad \forall x (\text{Human}(x) \rightarrow \text{Mortal}(x))$

(Modus Ponens)

x is Socrates $\quad$ So ① $P \rightarrow Q$ $\quad$ ② P $\quad$ i.e. Mortal(Socrates)