

# ML in Fashion: Using Deep Learning to recognise Indian Classical Sarees



Sudip Kumar Kar & Ritadip Bharati  
Batch 19

Instructor: Dr. Subhankar Mishra

# Recap

---

- The Original idea was to train a Machine learning algorithm to learn to recognize Indian Classical Sarees.
- This was to be done in two steps: first, a dataset was to be collected and next, this dataset was to be used to train the model.
- As a starting point, it was decided that a simple algorithm like TinyVGG would suffice.
- If successful then we were to move to more sophisticated algorithms like Mobilenet.
- Here, we present our progress on this path thus far.

# Dataset

- Web Scraping was done using Python selenium libraries
- Four different kinds of sarees were chosen: Sambalpuri, Ilkal, Kasavu, and Bandhani.
- After this we had about 4.5k images to work with.



Table 1: Indian Saree Dataset description

Saree type	Total number	Training	Testing
Bandhani	288	231	57
Ilkal	476	382	94
Kasavu	1536	1229	307
Sambalpuri	2388	1911	477
Total	4688	3753	935



Bandhani



Ilkal



Kasavu



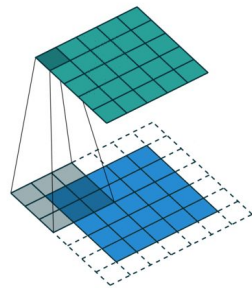
Sambalpuri

# Convolutional Neural Networks

- In one dimension, convolution is given as

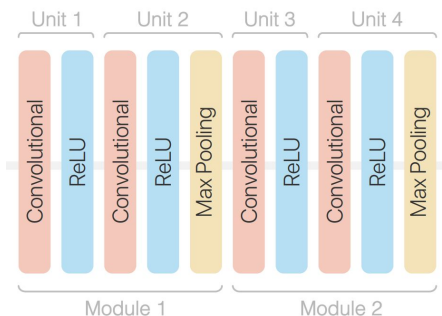
$$h(z) = \int_{-\infty}^{\infty} f(x)g(z-x)dx$$

- Convolution can be thought of as looking at something through a different perspective, literally! Visualise it as a window sliding over the original distribution to yield a new one.
- Doing so augments the original image data, this is leveraged by the algorithm to classify effectively.



# Tiny VGG

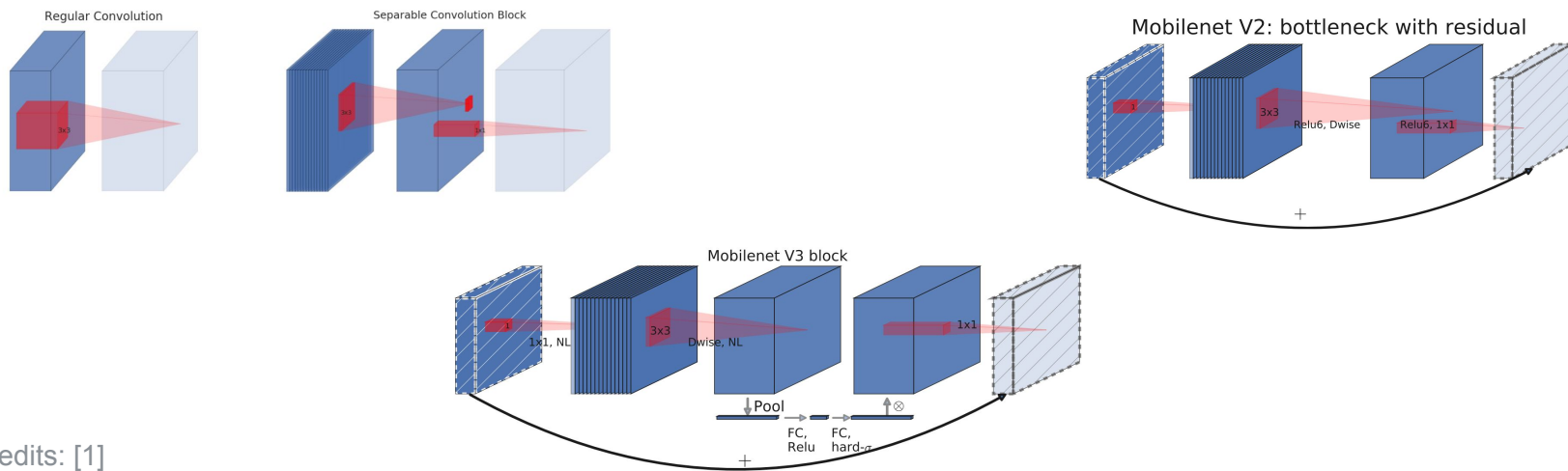
- The Tiny VGG architecture uses multiple convolutional layers to effectively classify images.



- The output of this is fed into a flattening layer with linear activation to give logits that can be used to determine the linear activation

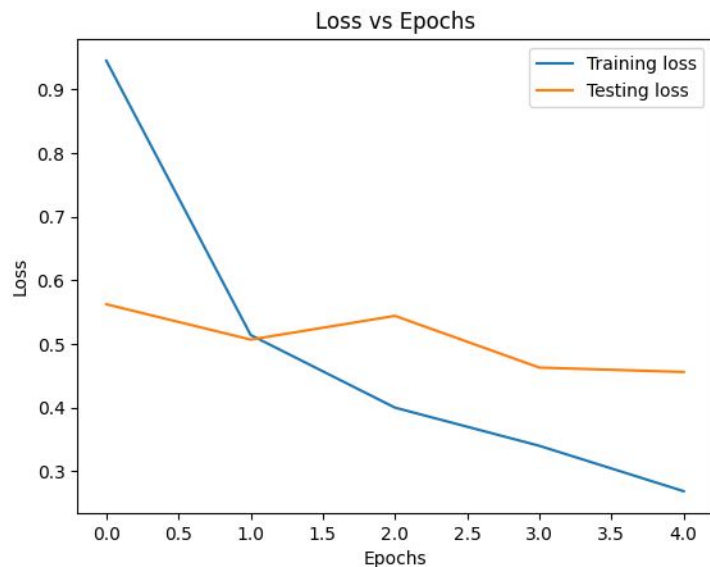
# Mobilenet

- Mobilenet is a series of Image Classification models designed for low end mobile devices, it reduces computational cost by splitting convolution into a Depthwise and a Pointwise block, the upgraded version uses additional Inverted residuals and squeeze and excite layer to improve efficiency.



# Experiments

- Tiny VGG performed surprisingly well for a small model.

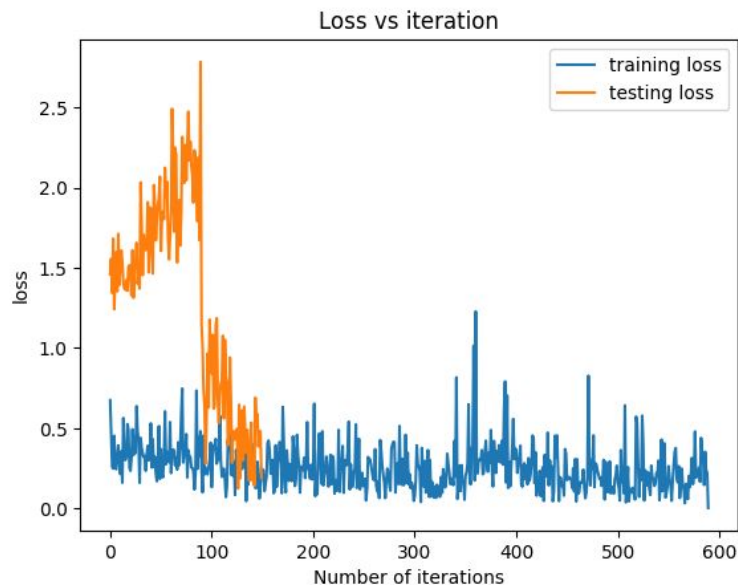


TinyVGG



# Experiments

- MobileNet V3, despite being a larger model, underperformed.



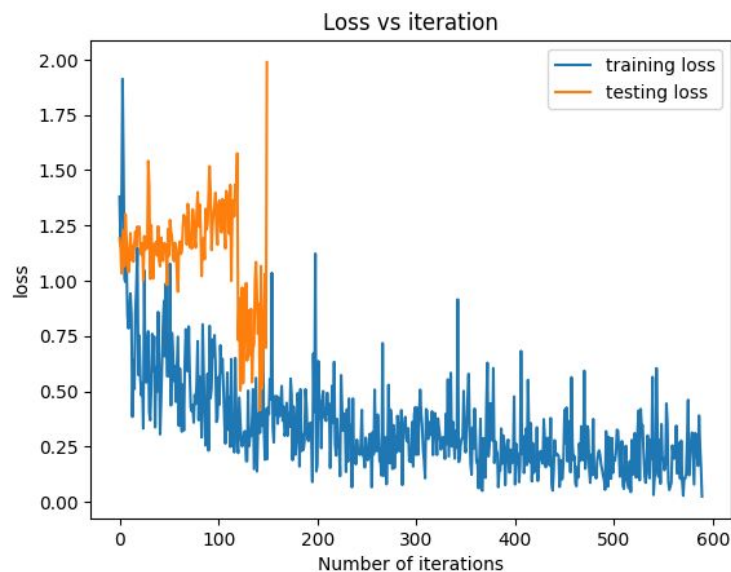
MobileNET v3 after 5 epochs



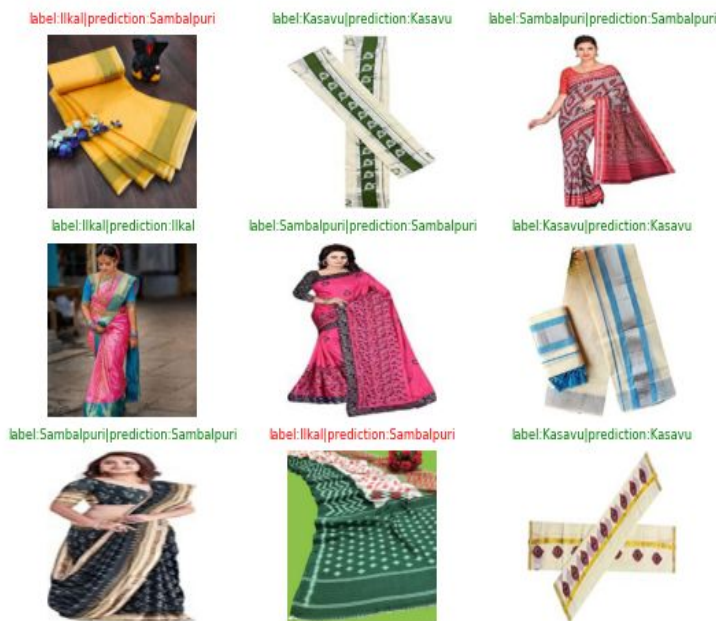


# Experiments

- The modification of MobilenetV3 by including ECA layers did not affect its effectiveness.



MobileNET v3 ECANET



# Future Plans

---

- Improve upon the current accuracy by extending dataset, tweaking hyperparameters and trying out new models that are specialised for fashion.
- Try out some non-CNN based models to gauge the relative accuracy of CNN models
- Try to build a Generative Adversarial network that can replicate Indian Saree patterns.

# References

1. Andrew Howard, Mark Sandler, Grace Chu, Liang-Chieh Chen, Bo Chen, Mingxing Tan, Weijun Wang, Yukun Zhu, Ruoming Pang, Vijay Vasudevan, Quoc V. Le, and Hartwig Adam. Searching for mobilenetv3.CoRR, abs/1905.02244, 2019.
2. Harshil Jain, Rohit Patil, Utsav Jethva, Ronak Kaoshik, Shaurya Agarawal, Ritik Dutta, and Nipun Batra. Generative fashion for indian clothing. In Proceedings of the 3rd ACM India Joint International Conference on Data Science & Management of Data (8th ACM IKDD CODS & 26th COMAD), CODS-COMAD '21, page 415, New York, NY, USA, 2021. Association for Computing Machinery.
3. Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. 2018.
4. Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2014.
5. Qilong Wang, Banggu Wu, Pengfei Zhu, Peihua Li, Wangmeng Zuo, and Qinghua Hu. Eca-net: Efficient channel attention for deep convolutional neural networks. CoRR, abs/1910.03151, 2019.
6. Zijie J. Wang, Robert Turko, Omar Shaikh, Haekyu Park, Nilaksh Das, Fred Hohman, Minsuk Kahng, and Duen Horng Chau. CNN explainer: Learning convolutional neural networks with interactive visualization.CoRR, abs/2004.15004, 2020.
7. Vincent Dumoulin and Francesco Visin. A guide to convolution arithmetic for deep learning.ArXiv e-prints, mar 2016.