**Things That didn’t work:**

* Polynomial LR
* Pointwise convolutions to reduce the Z-axis
* L2 regularization (Weight Decay) (not a significant improvement in generalization at 1e-5
* SGD with momentum

**Things That Did work:**

* Rohan’s LR
* Adam optimizer converged faster
* Padding in all Conv layers
  + Padding increased layers. More layers better accuracy and better generalization in this case
  + Padding helped in identifying images that are extended till the borders of the image

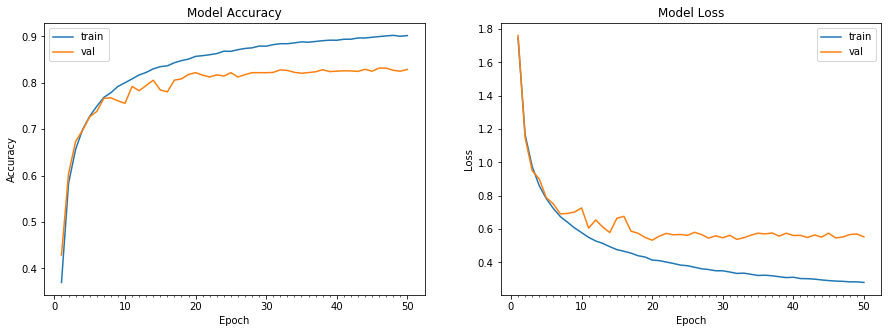
**Tips and Tricks**

* Overfitting is not a major problem as we must improve the accuracy by only 3-4 percentage points
* As overfitting is starting after 10 epochs, we can quickly prototype the hyperparameters or different architectures within 20 epochs
* Any better strategy should yield clear gains in first 10 epochs.
  + In my case both padding, increase in layers and Rohan’s LR improved accuracy within 10 epochs

**Things That Might Work:**

* Data augmentation would have increased validation accuracy

Final results (with 0.1 dropout)



Results with 0.2 dropout

