Surgical Tool Annotation Using Deep Convolutional Neural Network

Avinash Kori*[0000−0002−5878−3584], Varghese Alex*[0000−0001−5095−2358], and Ganapathy Krishnamurthi*[0000−0002−9262−7569] ⋆

Department of Engineering Design, Indian Institute of Technology Madras, India
gankrish@iitm.ac.in

1 Introduction

Automatic detection of surgical tools in the field of view is the preliminary step for developing automatic or robotic surgical arms. We make use of a multi-resolution CNN based approach for automatic detection of the tools.

2 Material and Methods

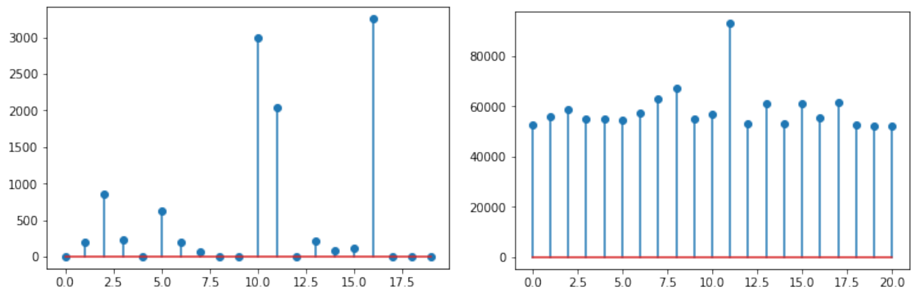
2.1 Data

In this manuscript we make use of the dataset that was made available as part of the Cataract-2018 challenge. No additional data was used to train, validate the processed the network.

2.2 Pre-Processing of Data

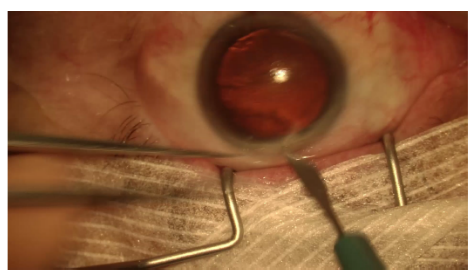
Oversampling of data was done to address the problem of huge class imbalance, bellow figure shows the frequency of occurrence of each class in training set before and after oversampling. Data was split into 70:15:15 ratio for training, validation and testing respectively.

⋆ All authors contributed equally.

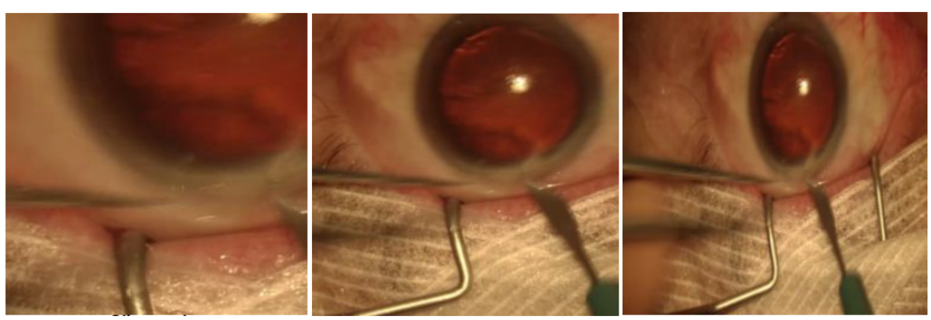


(a) Before (b) After

Fig. 1: Distribution of various classes in the entire training database. a) Before Over-sampling, b) After oversampling under represented classes.



(a) Original Image



(b) Glimpse-1 (c) Glimpse-2 (d) Glimpse-3

Fig. 2: Original Images & various Glimpses fed as input to the Network. Glimpses at various resolution provide information about contact with eye ball & further information about tool

The images were further normalized to using the statistics (Mean & Standard Deviation) derived from the ImageNet dataset.

2.3 Training and Inference

Three different glimpses (around center) were extracted from the input image. The original image, centre crop of of 256×256 and 512×512 form the glimpse. These images with 9 channels were used for training the proposed network. A convolutonal layer was appended to a pre-trained DenseNet 121 to form the model for the required classification task. The number of classes in the pre-trained model was modified to be equivalent to the number of classes in the Cataract-2018 dataset.

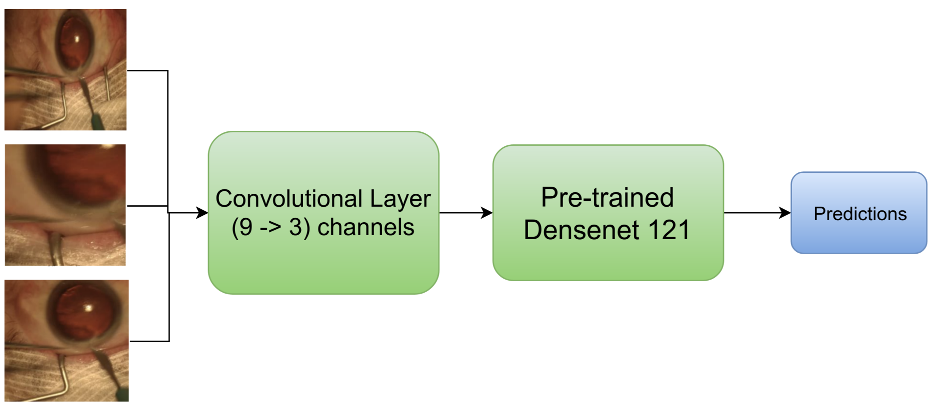


Fig. 3: Network architecture used for surgical tool annotation

3 Results

On the held out test data, the proposed methodlogy achieved a mean AUROC equivalent to 0.998. The class-wise AUROC is tabulated in Table 1.

4 Conclusion

This report presents a framework used for surgical tool annotation by using deep convolutional neural networks. Based on our evaluation of algorithm on held out test data, network achieves over 99.8 % AUROC performance on average.

Table 1: Performance on held out test data

Class	Class Wise AUROC	Class	Class Wise AUROC
Biomarker	0.99994	needle holder	0.9999
Charleux cannula	0.99991	Irrigation / aspiration handpiece	0.9998
hydrodissection cannula	0.99993	phacoemulsifier handpiece	0.9999
Rycroft cannula	0.9991	vitrectomy handpiece	0.9999
viscoelastic cannula	0.9993	implant injector	0.9999
cotton	0.9998	primary incision knife	0.9997
capsulorhexis cystotome	0.9998	secondary incision knife	0.9996
Bonn forceps	0.9970	micromanipulator	0.9993
capsulorhexis forceps	0.9989	suture needle	0.9997
Troutman forceps	0.9983	Mendez ring	0.99999