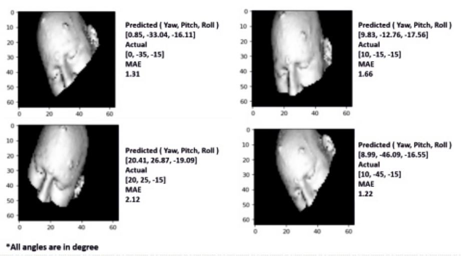
Deep Learning

9

Deep Learning based Head Pose Estimation & Fiducial Localization for Neuro Navigation

A. Bhardwaj*, P. K. Gaur, P. P. Shete, Venkata P. P. K., and D. M. Sarode

Computer Division, Bhabha Atomic Research Centre, Mumbai 400085, INDIA



Predictions over 3D head volume (MRI data) with ground truth

ABSTRACT

Image guided neurosurgical systems have proven very effective in terms of reducing hospitalization and recovery time due to their minimally invasive and precise nature of neurosurgical procedures. In this regard, an initiative has been undertaken by DRHR and CD[1]to extend affordable high-quality neurosurgical practice in India. Estimation of head pose of patient and establishing its relationship with medical imaging data is crucial for surgery planning and patient placement. Once the patient is in surgical setting, fiducial markers affixed on patient's head are localized[2] to establish a complete process of registration between the patient and medical imaging data. Deep learning-based head pose estimation system described in this article has achieved an accuracy of 4-degree MAE (Mean Absolute Error)over a typical patient as well as for medical imaging data. Further, fine sub-slice based fiducial localization technique has achieved 30% reduction in localization time with significant improvement in precision and recall.

KEYWORDS: *Medical imaging, Head pose estimation, Image guided brain surgery, deep residual network, Fiducial localization*

Introduction

Minimally invasive surgery can be performed with distinctness, less bleeding and having faster healing time using robots. Owing to the anatomical complexities of the brain, application of robotic systems in neurosurgery is limited as compared to other medical fields. Even a minor mistake during surgery can lead to disastrous effect. Hence, the biggest problem faced by the neurosurgeons is the visibility inside the patient's skull while performing the brain related procedures. To have visibility of the brain interiors, surgeons are heavily dependent on the medical imaging data both for diagnosis as well as for surgery planning.

Head pose of patient during medical imaging and surgery may not be same. Also, due to constrained motion of robot, sometimes we need to orient patient to ensure reachability of robot and surgical tool. Hence, we need to align the medical imaging head pose with the head pose of patient during the surgery, i.e., registration. This article presents two crucial components of Neuro navigation system,

1. Head pose estimation of patient and medical imaging data

2. Gross localization of fiducials

The estimated head pose information is utilized to accomplish coarse alignment of patient and medical imaging head volume, where as the localized fiducial coordinates are utilized to provide guidance in establishing a fine registration between patient and medical imaging data.

Head pose estimation for patient and medical imaging data

Head poses of both patient and medical imaging data are required to establish correspondence between them. This correspondence is required in surgery planning, patient placement and successful execution of surgery. For estimating the pose angles, Deep learning based CNN architectures are used. The following sections will provide details of the data sets, architecture, training of the model and results.

Dataset

For human head pose estimation, we have used open source labelled data sets like BIWI Kinect head pose[4], 300W_LP[5] and AFLW 2000[6]. For medical imaging head pose we have generated two different data sets (of different sizes, 50k and 150k) using anonymized MRI samples.

Solution architecture

Proposed method for patient head pose estimation takes RGB frame as input and applies face detection over it. Detected face data is passed to a Convolutional Neural Network to predict Yaw, Pitch and Roll angles. In our initial experiments, we have modified AlexNet architecture for prediction, where we observed accuracy saturation below our desired level (Fig.2). In our later experiments, we have used a modified ResNet50[3] architecture to predict the corresponding pose angles. Fig.1

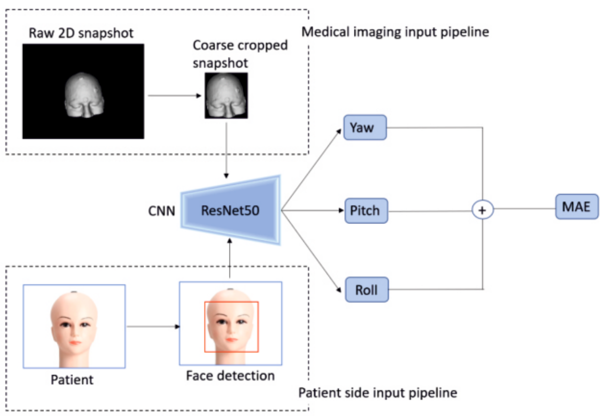


Fig.1: Head pose estimation pipeline for patient and medical imaging data.

**Author for Correspondence: A. Bhardwaj*
E-mail: abhilashb@barc.gov.in

depicts the graphical representation of patient and medical imaging head pose estimation pipeline.

3D reconstruction of MRI data contains only half of the face; hence the use of face detector is restricted. In this case, we have applied a coarse cropping mechanism to crop out the face area from medical imaging data. This cropped part is provided to backbone CNN to predict the yaw, pitch and roll angles.

Evaluation metric

The evaluation work estimates yaw, pitch and roll angles associated with head pose and attempts to minimize mean absolute error(MAE) across all the three angles as given in Equation 1.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \quad (1)$$

Equation 1: Formula to calculate Mean Absolute Error, in our case n=3 (yaw, pitch, roll), y (predicted angle) and x (true angle)

Training

Fig.2 represents validation loss for different training experiments conducted using modified AlexNet architecture for patient's head pose estimation. Here, we can observe the variation of Mean Absolute Error(MAE) with respect to epochs (the number of passes of the entire training dataset through neural network). Fig.2 also compares the variation of MAE for different batch sizes (BS). Fig.3 depicts the training and validation loss achieved during experimentation with ResNet50 architecture for the same problem, where we can observe significant reduction in validation loss compared to AlexNet.

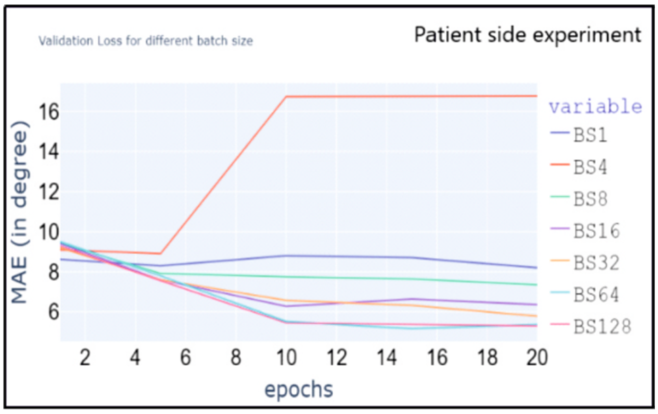


Fig.2: Validation loss-AlexNet.

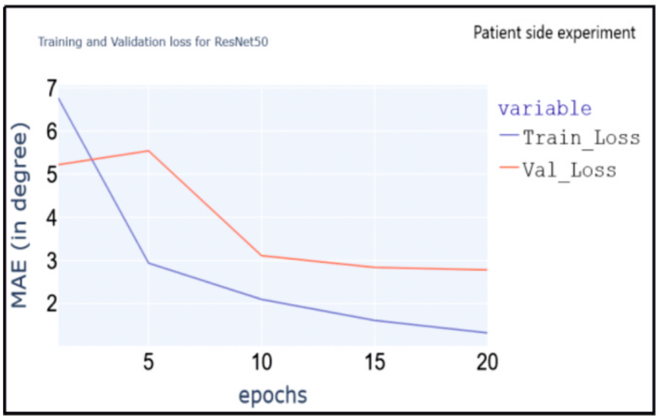


Fig.3: Validation loss - ResNet50.

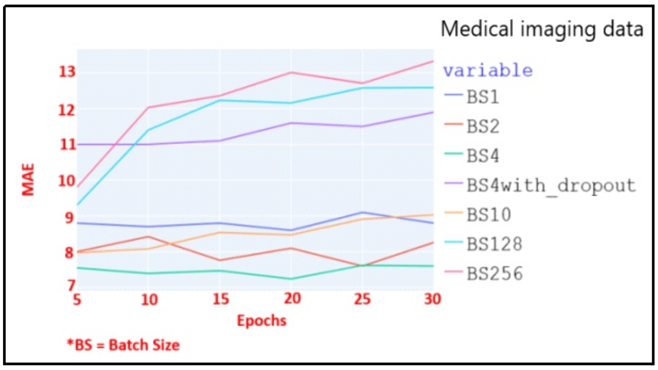


Fig.4: Validation loss-AlexNet.

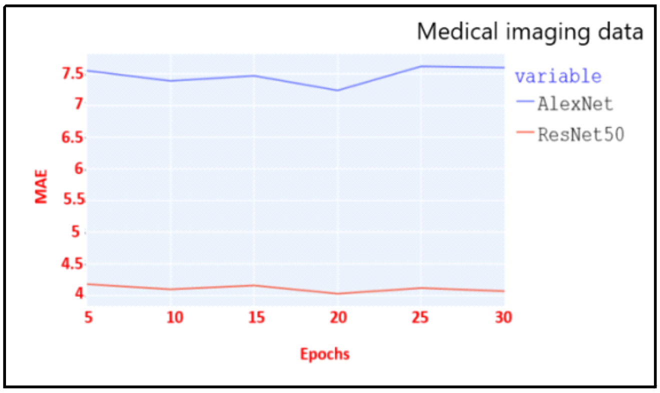


Fig.5: Comparison of AlexNet and ResNet50.

Table 1: Comparison of MAE over BIWI dataset for patient head pose.

Method	MAE (in degree) over complete BIWI dataset
Modified AlexNet trained over BIWI	3.58
Modified AlexNet trained over 300W_LP	7.77
Modified ResNet50 trained over 300W_LP	4.31

Fig.4 represents validation loss for different training experiments using modified AlexNet for medical imaging data. We can observe a significant reduction in validation loss using ResNet50 architecture for medical imaging head pose as depicted in Fig.5.

Results

Table 1 shows the performance of patient head pose estimation models over BIWI dataset, while Fig.6 depicts the predicted yaw, pitch and roll angles for medical imaging data.

Gross Localization of Fiducials

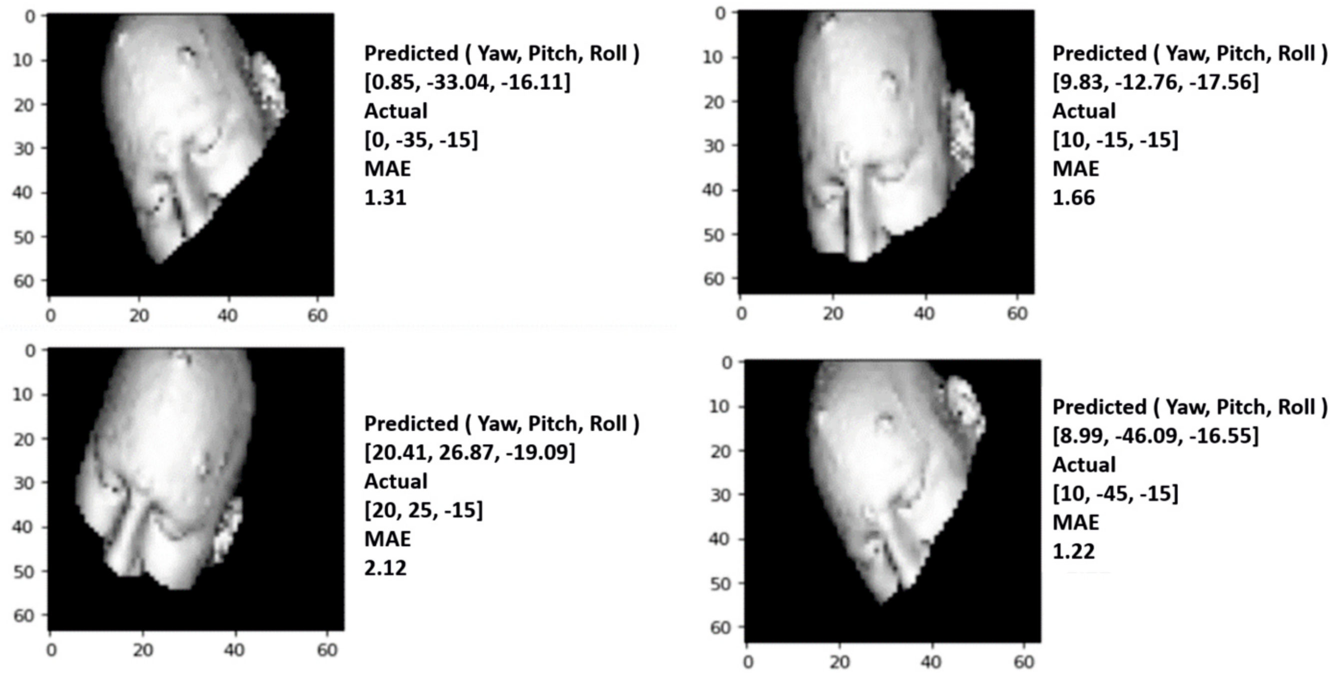
Fiducial localization is a complex and critical process, which requires manual annotation of fiducials over entire MRI volume. Gross localization of fiducials results in a bounded region of interest where fiducials are located. This accelerates the localization process and improves the overall efficiency of surgery. For gross localization of fiducials, deep learning based CNN architectures are used. The following sections will detail the data sets, architecture, training of the model and results.

Dataset

We have anonymized MRI data as received from a Hospital. Each MRI volume has 256x256x84 voxels with a resolution of 1mm x 1mm x 2mm.

Solution architecture

This work takes an axial slice of MRI as input and extracts the boundary (pronounced) of the slice which may contain



***All angles are in degree**

Fig.6: Predictions over 3D head volume (MRI data) with ground truth.

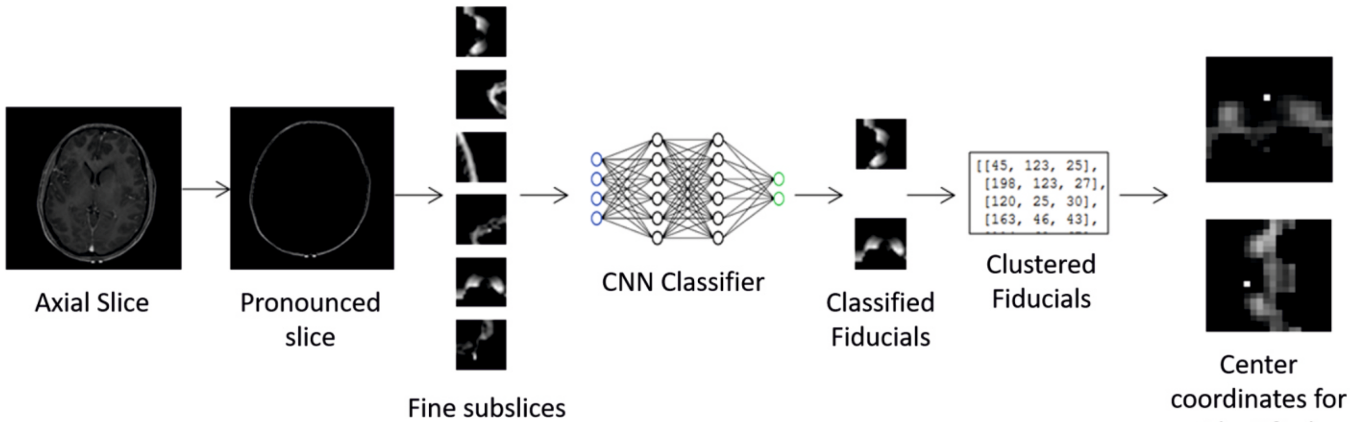


Fig.7: Solution architecture for fiducial localization.

fiducials. The pronounced slice is then divided into small fine sub-slices and passed on to the sub-slice classification network. Fig.7 depicts the solution architecture for gross localization of fiducials.

The sub-slice classification network as shown in Fig. 8 is built by aggregating the standard Deep Learning blocks viz. Convolution-Relu-Pooling (CRP), CR (Convolution-Relu), FCR (Fully Connected Relu) etc. The final sigmoid layer provides probability of whether the input sub-slice contains fiducial or not. We have used DBSCAN[7] clustering algorithm to identify the number of fiducials in an input volume by merging the neighborhood sub-slices(or fiducial candidates). The center of cluster of fiducials is taken as localized fiducial.

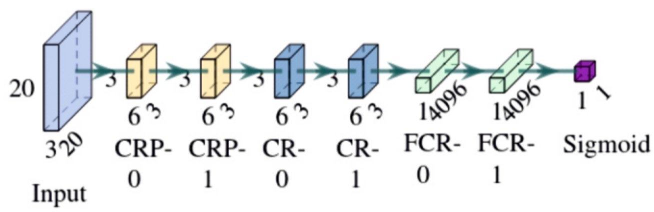


Fig.8: Sub-slice classification network.

Results

This method achieved a fiducial localization error of 4.3 mm, which implies a bounded 4.3mm spherical region of interest where fiducial is located. Now the doctors can refine the localization within the bounded region instead of scanning through the entire MRI volume, which reduces the localization time by 30%. Fig. 9 shows the localized fiducials over 3D head volume.

Conclusion

This article is a step towards head pose estimation for a neurosurgery and it presents a solution to establish a coarse alignment between patient's head and medical imaging data by estimating the head pose. Further, gross localization of fiducials reduces the localization time hence providing an assistance to doctors. The software modules developed here are currently under integration phase with the Neuro navigation system developed by BARC.

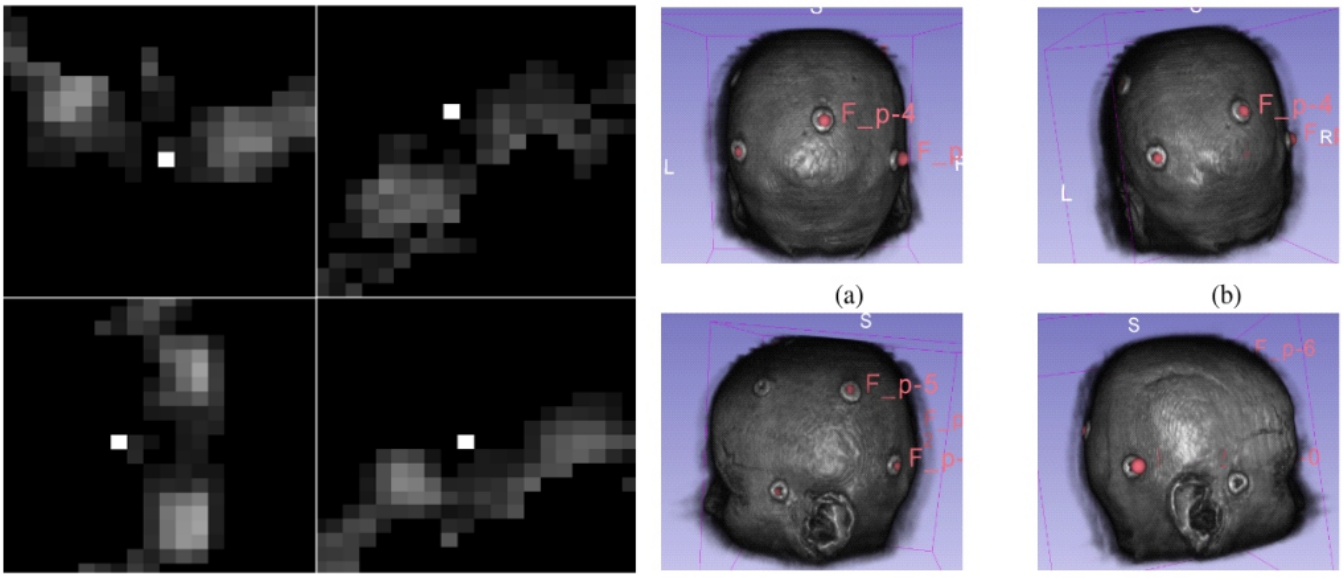


Fig.9: Fiducial localization in MRI sub-slices (left), Fiducial localization over test dataset (right).

References

[1] Gaurav Bhutani, T. A. Dwarakanath, K. D. Lagoo, Aliasgar Moiyadi and Venkata P P K: Neuro-Registration and Navigation Unit for Surgical Manipulation. In: 1st International and 16th National Conference on Machines and Mechanisms (iNaCoMM2013), IIT Roorkee (2013).

[2] Gaur, P., Bhardwaj, A., Venkata, P., Sarode, D. and Mahajan, A.: Assistance technique for Fiducial Localization in MRI images with Subslice Classification Neural Network. 6th International Conference on Intelligent Computing and Control Systems (ICICCS) (2022).

[3] He, K., Zhang, X., Ren, S. and Sun, J.: Deep Residual Learning for Image Recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR).(2016)

[4] BIWI Dataset https://www.kaggle.com/datasets/kmader/biwi-kinect-head-pose-database

[5] 300W_LP Dataset https://www.tensorflow.org/datasets/catalog/the300w_lp

[6] AFLW Dataset https://www.tensorflow.org/datasets/catalog/aflw2k3d

[7] Deng, D. “DBSCAN clustering algorithm based on density,” 2020 7th International Forum on Electrical Engineering and Automation (IFEEA) (2020).