DEPARTMENT OF ELECTRONICS AND
COMMUNICATION ENGINEERING
INDIAN INSTITUTE OF INFORMATION
TECHNOLOGY, DESIGN AND
MANUFACTURING KANCHEEPURAM
CHENNAI - 600127

Synopsis of

COMPUTER-AIDED CATARACT DIAGNOSIS USING FUNDUS RETINAL IMAGES

A Thesis

To be submitted by

TURIMERLA PRATAP

For the award of the degree

of

DOCTOR OF PHILOSOPHY

1 Abstract

Cataract is the cloudiness present in the eye lens due to denaturation of active protein cells. Cataract affects the quality-of-life and thereby troubling the daily routine activities. If a cataract is not diagnosed at an earlier stage, then it may lead to blindness. Early diagnosis and treatment may reduce vision loss and delays the cataract progression. To diagnose a large-screen population, the computer-aided cataract diagnosis (CACD) method using fundus retinal (FR) images is required. Hence, three CACD methods are proposed in this dissertation. Among the three CACD methods proposed, one method aims at improving the diagnostic performance based on accuracy, whereas the other two methods are focused on improving the robustness. In addition to the proposed CACD methods, a noise level estimation (NLE) technique is also proposed in this thesis to ensure robust performance against noisy environment. Initially, a CACD method is proposed to detect various stages of the cataract such as normal, mild, moderate, and severe from the FR images. The proposed method uses the pre-trained deep neural network (DNN) for transfer learning to carry out automatic cataract classification. Then pre-trained DNN model is used for the feature extraction and the extracted features are then applied to a support vector machine classifier. It is observed that the pre-trained DNN based CACD methods are suffering from performance diminution due to the presence of noise in digital fundus retinal images. To develop robust CACD methods, a simple, fast, and accurate NLE method for additive white Gaussian noise (AWGN) is proposed in this thesis. In subsequent to the proposed NLE method, an efficient network selection based robust CACD method under AWGN is proposed. The presented robust CACD method consists of a set of locally- and globally-trained independent support vector networks with features extracted at various noise levels. A suitable network is then selected based on the noise level present in the input image. Eventually, a robust CACD method using FR images is proposed to achieve better diagnostic accuracy. In this method, the features are extracted using combined feature extraction technique. The NLE-based classification is adopted in the classification stage to further enhance diagnostic performance. From the experimental results, it is observed that the proposed CACD methods exhibit superior performance to existing CACD methods under noisy conditions.

2 Objectives

The major objectives of the thesis are stated as follows:

- To develop an accurate computer-aided cataract diagnosis (CACD) method using fundus retinal (FR) images.
- To develop a simple noise level estimation (NLE) technique that is useful to estimate the noise level present in the input retinal image.
- Analysis and development of the robust computer-aided cataract diagnosis systems under noisy environment.
- To develop a robust deep neural network (DNN) based CACD method under practical noise conditions.

3 Existing Gaps Which Were Bridged

The existing gaps which were bridged in this thesis are listed as follows:

(a) The transfer learning from DNN is a more efficient classification method, and it is having less over-fitting issues as compared to end-to-end DNN constructed from scratch (Zhang *et al.*, 2017). A transfer learning-based CACD method has been exploited using FR images. The proposed CACD method outperforms the previous hand-crafted feature extraction methods (Yang *et al.*, 2013; Zheng *et al.*, 2014; Guo *et al.*, 2015; Yang *et al.*, 2016; Xiong *et al.*, 2017; Zhang *et al.*, 2017; Qiao *et al.*, 2017; Pratap and Kokil, 2019*a*). The presented method obtained 92.91% of classification accuracy which is better than the existing methods.

(b) It is observed that the existing CACD methods are suffering from performance diminution due to the presence of noise in digital FR images (Dodge and Karam, 2016). The performance of existing pre-trained DNN based CACD methods (Li *et al.*, 2018; Pratap and Kokil, 2019*b*) are performing poorly when tested with noisy input fundus retinal images. To overcome this drawback, robust CACD methods are proposed in this thesis. The presented methods improved the performance of CACD under noisy environment.

(c) The existing NLE methods (Donoho and Johnstone, 1994; Immerkaer, 1996; Yang and Tai, 2010; Liu *et al.*, 2013; Liu and Lin, 2012; Pyatykh *et al.*, 2012; Rakhshanfar and Amer, 2016; Xu *et al.*, 2017) are having poor performance to estimate a wide range of noise levels. Hence, a simple NLE method is proposed in this report. The proposed method is a direct approach and requires only Gaussian filters to estimate noise levels. This method is always free from the tedious segmentation process used for the extraction of super-pixels and homogeneous regions in a noisy image.

4 Most Important Contributions

4.1 CACD using deep transfer learning

In handcrafted feature engineering, the features are extracted manually by proper observations made from input feature space. In practice, the extraction of features using manual feature extraction techniques involves subject experts. Hence, this work focuses on the automatic feature extraction technique based CACD method. The framework of transfer learning based cataract detection method is shown in Figure 1. In this work, the features are extracted automatically from the pre-trained deep neural network. The extracted features are then given to support vector machine classifier for accurate cataract grading. **This proposed method makes Chapter 3 of the thesis.**

The proposed CACD method consists of an input image quality selection (IQS) module, pre-processing, feature extraction, and classification stages as shown in Figure 1. The input IQS module is used to filter out poor-quality fundus retinal images and allow only good-quality fundus retinal images to the input section. The poor-quality fundus retinal images are ignored in the analysis of the proposed CACD method for effectiveness. The G-filter is used as a pre-processing technique. The presented transfer learning approach achieved better performance as compared to the existing state-of-art handcrafted feature based CACD methods as shown in Table 1.

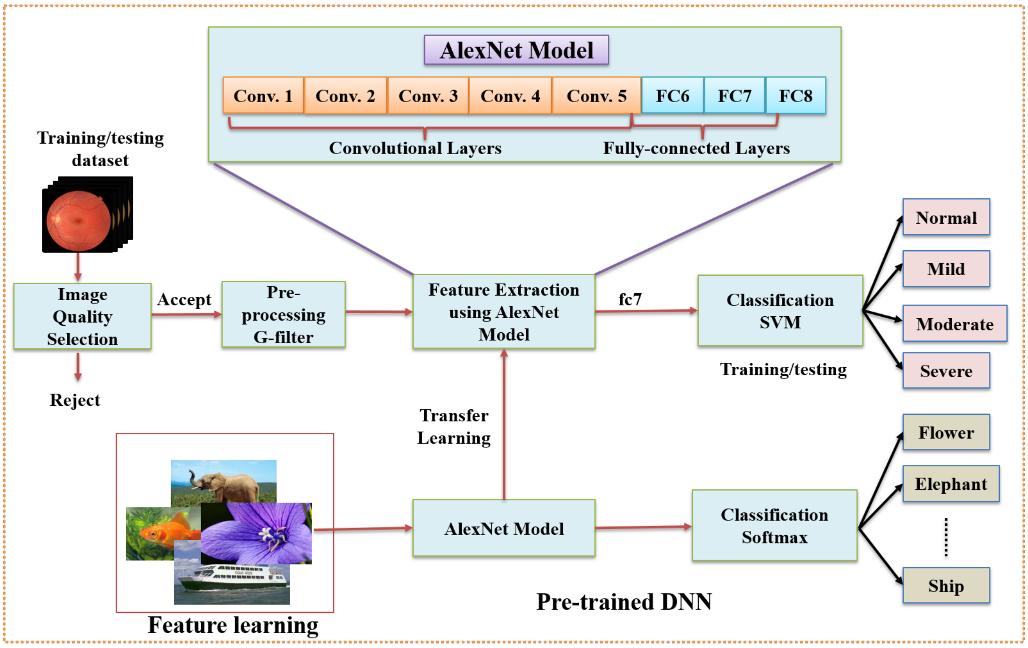

Figure 1: Transfer learning based computer-aided cataract diagnosis.

4.2 NLE for images using linear scale-space features

The performance of the proposed CACD method is dependent on noise level present in the input images. So NLE is important to achieve robust performance under noisy conditions. Hence, an NLE method is proposed and presented in this section. The framework of proposed NLE method is shown in Figure 2. The basic idea behind the proposed NLE is to train an independent set of linear regression (LR) models locally with respect to specific noise level, which is the nearest value to the actual noise level. Since the LR model trained on a narrow range of noise levels around actual noise level performs better as compared to a single LR model trained on a wide range of noise levels. The set of LR models trained individually and independently, called as LTLR for noise levels of range [0 100] are used to estimate noise level. The proposed feature extraction technique is developed via scale-space representation of the images for calculation of feature vector. The Gaussian scale representation is equivalent to low-pass filtering of a given noisy image with a Gaussian kernel. In this work, the noisy image is represented at three different scales using 3×3 Gaussian kernel with varying variances $\sigma' = 0.5, 1,$ and 5. The scale-space framework encompasses a theory for Gaussian derivative operators which can be used as a basis for expressing a large class of visual operations for computerized methods that process visual information. **The detailed description of this work is presented in Chapter 4 of the thesis.**

Table 1: Comparison of results with existing methods.

S.No	Reference	Feature Extraction	Classification	Two class accuracy (%)	Four class accuracy (%)
1	(Yang *et al.*, 2013)	Luminance	BPNN	85.96	82.29
		Gray co-occurrence matrix			
		Gray gradient co-occurance matrix			
2	(Zheng *et al.*, 2014)	DFT features	AdaBoost	95.22	81.52
3	(Guo *et al.*, 2015)	DWT	LDA	90.9	77.1
		Sketch		86.1	74
		DWT+Sketch		89.3	73.8
4	(Yang *et al.*, 2016)	DWT	SVM	91.6	82.5
		Sketch	SVM	87.9	78.2
		Texture	SVM	90.4	81.9
		DWT+Sketch+Texture	SVM	90.5	83.2
		DWT	BPNN	91.9	81.3
		Sketch	BPNN	87.8	75.7
		Texture	BPNN	90.4	80.4
		DWT+Sketch+Texture	BPNN	89.9	82.9
		DWT+Sketch+Texture	Ensemble learning	93.2	84.5
5	(Xiong *et al.*, 2017)	Statistical features	Decision tree	88.4	83.8
6	(Qiao *et al.*, 2017)	Genetic algorithm	SVM	95.33	87.52
7	(Zhang *et al.*, 2017)	DNN		93.52	86.69
8	(Pratap and Kokil, 2019*a*)	SVD	SVM	97.78	-
9	**Proposed**	**Pre-trained CNN**	**SVM**	**100**	**92.91**

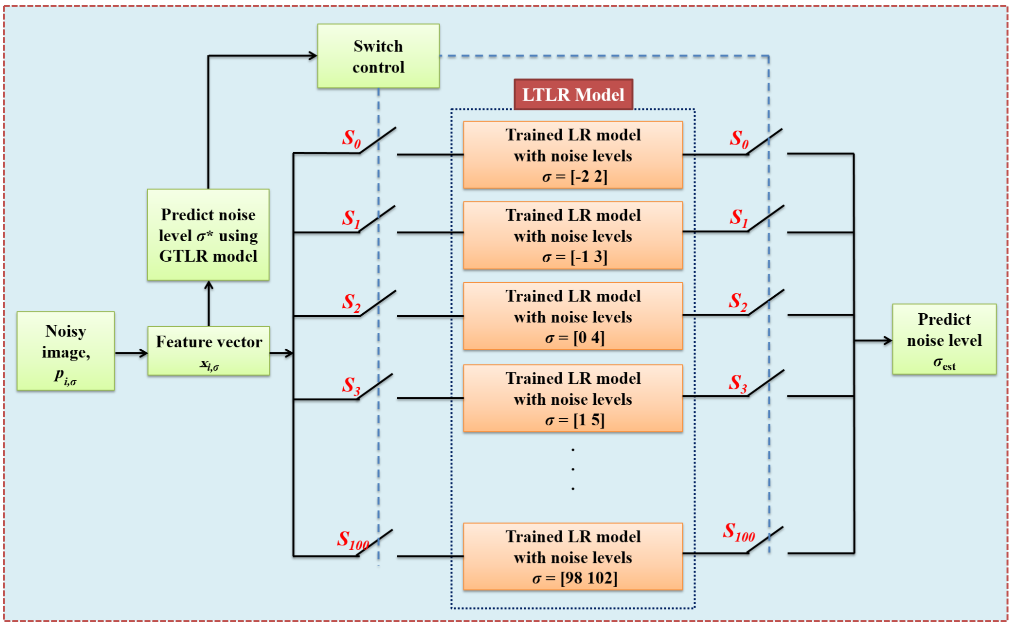

Figure 2: Framework of proposed NLE method.

4.3 Efficient network selection for CACD under noisy conditions

The proposed transfer learning based CACD method proved as an accurate method when compared with existing CACD methods (Yang *et al.*, 2013; Zheng *et al.*, 2014; Guo *et al.*, 2015; Yang *et al.*, 2016; Xiong *et al.*, 2017; Zhang *et al.*, 2017; Qiao *et al.*, 2017; Pratap and Kokil, 2019*a*). The use of different DNNs in CACD methods has been increased in recent years. It is observed that the performance of DNN based CACD methods is good when training and testing are done with good-quality fundus retinal image datasets. However, the performance of pre-trained DNN based CACD methods gets worse when the network trained with a good-quality fundus retinal image dataset is then tested with a noisy fundus retinal dataset at different noise levels (σ). It is perceived that pre-trained DNN based CACD methods that are trained with good-quality fundus retinal images are performing poorly when tested with noisy input fundus retinal images. So the presented CACD method considered this problem as a serious concern and then provided a solution by combining locally-trained and globally-trained support vector networks in order to achieve better performance under noisy conditions. **This work is presented in Chapter 5 of the thesis.**

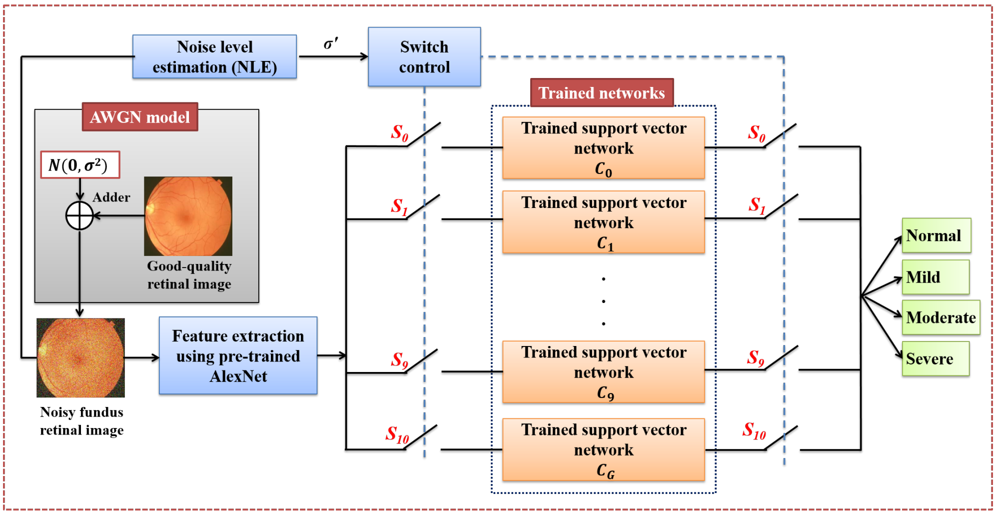

Figure 3: Framework of the proposed CACD method based on selection of efficient network.

The presented CACD method consists of locally- and globally-trained independent support vector network classifiers in which a suitable network is selected based on the noise level present in an input fundus retinal image. The framework of the proposed CACD method is shown in Figure 3. The ability of cataract detection is depending on the noise level of the input fundus retinal image and the selection of a trained support vector network for the corresponding input image. So the selection of an efficient network comprises two stages as.

(a) Noise level estimation (NLE) from input fundus retinal image.

(b) Mapping the input fundus retinal image to the desired network for cataract diagnosis based on NLE from input fundus retinal image as shown in Figure 3.

The noise estimation from noisy input fundus retinal image is made possible with the use of existing NLE techniques. The noise levels used in this work are integers so it is required to round off the output estimated noise levels (σ') in practice. For sake of convenience, we assumed the estimation of the noise level in ideal conditions ($\sigma' = \sigma$). The switch control module shown in Figure 3 is used to control the switches S_0, S_1, ... , S_{10}. The operation of k^{th} switch S_k under switch control is given as

$$S_k = \begin{cases} \textit{Close to select SVM } C_k, & \textit{for } k < 10 \\ \textit{Close to select SVM } C_G, & \textit{for } 10 \leq k \leq 55 \\ \textit{Open}, & \textit{elsewhere}. \end{cases} \tag{1}$$

4.4 CACD method under practical noise conditions

The quality level present in the training and testing must inevitably be the same to achieve better accuracy for distorted input images. At the same time, the fine-tuning of a network with noisy images increases the robustness against noisy environment (Rodner *et al.*, 2016). So the combination of fine-tuned DNN in association with quality level matching further improves the robustness than individual method. Hence, the proposed CACD method consists of fine-tuning based combined feature extraction (CFE) technique then followed by NLE-based classification. The framework of proposed method is shown in Figure 4. **The presented method makes Chapter 6 of the thesis.**

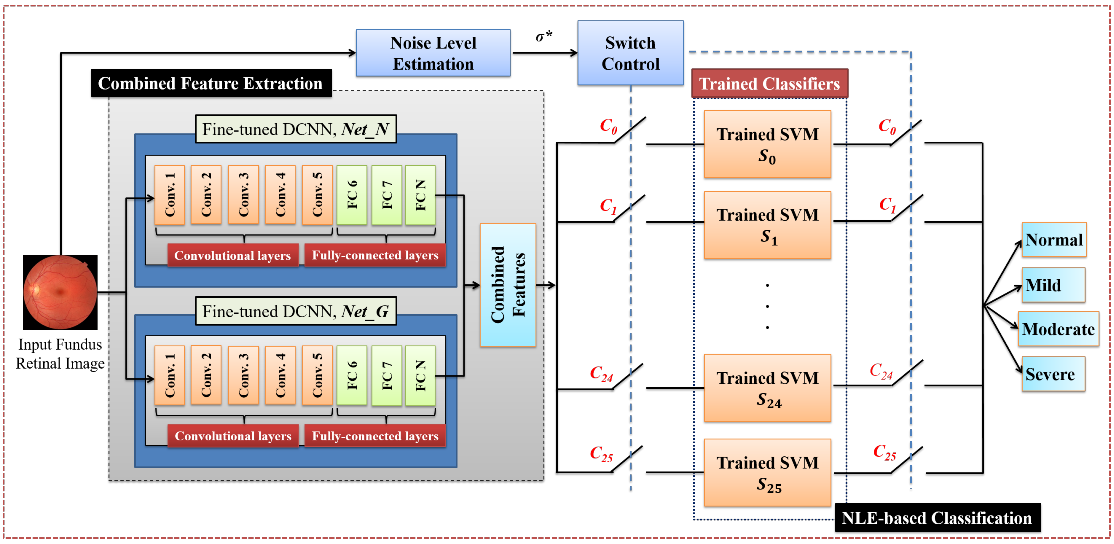

Figure 4: The CACD method under practical noise conditions.

The performance of networks Net_N and Net_G in the proposed CACD method are evaluated in terms of accuracy and presented in Table 2. The performance of the fine-tuned network Net_G in association with support vector machine (SVM) classifier is good at $\sigma_a = 0$ and poor as noise level σ_a increases. The performance of the network Net_N in association with SVM classifier is good and consistent but exhibiting lower performance at $\sigma_a = 0$ than actual performance. Hence, the features that are extracted from the fine-tuned networks Net_N and Net_G are combined. The performance of CFE technique followed by SVM classifier and proposed NLE-based classification are presented in Table 2.

Table 2: The performance of proposed CACD method in terms of accuracy.

Method	Noise levels $(\sigma_m = 0, \sigma_a = \{0, 5, 10, 15, 20, 25\})$					
	$\sigma_a = 0$	$\sigma_a = 5$	$\sigma_a = 10$	$\sigma_a = 15$	$\sigma_a = 20$	$\sigma_a = 25$
Net_G+SVM	93.23±0.00	89.92±0.46	71.86±0.90	46.98±0.71	41.99±0.84	40.10±0.72
Net_N+SVM	91.67±0.00	90.01±0.59	88.52±0.57	87.19±0.70	85.25±0.78	83.79±0.78
CFE+SVM	**93.49±0.00**	**91.03±0.47**	88.15±0.63	84.67±0.82	82.41±0.98	77.49±1.18
Proposed	**93.49±0.00**	90.81±0.35	**88.98±0.69**	**87.67±0.61**	**85.36±0.77**	**84.27±0.88**

The proposed CACD method is compared with existing methods (Zhang *et al.*, 2017; Li *et al.*, 2018; Pratap and Kokil, 2019*b*, 2020) and presented in Table 3. In result comparison, the existing CACD methods developed based on different variations of the DNN models such as AlexNet trained from scratch (Zhang *et al.*, 2017), pre-trained AlexNet using transfer learning (Pratap and Kokil, 2019*b*), ResNet-50 using fine-tuning (Li *et al.*, 2018), and AlexNet with globally trained SVM (Pratap and Kokil, 2020) are considered. From the Table 3, it is clear that the proposed CACD method is superior when compared with existing CACD methods for noise levels ranging from 0 to 25.

Table 3: Comparison of the proposed CACD method with existing methods.

Method	Noise levels $(\sigma_m = 0, \sigma_a = \{0, 5, 10, 15, 20, 25\})$					
	$\sigma_a = 0$	$\sigma_a = 5$	$\sigma_a = 10$	$\sigma_a = 15$	$\sigma_a = 20$	$\sigma_a = 25$
(Zhang *et al.*, 2017)	88.54±0.00	61.17±0.22	46.42±0.12	41.52±0.16	32.75±0.36	26.41±0.32
(Li *et al.*, 2018)	86.20±0.00	62.93±1.08	32.57±0.68	24.99±0.06	25.00±0.00	25.00±0.00
(Pratap and Kokil, 2019*b*)	92.97±0.00	87.46±0.66	72.14±0.80	58.19±0.78	39.54±1.01	28.78±0.70
(Pratap and Kokil, 2020)	88.28±0.00	88.92±0.70	85.51±0.77	81.28±1.08	78.14±1.44	75.86±1.39
Proposed	**93.49±0.00**	**90.81±0.35**	**88.98±0.69**	**87.67±0.61**	**85.36±0.77**	**84.27±0.88**

The proposed CACD method is also applicable to practical noise such as Poisson-Gaussian. The applicability of the proposed CACD method is evaluated with dataset $\mathbf{U}_{\sigma_m,\sigma_a}$ by varying noise levels σ_m and σ_a and then presented in Table 4. The noise level for multiplicative and Poisson-Gaussian noises is estimated from the noise-alone image obtained after subtracting the original image from the corresponding noisy image (Laligant *et al.*, 2013). The standard deviation of noise-alone image (σ^*) is then given to switch control in the proposed CACD method for proper operation.

Table 4: The evaluation of proposed CACD method with practical noise (Poisson-Gaussian).

Method	Noise levels $(\sigma_m = \{0, 0.5, 1, 1.5\}, \sigma_a = 5)$			
	$\sigma_m = 0$	$\sigma_m = 0.5$	$\sigma_m = 1$	$\sigma_m = 1.5$
Net_G+SVM	89.88±0.59	81.99±0.38	66.41±0.87	45.13±0.66
Net_N+SVM	90.25±0.37	89.52±0.57	**87.53±0.89**	86.00±1.02
CFE+SVM	**91.05±0.42**	89.30±0.48	86.18±0.45	83.91±0.84
Proposed	91.02±0.39	**90.10±0.62**	87.47±0.57	**86.12±0.77**

5 Conclusions

The performance diminution has been observed in CACD methods for noisy input FR images. The CACD methods using efficient selection of support vector machines and CFE technique were proposed. The proposed robust CACD methods corrected the performance diminution against noisy FR images. The analysis has been carried out with AWGN and signal-dependent practical noises. The performance of presented CACD methods were analysed and then compared with existing methods. From the experimental results, the proposed CACD methods are found to be efficient and robust for good-quality and noisy input FR images. In future, the image degradations such as blur, contrast and JPEG artifacts need to be considered in the design of CACD methods. The generalized IoT based robust CACD method against variety of image distortions is required and can be considered as future scope.

6 Organization of the Thesis

The proposed outline of the thesis is as follows:

(a) Chapter 1: Introduction

(b) Chapter 2: Literature survey

(c) Chapter 3: CACD using deep transfer learning

(d) Chapter 4: Noise level estimation for images using linear scale-space features

(e) Chapter 5: Efficient selection of SVM for CACD under noisy environment

(f) Chapter 6: DNN based robust CACD method

(g) Chapter 7: Conclusion and future scope

7 List of Publications

I. REFEREED JOURNALS BASED ON THE THESIS

1. **Turimerla Pratap** and Priyanka Kokil. "Computer-aided diagnosis of cataract using deep transfer learning", *Biomedical Signal Processing and Control*, 53, 101533, (2019). [I.F. 3.88]
2. **Turimerla Pratap** and Priyanka Kokil. "Efficient network selection for computer-aided cataract diagnosis under noisy environment", *Computer Methods and Programs in Biomedicine*, 200, 105927, (2021). [I.F 5.428]
3. **Turimerla Pratap** and Priyanka Kokil. "Deep neural network based robust computer-aided cataract diagnosis system using fundus retinal images", *Biomedical Signal Processing and Control*, 70, 102985, (2021). [I.F. 3.88]

4. Priyanka Kokil and **Turimerla Pratap**. "Additive white Gaussian noise level estimation for natural images using linear scale-space features", *Circuits, Systems, and Signal Processing*, 40.1, 353-374, (2021). [I.F. 2.225]

II. PRESENTATIONS/PUBLICATIONS IN CONFERENCES BASED ON THE THESIS

1. **Turimerla Pratap** and Priyanka Kokil. "Automatic cataract detection in fundus retinal images using singular value decomposition", *IEEE International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, 373 - 377, (2019).
2. **Turimerla Pratap** and Priyanka Kokil. "Correcting automatic cataract diagnosis systems against noisy/blur environment", *IEEE National Conference on Communications (NCC)*, 1 - 6, (2020).

III. REFEREED JOURNALS (Others)

1. Sudharson S., **Turimerla Pratap**, and Priyanka Kokil. "Noise level estimation for effective blind despeckling of medical ultrasound images", *Biomedical Signal Processing and Control*, 68, 102744, (2021). [I.F. 3.88]

IV. PRESENTATIONS/PUBLICATIONS IN CONFERENCES (Others)

1. **Turimerla Pratap** and Priyanka Kokil. "Approximate optimization of gabor filter parameters in application to blood vessel segmentation in retinal images", *IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, 1 - 5, (2019).

References

1. **Dodge, S.** and **L. Karam** (2016). Understanding how image quality affects deep neural networks. *In IEEE International Conference on Quality of Multimedia Experience*. 1–6.

2. **Donoho, D. L.** and **J. M. Johnstone** (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, **81**(3), 425–455.

3. **Guo, L.**, **J.-J. Yang**, **L. Peng**, **J. Li**, and **Q. Liang** (2015). A computer-aided healthcare system for cataract classification and grading based on fundus image analysis. *Computers in Industry.*, **69**, 72–80.

4. **Immerkaer, J.** (1996). Fast noise variance estimation. *Computer Vision and Image Understanding*, **64**(2), 300–302.

5. **Laligant, O.**, **F. Truchetet**, and **E. Fauvet** (2013). Noise estimation from digital step-model signal. *IEEE Transactions on Image Processing*, **22**(12), 5158–5167.

6. **Li, J.**, **X. Xu**, **Y. Guan**, **A. Imran**, **B. Liu**, **L. Zhang**, **J.-J. Yang**, **Q. Wang**, and **L. Xie** (2018). Automatic cataract diagnosis by image-based interpretability. *In IEEE International Conference on Systems, Man, and Cybernetics*. 3964–3969.

7. **Liu, W.** and **W. Lin** (2012). Additive white gaussian noise level estimation in SVD domain for images. *IEEE Transactions on Image processing*, **22**(3), 872–883.

8. **Liu, X.**, **M. Tanaka**, and **M. Okutomi** (2013). Single-image noise level estimation for blind denoising. *IEEE transactions on image processing*, **22**(12), 5226–5237.

9. **Pratap, T.** and **P. Kokil** (2019*a*). Automatic cataract detection in fundus retinal images using singular value decomposition. *In IEEE International Conference on Wireless Communications Signal Processing and Networking*. IEEE, 373–377.

10. **Pratap, T.** and **P. Kokil** (2019*b*). Computer-aided diagnosis of cataract using deep transfer learning. *Biomedical Signal Processing and Control*, **53**, 101533.

11. **Pratap, T.** and **P. Kokil** (2020). Correcting automatic cataract diagnosis systems against noisy/blur environment. *In IEEE National Conference on Communications*. 1–6.

12. **Pyatykh, S.**, **J. Hesser**, and **L. Zheng** (2012). Image noise level estimation by principal component analysis. *IEEE Transactions on Image Processing*, **22**(2), 687–699.

13. **Qiao, Z.**, **Q. Zhang**, **Y. Dong**, and **J.-J. Yang** (2017). Application of SVM based on genetic algorithm in classification of cataract fundus images. *In IEEE International Conference on Imaging Systems and Techniques*. 1–5.

14. **Rakhshanfar, M.** and **M. A. Amer** (2016). Estimation of gaussian, poissonian–gaussian, and processed visual noise and its level function. *IEEE Transactions on Image Processing*, **25**(9), 4172–4185.

15. **Rodner, E.**, **M. Simon**, **R. B. Fisher**, and **J. Denzler** (2016). Fine-grained recognition in the noisy wild: Sensitivity analysis of convolutional neural networks approaches. *arXiv preprint arXiv:1610.06756*.

16. **Xiong, L.**, **H. Li**, and **L. Xu** (2017). An approach to evaluate blurriness in retinal images with vitreous opacity for cataract diagnosis. *Journal of Healthcare Engineering. Volume 2017, Article ID: 5645498*.

17. **Xu, S.**, **X. Zeng**, **Y. Jiang**, and **Y. Tang** (2017). A multiple image-based noise level estimation algorithm. *IEEE Signal Processing Letters*, **24**(11), 1701–1705.

18. **Yang, J.-J.**, **J. Li**, **R. Shen**, **Y. Zeng**, **J. He**, **J. Bi**, **Y. Li**, **Q. Zhang**, **L. Peng**, and **Q. Wang** (2016). Exploiting ensemble learning for automatic cataract detection and grading. *Computer Methods and Programs in Biomedicine.*, **124**, 45–57.

19. **Yang, M.**, **J.-J. Yang**, **Q. Zhang**, **Y. Niu**, and **J. Li** (2013). Classification of retinal image for automatic cataract detection. *In IEEE International Conference on e-Health Networking, Applications & Services*. 674–679.

20. **Yang, S.-M.** and **S.-C. Tai** (2010). Fast and reliable image-noise estimation using a hybrid approach. *Journal of Electronic Imaging*, **19**(3), 033007.

21. **Zhang, L.**, **J. Li**, **H. Han**, **B. Liu**, **J. Yang**, and **Q. Wang** (2017). Automatic cataract detection and grading using deep convolutional neural network. *In IEEE International Conference on Networking, Sensing and Control*. 60–65.

22. **Zheng, J.**, **L. Guo**, **L. Peng**, **J. Li**, **J. Yang**, and **Q. Liang** (2014). Fundus image based cataract classification. *In IEEE International Conference on Imaging Systems and Techniques*. 90–94.