Abstract: More than 42 Cr new diabetes Patients added worldwide as per the World Health Association Annual Report Statistics [3, 7]. The World Health Organization (WHO) reports that there is measurable hike in the number of individual Diabetes cases in the various regions and sectors of WHO Survey [9]. Because of the high level of stress, irrespective of the Gender and income, the Death Toll increasing every year. In this paper, hypothetical analysis-based Survey done of diabetes mellitus for early prediction and Automatic Detection of Exudates for Diabetic Retinopathy [8, 17]. The Hypothetical analysis results indicate the severances of the issue and significant importance of the need for early prediction and Automatic Detection [13]. With hypothetical analysis across various models we proposed to provide a vision into various machine learning models and its prognostic precision in relations of the recital, accuracy improvement from 2+% to 12+%.

Keywords: Exudates, Diabetic Retinopathy (DR)

I. INTRODUCTION

The different sorts of diabetes in patients create different sorts of retinopathy get creates after long term several categories of the diabetes [17].Diabetic retinopathy caused by diabetes for long duration and end up with the blindness across the young age to the people who have reached their maturity [18]. Diabetic retinopathy is categorized by the development of Retinal Microneurysms, Hemorrhages and Exudates. Because of its occurrence and clinical importance, the investigators have attempted to improve its analysis and cure by emerging procedures to complete retinal image analysis, fundus image improvement and observing [5, 16, 19]. Different types of Diabetic Retinopathy [18, 19] are as follows: Mild Non-Proliferative Retinopathy, Moderate Non-Proliferative Retinopathy, Severe Non-Proliferative Retinopathy and Proliferative Retinopathy.

II. OBJECTIVE OF THE PROPOSED METHODOLOGY

The major Objective of the Proposed Methodology and Hypothetical Research Survey Analysis for recognition of Exudates in diabetic retinopathy images for diverse categories of image considerations [17, 18].

1. To habitually sense the subsequent normal structures in image to procedure the pathology finding
   (i) Automatic finding of optic circle boundary
   (ii) Automatic finding of retinal blood vessels
2. To consequently distinguish lesion, i.e., exudates within the retinal image for the initial discovery of diabetic retinopathy [6, 12].
3. To create a programmed retinal examination framework to classify the harshness of the illness [14, 15].

Fig 1: Data Flow Diagram (DFD) of the Methodology for Analysis

1. Stage 1 gives the data about the information database (discloses method embraced to gather set of eye pictures).
2. In the next stage pre-processing done to remove noise.
3. Further, the pre-processed stage is segmented using simple linear iterative clustering algorithm super-pixel segmentation is done (stage 3).
4. Stage 4 gives information about optic disc elimination using key point extraction and template matching.
5. The normal feature extraction of the diabetic retinopathy is decided within the stage-5 [17, 18].
6. In block-6, abnormal feature extraction using super pixel multivariable classification algorithm.
7. Finally, the outcome is acquired and will be introduced in the stage 7, which finishes up the adequacy of system created by us.
III. RESULTS AND DISCUSSIONS

For Hypothetical Research Survey Analysis purpose, the data was assimilated from freely obtainable quantities, i.e. Kaggle and Messidor.

A. Kaggle dataset
- 35126 fundus images, marked for 5 class documentation
- No DR, Mild DR, Moderate DR, Extreme DR, Proliferative DR
- Contains a huge number of noisy and often images which are misannotated. The underdone Kaggle data more carefully replicates real-world situation, where images are taken under the unlike circumstances, thus subsequent in numerous superiority levels.

B. Messidor dataset
- 1200 fundus images marked for four class identification [19].
- In spite of its comparatively minor scale, is considered a high-fidelity source with consistent labeling.
- These datasets comprise color photographs of right & left eyes.
- The images proportions vary between low-100s to low-1000s.
- The eminence of data varies expressively among the datasets.

Fig 2: Precise statistics of images for each dataset/class
The experiment based on the possible eye lesions detection notwithstanding the observed noisiness in application dataset. Fig. 2 validates the relative data distribution for Messidor and Kaggle datasets between the individual DR classes.

Table I: Precise Data Set of images for each dataset/class

<table>
<thead>
<tr>
<th>Categories</th>
<th>Mild DR</th>
<th>Severe DR</th>
<th>Pro DR</th>
<th>No DR</th>
<th>Moderate DR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Messidor</td>
<td>20</td>
<td>60</td>
<td>10</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Kaggle</td>
<td>2000</td>
<td>100</td>
<td>500</td>
<td>28000</td>
<td>6000</td>
</tr>
</tbody>
</table>

In table 1 the precise statistics of images for each dataset/class displayed based on the references of the Data available publicly.

Table-II: Data Set of Messidor dataset Vs Kaggle (Raw and Augmented)

<table>
<thead>
<tr>
<th>Categories</th>
<th>Messidor (Raw)</th>
<th>Kaggle (augmented)</th>
<th>Messidor + Kaggle (Raw)</th>
<th>Messidor + Kaggle (augmented)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No DR</td>
<td>500</td>
<td>27000</td>
<td>51000</td>
<td>27000</td>
</tr>
<tr>
<td>Mild DR</td>
<td>20</td>
<td>3000</td>
<td>56000</td>
<td>3000</td>
</tr>
</tbody>
</table>

Associated to the Messidor dataset, the Kaggle dataset contains superior amount of low fidelity data. The images were taken with diverse fundus cameras, subsequent in numerous superiority stages as shown in the Fig. 3 of Messidor dataset Vs. Kaggle based on the dataset Statistics (Raw and Augmented) in Table 2 [5, 16].

Fig 3: Messidor dataset Vs Kaggle Statistics (Raw and Augmented)
Comparatively noisy eccentric of images is experiential through their blurriness, under/over-exposure, occurrence of unconnected relics, and so on. The raw arrangement of Kaggle dataset prudently reproduces the nature of DR detection in real-life settings, where considerable erraticism in data superiority is pragmatic among the organizations.

Figure 4 validates the assorted persons (gender and wage) start with youngsters in late 30s and up to Senior Citizens, level of passing as a result of hypertension.

Table 3: Dataset of % of Diabetic cases on financial status

<table>
<thead>
<tr>
<th>Categories</th>
<th>Poor</th>
<th>Middle Class</th>
<th>Rich</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>72</td>
<td>54</td>
<td>56</td>
</tr>
<tr>
<td>Women</td>
<td>52</td>
<td>45</td>
<td>21</td>
</tr>
</tbody>
</table>

Fig 4: % of Diabetes cases statistics on financial status
As the stress level is high now a days, regardless of the Gender and income, the Diabetes mellitus is long-lasting, a ceaseless illness where it instigated for the great sugar level in the cardiovascular structure. It is affects since of the unacceptable way of life of the pancreatic beta cells. Assorted bits of the physical make-up which incorporates pancreas glitch, threat of heart ailments, hypertension, pancreatic subjects, kidney disappointments, pancreatic subjects, nerve harm, foot matters, ketoacidosis, visual upsetting impacts, and other eye matters.
cascades and glaucoma and so on obstructed because of this. Around diverse determinations behind cause like a lifestyle of a man, the absenteeism of movement, nourishment tendencies, robustness, smoking, elevated cholesterol, hypertension and so forth which fundamentally increment the threat of treating diabetes. It influences a broad assortment of ages, tallying youths to grown-up and matured individuals.

### Table 4: A1C % Data Set

<table>
<thead>
<tr>
<th>Categories</th>
<th>Diabetes</th>
<th>Prediabetes</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1C %</td>
<td>6.7</td>
<td>126</td>
<td>300</td>
</tr>
</tbody>
</table>

Data Set Table 4 shows statistics based on the A1C % with respect to all 3 categories Diabetes, Prediabetes and Normal.

**Fig 5: A1C % Statistics**

Figure 5 shows the statistics based on the A1C % with respect to the all three categories Diabetes, Prediabetes and Normal.

### Table 5: Fasting Plasma Glucose Data Set

<table>
<thead>
<tr>
<th>Categories</th>
<th>Diabetes</th>
<th>Prediabetes</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fasting Plasma</td>
<td>6</td>
<td>112</td>
<td>170</td>
</tr>
<tr>
<td>Glucose</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data Set from Table 5 shows the statistics based on the Fasting Plasma Glucose with respect to the all three categories Diabetes, Prediabetes and Normal.

**Table 6: Oral Glucose Tolerance Test**

<table>
<thead>
<tr>
<th>Categories</th>
<th>Diabetes</th>
<th>Prediabetes</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral Glucose</td>
<td>5</td>
<td>99</td>
<td>139</td>
</tr>
<tr>
<td>Tolerance Test</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data Set from Table 6 shows the statistics based on the Oral Glucose Tolerance Test with respect to the all three categories Diabetes, Prediabetes and Normal. [3, 7]

### Table 7: Various Models

<table>
<thead>
<tr>
<th>Models</th>
<th>Accuracy</th>
<th>Improved Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB</td>
<td>73.5</td>
<td>82.5</td>
</tr>
<tr>
<td>Random Forest</td>
<td>75.5</td>
<td>79</td>
</tr>
<tr>
<td>SVM</td>
<td>77.5</td>
<td>80</td>
</tr>
<tr>
<td>KNN</td>
<td>71.6</td>
<td>75.6</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>74</td>
<td>75.5</td>
</tr>
</tbody>
</table>

Table 7 shows the analysis to give a vision into various AI models and its prognostic precision in relations of the presentation. There is an accuracy improvement across all the models, for the prophetic task will become quicker.

**Fig 6: Fasting Plasma Glucose Statistics**

Figure 6 shows the statistics based on the Fasting Plasma Glucose with respect to the all three categories Diabetes, Prediabetes and Normal. [1]

**Fig 7: Oral Glucose Tolerance Test Statistics**

Figure 7 shows the statistics based on the Oral Glucose Tolerance Test with respect to the all three categories Diabetes, Prediabetes and Normal. [8, 9]

**Fig 8: Hypothetical Analysis of Improvement of an Accuracy reference**

The evaluation of the accuracy of the various groupings is shown in Fig. 8. Recital, accuracy improvement from 2.05% to 12.4% across various models.
IV. CONCLUSION

Several subjects with respect to the diabetic retinopathy were considered in better complexity & the research problem was expressed. In this work, the different methods for DR diagnosis were discussed. It is constructed on studying texture insight competences in fundus images to distinguish vigorous patients from DR images [19]. By the former discovery of diabetic retinopathy using state of art of image technologies will have several applications based on the hypothetical analysis survey in this paper. The strategy talked about where less human correspondence offering increment to amazingly sterile procedure and making the framework recognizable proof completely programmed. Study will be supported out for the discovery & its pertinent constraints Wide-ranging collected works survey has been done in the hypothetical analysis in the domain of medical solicitations.

In the reference of hypothetical analysis to give a vision into numerous AI models and its diagnostic exactness in relations of the recital, accuracy improvement from 2.05% to 12.4% across various models.

REFERENCES

[18] Hsin-Yi Tsao, Pei-Ying Chan, Forecasting diabetic retinopathy and classifying interpretable biomedical landscapes by machine learning algorithms BMC Bioinformatics; 2018

AUTHORS PROFILE

Lubna Taranum MP received B.E. degree and M.Tech from VTU University in 2008 and 2011, respectively. She was Assistant professor in engineering college from 2011 to 2013. Currently she is a research fellow from VTU University, Karnataka. Her research interest includes medical image processing, Machine learning, pattern recognition.

Rajashekar J. S. received B.E. degree and M.Tech from VTU University, respectively and the PhD degree from Jain University. His primary area of interest Process Instrumentation and control. He started his career as a lecturer at P V P Institute of technology, Sangli, Maharashtra. Presently he is working as a Professor and head of the department in electronics instrumentation and engineering, Dayananda Sagar College of engineering, Bengaluru. Over 32 years of teaching and supervising research projects at various levels, he has multitudes of experience and publications both in practical and theoretical aspects in his areas of interest.