Developing a diagnosis system for the Otitis media diseases based on the color image segmentation

Musa Alrefaya[1], Nabil Ashour[2], Omar Ghaith[1] , Khaled Juba[1],Yousef Qwasmeh[1] and Haneen Dweik[1].

[1]*Department of Computer Science ,Palestine Polytechnic University, Hebron Palestine*

[2]*Department of ENT, AL Ahli Hospital, Hebron-Palestine.*

Abstract

Middle ear diseases diagnosis is highly dependent on the doctor's experience thus some of the abnormalities can be easily missed even by experienced doctors. The process of reading medical images, detecting symptoms, and reaching a true conclusive diagnosis is a process that depends heavily on the doctor's empirical knowledge, memory, intuition, and diligence. All of which are uncontrollable factors, which leaves the diagnostic process inconsistent, and prone to error for doctors who have been in the field for a while, and makes it extremely difficult for newly graduated doctors. This research aims to develop smart diagnosis system to support ENT (ear nose and throat) otolaryngologist that can identify and highlight lesions and regions of interests, and provide a second diagnosis opinion. For all patients fully Audiological and video otoscopic examination were done by the ENT otologists.

Keywords: Otitis media, Computer diagnosis systems, Middle ear, Segmentation.

1. Introduction

Computer-aided diagnosis (CAD) in medicine is the result of a large amount of effort and researches expended in the interface of medicine and computer vision [10]. CAD systems in medicine use diagnostic rules to emulate the way an expert dotcors makes diagnostic decisions. Middle ear diseases diagnosis is highly dependent on the doctor's experience thus some of the abnormalities can be easily missed even by experienced doctors. The goal of this work is to develop a smart diagnosis system to support otologists that can identify and highlight lesions and regions of interests, and provide a second diagnosis opinion. The system identify middle ear image features that can distinguish and set apart different diseases with. It is targeting to minimize misdiagnosis where one of these cases occur:

a) **False positives**: The error causes in data reporting, which a test result is positive but it is not present.

b) **False negatives:** The error causes in a test result, which it is negative but it is present.

For making a precisise daignosis systems, doctors need to inspect images (visual perception) and interpret what they see (cognition). There studies have shown that clearly present abnormalities are, at times, not reported or misinterpreted.

In this work, images of the middle ear (tympanic membrane/eardrum) are acquired from the live feed of endoscope that's attached to an ear nose & throat unit (ENT). After that, we perform image processing operations for enhancing the acquired images, and extract relevant features. The final stage is implanted to classify the image into the right disease category. For all patients fully audio logical and video endoscopic examination were done by the ENT otologists.

The rest of the paper is organized as following: the theoretical background in section 2. Recent related researches are summarized in section 3. The proposed system is illustrated in section 4. The experiments and results are presented in section 5. Finally, we conclude the paper in section 6. Figure 1 shows the used smart ENT unit.

Figure 1: The designed smart ENT unit

2. Theoretical background

In this section, we discuss the tympanic membrane, which is the part of ear, as well as the different types of middle ear conditions, and their most dominant characteristics starting with otitis media, which is the most common ear disease [2,3,4].

Tympanic membrane: tympanic membrane or (eardrum) is the barrier that separates the middle and outer ear from the external ear. It is a thin, circular layer of tissue that marks the point between the middle ear and the external ear. It is approximately 0.1 mm thick, 8 to 10 mm in diameter, and has a mass weight of around 14 mg. Figure 2 shows a healthy tympanic membrane.

Figure 2: Healthy tympanic membrane

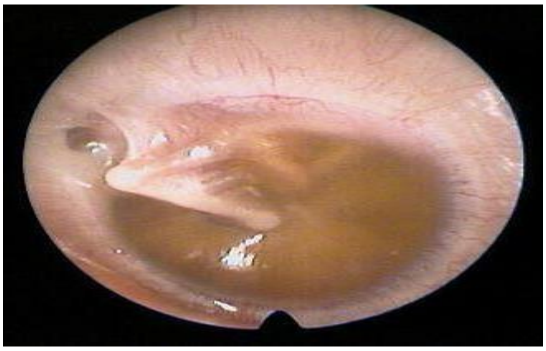
Otitis media (OM): Otitis media is a general term for middle-ear inflammation, and it is one of the most common childhood ear diseases worldwide and the second most important cause of hearing loss. Otitis media is classified into three different types:

1- Acute otitis media (AOM): is a common condition seen in primary care offices, as 1 in 4 children will have at least 1 episode of AOM by age 10 years. It results from infection of fluids that has become trapped in the middle ear. Signs of presence of AOM are redness, swelling, and bulging TM. In figure 3 we can see how AOT affects the tympanic membrane.

Figure 3: The effect of acute otits media on the tympanic membrane

2- Otitis media with Effusion (OME): otitis media with effusion is caused by the fluids pooling in the middle ear after the body fails to extract them, but unlike AOM the fluids are not infected, figure 4 shows Otitis media with effusion.

Figure 4: Otitis media with effusion



3- Chronic supportive otitis media (CSOM): CSOM is an escalation of AOM that occurs when the infection causes a perforation in the tympanic membrane. Figure 5 exhibits a case of CSOM and shows how does it affect the tympanic membrane.

4- Bullous myringitis: it is a type of ear infection in which small, fluid-filled blisters form on the eardrum. These blisters usually cause severe pain. Figure 6 shows how does the tympanic membrane look with this infection.

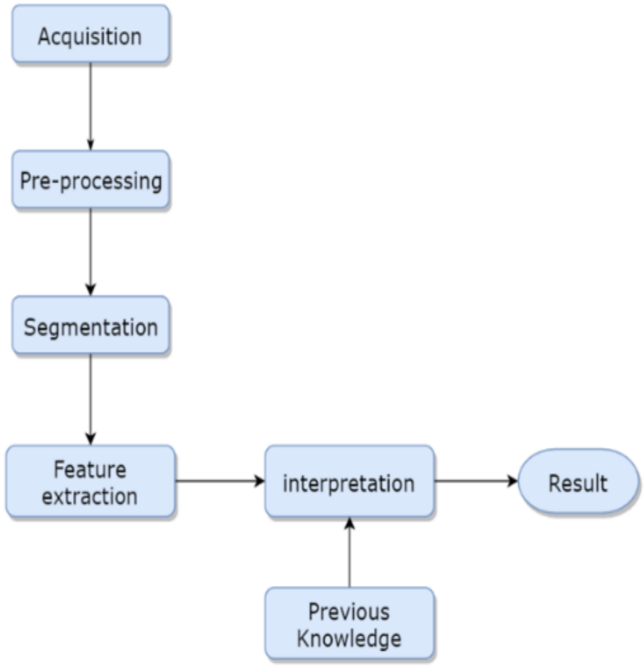
Figure 5: Chronic supportive otitis media and Figure 6: Bullous myringitis

2.1 Image processing:

Image processing is one of tools that enable computer vision. It is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image.

An image that goes through image processing operations generally follows the pipeline shown in figure 7:

Figure 7: Image processing pipeline



- Image acquisition: the image can be captured by cameras, using sensors like CCD to transform the analog signal into digital.
- Pre-processing: involves improving the quality of the image by removing noise and artifacts or perform image transformation.
- Segmentation: involves isolating an object in the image from the background to remove any unnecessary information.
- Feature extraction: involves extracting predefined relevant information from the image. It is basically a function that transforms the image into a compact form (feature) that holds enough information about the original image to allow application specific operations to be performed.
- Classification: involves taking a decision based on the extracted features, and sometimes previous knowledge.

2.2 Computer Aided System

Computer aided systems are used in many different areas, and have been a great assistance for people in terms of speed and accuracy. One of the main fields that extensively use CAD is the medical field. In theis area a computer aided system is defined as a computer based diagnosis system that helps doctors take decisions swiftly by providing a second opinion, using a set of technologies that usually involves artificial intelligence (AI), computer vision, and medical image processing [10]. CAD systems can be divided into two categories:

Computer aided detection systems (CADe): which are systems designed to decrease observational oversights and thus the false negative rates—of doctors interpreting medical images.

Computer aided diagnosis systems (CADx): which are systems designed to analyze the patterns and textures in medical images to distinguish and classify them into different classes.

3. Previous Work

In this section, recent publications that are relevant to this work, methodology, approach, and results are presented.

Kuruvilla et al. (2013) reported an automated algorithm based on image-analysis to distinguish between three types of Otits Media (AOM, OME and Otiits media with no effusion (NOE) [2].

In order to enable the system to classify images into the right categories, they needed to understand the thought process of doctors when they try to diagnose the disease through images, so an experiment was conducted to identify what needs to be accounted for in image classification and to understand the decision making behind it.

The outcome of the experiment was the set of features that the system is going to extract and use. The way a decision is taken based on these features to classify the images. Authors extracted eight features from the images**:**

- Bulging: Bulging is crucial for diagnosing AOM, a feature to measure the bulging percentage in the tympanic membrane is developed. The way it measures it is by giving each pixel a depth value that represents the bulging.

- Central concavity: The tympanic membrane is attached firmly to the malleus that is one of the three middle ear bones called auditory ossicles. In the presence of an infection, the tympanic membrane begins to bulge in the outer area, while leaving the area that is closely attached to the bone without bulging, which causes a concavity, and this feature test the presence of concavity.

- Light: The devices used to take pictures of the tympanic membrane have a light attached to them, the illumination is somewhat evenly distributed when the tympanic membrane is not infected, but when there is an infection causing bulging, the illumination is distributed uniformly, and this feature measures the uniformity of the illumination.
- Malleus presence: In the case of OME or no infection, the tympanic membrane stays neutral, which leaves the malleus bone present, this feature is responsible for detecting the presence of the bone in the taken image.
- Translucency: This feature describes the level of transparency of the tympanic membrane (the extent to which it allows the light to pass through it), it plays an important role to distinguish between OME and NOE.

- Amber level: In the case of OME the tympanic membrane's color is yellow (amber), so this feature describes the level of amber color in the image, and it is important to distinguish between OME and the other classes.
- Bubble presence: The presence of bubbles behind the tympanic membrane is a sign of OME, and this feature is designed to detect the presence of bubbles.
- Grayscale variance: This feature measures the variance of color values in the grayscale image. OME has a more uniform appearance than AOM and NOE and has consequently a much lower grayscale variance that can be used to distinguish it from the rest.

The algorithm responsible for classifying the images follows the following flow: After calculating each feature and assigning a value to it, the path is taken based on wither the value satisfies the threshold or not. The value of the threshold for each feature is pre-determined during the training phase of the algorithm. The results for this algorithm showed a high level of accuracy (89.9%). In Figure 8 shows Classification decision tree.

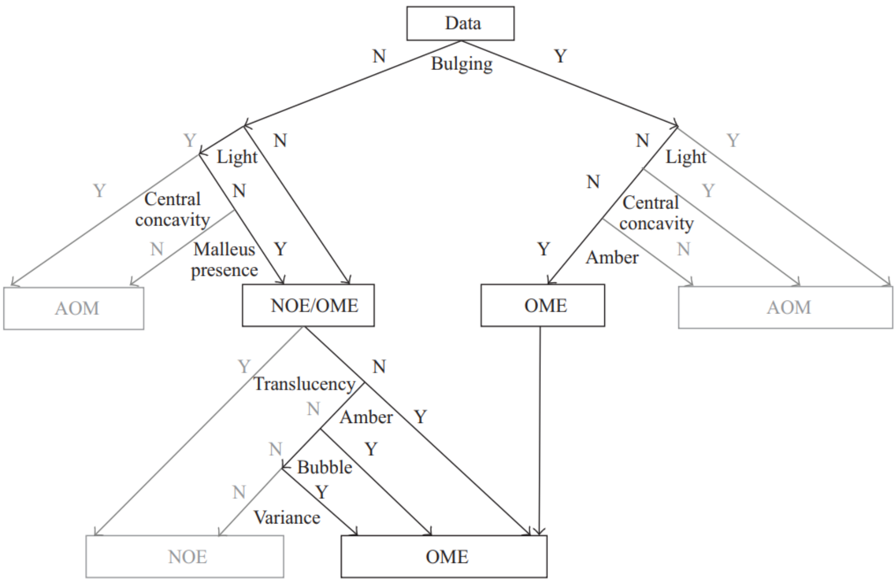


Figure 8: Classification decision tree

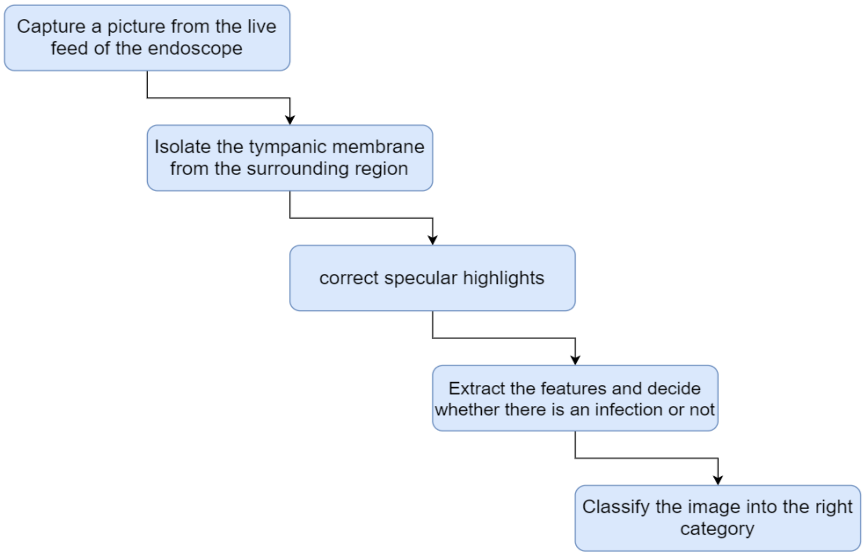
4. The proposed system: Otites media diagnosis system based on endoscope image

This section illustrates the propose system that involves five different stages starting with image acquisition and ending with classifying the image into one of the following categories:

1. Normal/healthy tympanic membrane.
2. Acute otitis media.
3. Otitis media with no effusion.

In this work, images of the middle ear (tympanic membrane/eardrum) are acquired from the live feed of endoscope that's attached to an ear nose & throat unit (ENT). After that, we perform image processing operations for enhancing the acquired images, and extract relevant features. The final stage is implanted to classify the image into the right disease category. Figure 9 shows the block diagram for the proposed system.

Figure 9: The proposed system block diagram



4.1 Image acquisition:

The system is captureing images from the live feed of endoscopes upon request from the doctor. The image ideally should clearly show only the tympanic membrane without any obstructions like hair and wax or extras from the surrounding area.

4.2 Preprocessing:

Since images are not going to always be in an ideal situation, in this stage we eliminate or minimize the impact of image artifacts associated with endoscopic images, which fundamentally consist of specular highlights (a bright spot of light that appears on shiny objects when illuminated) using median filter, and also isolate the tympanic membrane from the irrelevant regions through segmentation.

Segmentation: As explained previously segmentation is the process of isolating region of interest from noise or unwanted information in an image. In this work, we developed a simple but fast segmentation approach, called it the rigid segmentation approach.

4.3 Rigid segmentation:

The proposed approach at the beginning find the central point of an image by dividing the number of horizontal and vertical pixels by two, then this point is used as a center for a circle with constant radius of 100 (determined experimentally) and we isolate this region from the rest of the image. We developed this approach because we noticed that most of images have the tympanic membrane located at their center. However, in cases where this is not true, crucial parts of the image were lost as a result of this segmentation. Figure 10 shows sample images of this segmentation.

Figure 10: Rigid segmentation

4.4 Dynamic segmentation:

While dynamic segmentation produced better results, the same problem was still present in some images and some parts of the tympanic membrane were lost in the process. In an attempt to include the regions outside of the circular region we first created an image resembling the circle. After that we reduced the number of intensity values in the grey scale image to 9 levels by dividing the values by 30, we assume that regions that lie within the circle and extend even beyond are part of the tympanic membrane, and hence we include them in the segmented region. Figure 11 shows an example of image segmented using this approach.

4.5 Connected regions based segmentation:

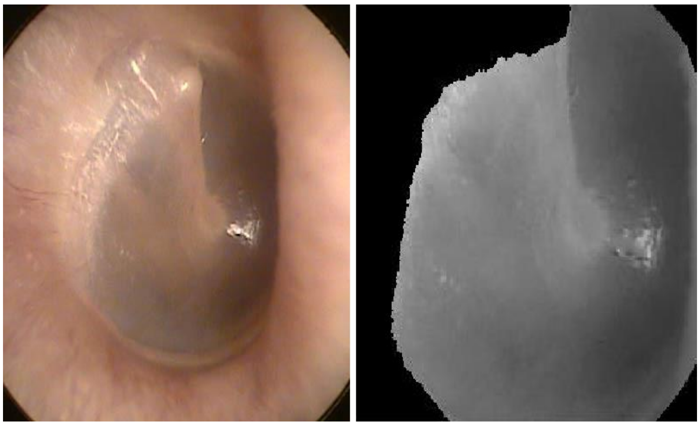
Connected regions based segmentation: While dynamic segmentation produced better results, the same problem was still present in some images and some parts of the tympanic membrane were lost in the process. In an attempt to include the regions outside of the circular region we first created an image resembling, after that we reduced the number of intensity values in the grey scale image to 9 levels by dividing the values by 30, we assume that regions that lie within the circle and extend even beyond are part of the tympanic membrane, and hence we include them in the segmented region. Figure 12 shows an example of image segmented using this approach.

Figure 11: Dynamic segmentation and Figure 12: connected regions based segmentation

4.6 Active contour based segmentation:

Connected regions segmentation gave us acceptable results, but we wanted to further enhance our segmentation so we tried using active contours to observe the results and decide whether it is worthwhile to use a time intensive approach or not. Active contours based segmentation is a powerful edge based segmentation where an initial boundary or contour is given to the algorithm as a starting point, then the algorithm shrinks this boundary through repetitive iterations, in each iteration we look for a smaller well defined boundary until we outline the object we are trying to extract, or until we finish a preset number of iterations. Figure 13 shows the result of this segmentation on one of the images.

Figure 13: Segmentation using active contours



4.7 Specular highlight:

The presence of wax on the tympanic membrane causes light coming from the endoscope to reflect, which results in white regions that heavily affect feature computation, and hence must be corrected.

First we tried to correct illumination using the region surrounding each pixel, we used a 25*25 window around the pixel in each channel. The formula we used takes every channel individually and calculates the summation of all values around the pixel with the exception of those affected by reflection, and divides by the number of pixels, then we used the new values from each channel in place of the pixel. Although the results eliminated the reflection problem, the resulting regions were mostly gray, and one could easily see that the image was modified. Figure 14 shows the result of using this formula.

Figure 14: Specular highlight

5. Experiments and Results:

In this work, incoming images classify into either normal or abnormal as a first step, then it becomes easier to classify abnormal images into acute otitis media, or otitis media with effusion solely based on color information. Although color based classification is simple and error prone, our observations and the doctor's input made it clear that when trying to classify only into two classes with such a big difference in colors, color information can be very distinctive.

The feature vector used for classification consists of 5 features, level of grayness (G), presence of the malleus bone (B), number of connected components (CC), redness level (R), and the level of yellow/amber color in the image (Y). The first 3 features (G, B, CC) are used to determine whether the image is normal or abnormal and the last two features (R,Y) are used to determine the class of the disease.

1- **Level of grayness:** As normal tympanic membrane's color is predominantly gray we wanted to calculate a feature that describes that, the feature is simply calculated by converting the image from the RGB model into the HSV model then finding the individual pixels that lie within the gray color range, we then calculate how much of the non-zero pixels does the gray pixels form (the percentage of gray pixels in the image). The resulting percentage is then tested against a threshold found experimentally to determine whether this feature indicates that the image is normal or abnormal. Figure 15 shows HSV color sapce.

2- Presence of the malleus bone: In a normal healthy ear the malleus bone is clearly visible and well defined, and this feature is designed to detect the presence of this bone. The detection is based on finding the edge that resembles the bone in an edge image, the edge image however is not computed from the original image but from a grey level image with intensity values reduced to only 5 by dividing the values by 60, edges are computed using canny edge detection, then we use dilation to thicken the edges and make the longest line of contiguous pixels more clear. Our computation of this feature based on the assumption that if there is a bone, then the longest line of contiguous pixels found in the dilated edge image is the edge of that bone. The way we detect the longest line is by defining a horizontal region along the image with a set width, then we iterate through the rows in that region to find the longest line. In order for us to traverse the whole image we then rotate the image by 3 degrees time and time again until we have rotated the image 180 degrees, repeating the process of finding the longest line each time and with each time holding the number of row in which this line has occurred and the degree of rotation. We then go back to the original image, rotate it by the degree in which we found the longest line, then we define a fixed region around the row number of the found line. In the last step of this process we find the edge image of the original image using canny edge detection, then we count the number of edge points within the region. Then we test the number against a threshold found experimentally, if the count is less than the threshold then there is a bone if not then this is not a bone. This calculation is based on our observation that in healthy images the shape of the tympanic membrane is uniform without any irregularities which leaves the number of edge points of the surrounding region relatively small compared to abnormal images. Figures from (16-18) demonstrate the various stages in this operation.

Figure 16: presence of the malleus bone

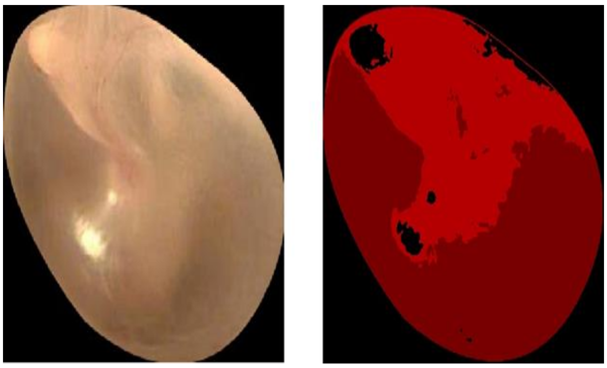


Figure 17: presence of the malleus bone

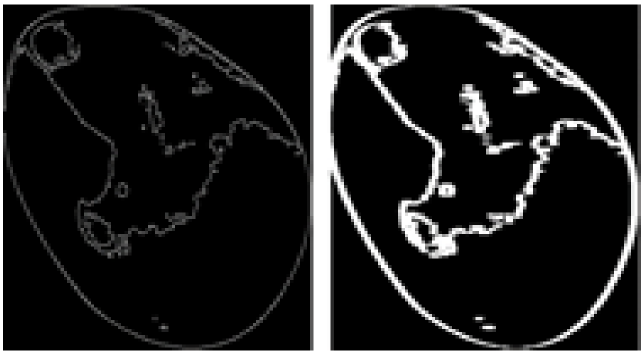
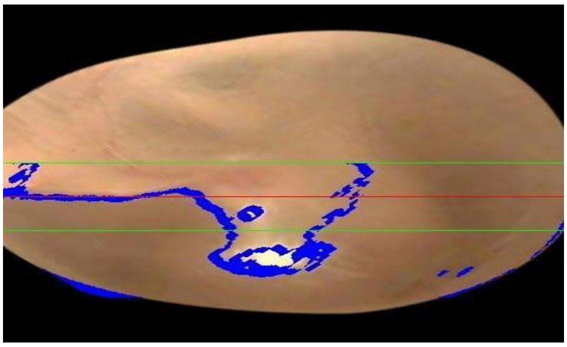


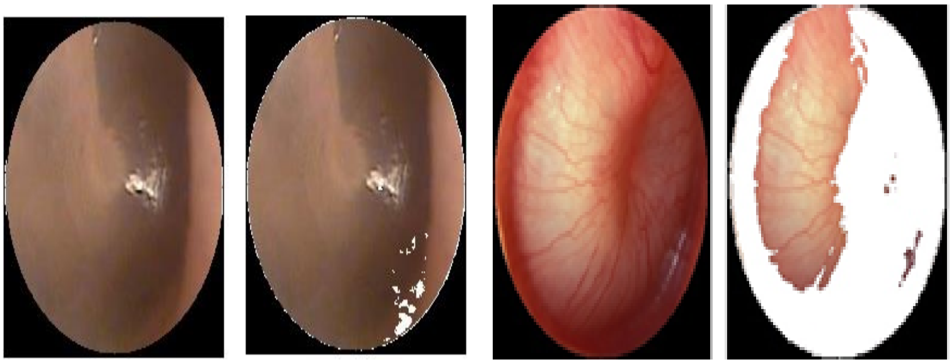
Figure 18: presence of the malleus bone



2- **Connected component:** As mentioned previously, the uniform shape of the normal ear and the irregular form of the abnormal ear is one of the most distinctive clues that aid the diagnosis process, to describe the form of the tympanic membrane in Kuruvilla et al. (2013) bulging feature was used to describe the depth of the object [2], however for our inability to implement this feature we resorted into using this feature to describe the shape. We calculate this feature by first reducing the number of intensity values in the gray level image to half, this way we can group pixels close to each other in value into a single component, and because of the uniform nature of the normal ear the resulting number of regions is small relative to the regions resulting from an abnormal image, the reason for that is the non-uniform distribution of light over an infected tympanic membrane that results from its irregular shape. To count the number of connected components in an image we iterate over it pixel by pixel, and for each pixel mark it as visited and check the pixels around it for the same value, once a pixel is found we mark it as visited and check its neighbors, the algorithm traverses the surrounding pixels in a depth first search manner, and goes back recursively when it's finished with each pixel. We repeat the process for each unvisited pixel till we have traversed the whole image. The counter that holds the number of regions is then checked against a threshold calculated experimentally to determine whether the image is normal or abnormal. Images for this feature are not necessary, because the differences cannot be realized visually.

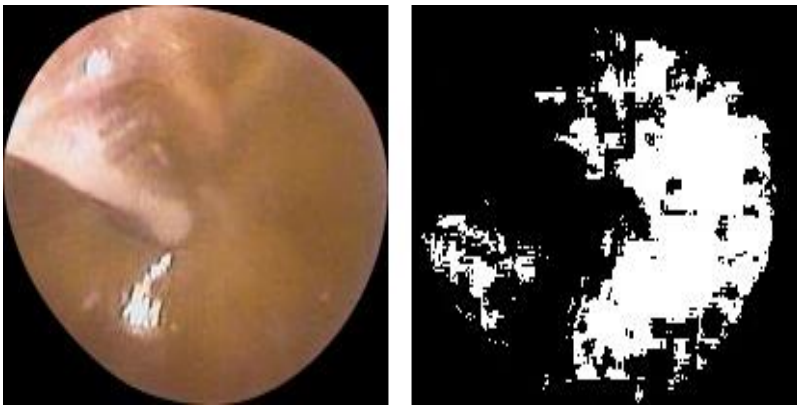
4- Redness level: From our observation of the available images, and after consulting with the doctor we found out that one of the most prominent features in an acute otitis media infected ear is the redness, so we decided to use this feature to classify images into this class. The feature is simply calculated by counting the number of pixels that has a value of R that is 65% or higher relative to the other colors (G,B), for each pixel we calculate the summation of the three channels RGB at that pixel, then divide the value of the R channel at that pixel's location by the summation, then we compare it to the threshold. In the end we divide the number of pixels with high value of R by the total number of non-zero pixels to calculate percentage of red pixels relative to others. If the percentage is higher than the defined threshold then the image is classified as acute otitis media. Figures 19 show examples with both cases, healthy and infected.

Figure 19: Redness level (helathy and infected)



5- Yellowness level: This feature is used to classify images into otitis media with effusion. Similar to the previous one, this feature is calculated by counting the number of pixels that contain a large percentage of yellow, however finding those pixels is slightly different, we have to ensure that the percentage of red in each pixel is 35% or more and at the same time the percentage or green is 35% or higher, those are calculated in the same manner by dividing each channel (R,G) by the summation of all three channels. The number of pixels is then divided by the total number of non-zero pixels in the image to calculate the percentage, which is then compared to the threshold to make the classification decision. Figure 20 shows an example where the pixels are detected and marked.

Figure 20: the level of yellow/amber color in the image (Y)



6. Conclusion

In this work, an automated feature extraction and image classification system for otitis media was proposed and developed. Images of the middle ear acquired from the live feed of endoscope that attaches to an ENT unit. After that, an image processing procedures have been applied for enhancing and extracting significant featerures for the input images. The last stage was to classify normal and abnormal images. The results shows that the proposed system classified the input images and diagnose the abnormal cases.

References

[1] Elaine C. Jong, Dennis L. Stevens, Netter's Infectious Diseases Elsevier Health Sciences,2012.

[2] Kruvilla, A., Shaikh, N., Huberman, A., & Kovačević, J. Automated diagnosis of otitis media: Vocabulary and grammar. International Journal of Biomedical Imaging. Volume 2013, 15 pages,2013.

[3]Senaras.C,Moberly.C.A,Teknos.TH,Essig.G,Elmaraghy.CH,Taj-Schaal.N, Yu.L,and Gurcan.M, (March 2017)." *Autoscope: automated otoscopy image analysis to diagnose ear pathology and use of clinically motivated eardrum features*", Proc. SPIE 10134, Medical Imaging 2017: Computer-Aided Diagnosis, 101341X 3.

[4]Lieberthal.A,Carroll.A,Chonmaitree.T,Ganiats.TH,Huberman.,Jackson.M, Joffe.M, Millard,Rosenfeld.R,Sevilla.X,Schwartz, Thomas.P and Tunkel.D, (March 2013). " *The Diagnosis and Management of Acute Otitis Media,"*Pediatrics Journal.

[5]Chuen-Kai Shie ; Chung-Hisang Chuang ; Chun-Nan Chou ; Meng-Hsi Wu ; Edward Y. Chang, 2015." *Transfer representation learning for medical image analysis"*, Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

[6] Michael E. Pichichero,(2013)." *Otitis Media"*,Pediatric clinics of North America Volume 60, Issue 2, Pages 391–407.

[7] Hermanus C.,Myburgha Willemien H.van ZijlaDe Wet Swanepoelbcg Sten Hellströmde (March 2016).*"Claude Laurent Otitis Media Diagnosis for Developing Countries Using Tympanic Membrane Image-Analysis"*, EBio Medicine, Pages 156-160.

[8] Chuen-KaiShie ; Hao-Ting Chang ; Fu-Cheng Fan ; Chung-Jung Chen ; Te-Yung Fang ; Pa-Chun Wang,(2014)." *A hybrid feature-based segmentation and classification system for the computer aided self-diagnosis of otitis media*",Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

[9] Ionuţ Mironică ; Constantin Vertan ; Dan Cristian Gheorghe,(2011)." *Automatic pediatric otitis detection by classification of global image features* ",E-Health and Bioengineering Conference (EHB).

[10] Yanase J, and Triantaphyllou E, 2019, A Systematic Survey of Computer-Aided Diagnosis in Medicine: Past and Present Developments, Expert Systems with Applications, Vol. 138, in press. https://doi.org/10.1016/j.eswa.2019.112821.

[11] Cha D, Pae C, Seong SB, Choi JY, Park HJ. Automated diagnosis of ear disease using ensemble deep learning with a big otoendoscopy image database. EBioMedicine. 2019; 45:606–614. doi:10.1016/j.ebiom.2019.06.050