DESIGNING, WORKING & FUNCTIONING OF BIONIC KIDNEY

PROJECT REPORT

Submitted for the course: Introduction to Innovative Projects (PHY1999)

By

MEERA L	**18BEC0597**
DEVYANI BISWAS	**18BME0658**
ASHWER VASHIST	**18BME0977**
KAUSTUBH YADAV	**18BCB0150**

Slot: TF1

Name of faculty: Prof. Ezhil Vizhi R

(SCHOOL OF ADVANCED SCIENCES)

VIT UNIVERSITY
(Estd. u/s 3 of UGC Act 1956)
VELLORE ■ CHENNAI
www.vit.ac.in

October, 2020

CERTIFICATE

This is to certify that the project work entitled "***Designing, Working and Functioning of Bionic Kidney***" that is being submitted by "***Meera L, Devyani Biswas, Kaustubh Yadav and Ashweer Vashist***" for Engineering Physics (PHY1001) is a record of Bonafide work done under my supervision. The contents of this Project work, in full or in parts, have neither been taken from any other source nor have been submitted for any other CAL course.

Place : Vellore

Date : 28-10-2020

Signature of Students:

Meera L
Devyani Biswas
Kaustubh Yadav
Ashwer Vashist

Signature of Faculty:

Prof. Ezhil Vizhi R

ACKNOWLEDGEMENTS

We would like to express my gratitude to all those who have helped me in the successful completion of this project. Without their support, we would not have been able to achieve the goal of the project successfully. We would like to take this opportunity to thank our guide, ***Prof. Ezhil Vizhi R***, for her constant support, guidance and mentorship to complete the project on time.

Finally, we would like to express my sincere gratitude to VIT University, which provided me with a platform to hone our skills.

18BEC0597
18BME0658
18BCB0150
18BME0977

ABSTRACT

This project deals with the cumbersome disease that is faced by a majority of the world's population- Kidney Diseases and the toll it takes on people.The usual process of medical aid revolves around the painful procedure of dialysis and the financial and emotional burden it takes on the patients. In this project we have focused on four primary domains namely: Data Science , Image processing , Material selection and Designing. With data science codes we predict the functionality of the kidney by comparing it to a the parameters of a healthy kidney, image processing helps us understand and track the development of kidney stones post-op in the bionic organ , the martial has been selected with utmost care keeping in mind the proper compatibility of the material used and the kind of enzymes of the body and finally the kidney structure has been designed to try and replicate the original organ to reduce the difficulties of the patient and the doctor. Towards the end of the report we have displayed the results we have obtained in these domains respectively.

1. INTRODUCTION:

1.1 OBJECTIVE AND GOAL OF THE PROJECT

Bio Mechanical Field has always been one of the most promising fields in terms of innovative prospects. With the evolution of Data Science that collects data from the targeted patients and compares them with another set of pre-existing data collected from thousands of other sources, the scope of a perfect diagnosis is almost guaranteed. With the chronic shortage of kidney donors, researchers from Cal Tech and University of Vanderbilt have combined forces to come up with artificial kidneys as a long-term solution. While the designing of the kidney is completed, it still suffers the problem of regulating blood flow to avoid the formation of blood clots. This is done by the means of simulations as the flow of blood follows mechanical forces. We intend to design a similar kidney model and reduce the chance of blood clotting in it using data science. In case the patient shows symptoms of artificial kidney malfunction, he/she can be given immediate diagnosis at the hospital after providing some values based on specific parameters as input to the code we will be writing. The accuracy of the result will determine the extent of damage if any within seconds compared to the usual time period of waiting for the results of a blood test.

1.2 DETAILED LITERATURE SURVEY

A) DETECTING CHRONIC KIDNEY DISEASE USING MACHINE LEARNING ALGORITHMS

Chronic kidney disease, also called chronic kidney failure, is the gradual loss of kidney function. Kidneys filter wastes and excess fluids from the blood, which are then excreted. When chronic kidney disease reaches an advanced stage, dangerous levels of fluid, electrolytes and wastes can build up in the body. In the early stages of chronic kidney disease, the symptoms are negligible. Chronic kidney disease may not become apparent until the kidney function is significantly impaired. Chronic kidney disease can progress to end-stage kidney failure, which is fatal without artificial filtering (dialysis) or a kidney transplant.

B) USING IMAGE PROCESSING AND DEEP LEARNING TO DETECT CHRONIC KIDNEY INFECTION

Kidney Stones are one of the most common diseases infecting about 1 million Indians a year. It's caused due to the deposition of acid salts near the Glomerulus of the Kidney, which in turn causes extremely sharp pains. Kidney Stones might be a result of a bad lifestyle or might be a genetic implication. The initial methods of Kidney Stone Detection were using X-Rays, but currently, CT-scans provide a much accurate alternative. In our method we use, both X-Ray Scans and CT-Scans to detect Kidney Stones.

C) MATERIALS

Regardless of the type of artificial kidney used the patient's blood must be rendered incoagulable by the injection of heparin. The blood is then guided along a cellophane membrane on the other side of which is the rinsing fluid. By the process of dialysis a large part of the abnormally retained products of metabolism. In addition to this process an exchange takes place between the necessary electrolytes of the blood plasma water and those in the rinsing fluid. Hence, if the patient's blood electrolyte pattern is abnormal before treatment, it will tend to be corrected as it approaches the composition of the rinsing fluid in as much as the fluid contains normal concentrations of these ions. Since water passes through the cellophane easily, careful attention must be given the osmotic pressure on the two sides of the membrane. On the inside the blood plasma protein tends to draw water from the rinsing fluid to the blood. This can be prevented by making the rinsing fluid isotonic with the addition of glucose. If desired it can be made hypertonic and capable of withdrawing fluid from the patient, a maneuver which may be particularly gratifying to patients with pulmonary edema.

D) MODEL

2. Methodology

A) DETECTING CHRONIC KIDNEY DISEASE USING MACHINE LEARNING ALGORITHMS

Google Collaboratory, or "Collab" for short, a product from Google Research is used for simulating the code. Collab allows to write and execute arbitrary python code through the browser, and is well suited to machine learning and data analysis.

DATASET DESCRIPTION:

The data set was obtained from Kaggle. The data set consists of various entries of patients such as age, blood pressure, RBC count and other parameters which can affect the patient's kidney.

Name of attribute	Attribute Abbreviation	Description of attribute
Age	ag	Patient age as numeric attribute with minimum age of 2 years and maximum age of 90 years.
blood pressure	bp	BP with minimum 50 mmHg and maximum 180 mmHg.
specific gravity	sg	Nominal attribute with five distinct values.
Albumin	al	Nominal attribute with six distinct values.
sugar	su	Nominal attribute with six distinct values.
red blood cells	rbc	Nominal attribute with normal and abnormal blood counts.
pus cell	pc	Nominal attribute categorized as normal and abnormal cells.
pus cell clumps	pcc	Nominal attribute categorized as present and notpresent cell clumps.
Bacteria	ba	Nominal attribute divided into present and notpresent.
blood glucose random	bgr	Numeric attribute ranging from 22 mg/dl to 490 mg/dl.
blood urea	bu	Numeric attribute ranging from 1.5 mg/dl to 391 mg/dl.
serum creatinine	sc	Numeric attribute ranging from 0.4 mg/dl to 76 mg/dl.
Sodium	sod	Numeric attribute ranging from 4.5 mEq/L to 163mEq/L.
Potassium	pot	Numeric attribute ranging from 2.5 mEq/Lto 47mEq/L.
Hemoglobin	hemo	Numeric attribute ranging from 3.1 gms to 17.8 gms.
packed cell volume	pcv	Numeric attribute ranging from 9 to 54.
white blood cell count	wc	Numeric attribute with minimum 2200 cells/cumm and maximum 26400 cells/cumm.
red blood cell count	rc	Numeric attribute with minimum 2.1 millions/cmmand maximum 8 millions/cmm.
hypertension	htn	Nominal categorization as yes or no.
diabetes mellitus	dm	Nominal categorization as yes or no.
coronary artery disease	cad	Nominal categorization as yes or no.
appetite	appet	Nominal categorization as good or poor
pedal edema	pe	Nominal categorization as yes or no.
Anemia	ane	Nominal categorization as yes or no.
Class	class	Categorized as having CKD or notCKD.

INITIAL DATASET AND CLEANING

The raw data consists of lot of missing values, and the datatype is not uniform. In this data set, few values are of the type 'int' and the rest are of type 'string'

```
data
```

	'age'	'bp'	'sg'	'al'	'su'	'rbc'	'pc'	'pcc'	'ba'	'bgr'	'bu'	'sc'	'sod'	'pot'	'hemo'	'pcv'
0	48	80	1.02	1	0	?	normal	notpresent	notpresent	121	36	1.2	?	?	15.4	44
1	7	50	1.02	4	0	?	normal	notpresent	notpresent	?	18	0.8	?	?	11.3	38
2	62	80	1.01	2	3	normal	normal	notpresent	notpresent	423	53	1.8	?	?	9.6	31
3	48	70	1.005	4	0	normal	abnormal	present	notpresent	117	56	3.8	111	2.5	11.2	32
4	51	80	1.01	2	0	normal	normal	notpresent	notpresent	106	26	1.4	?	?	11.6	35
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
395	55	80	1.02	0	0	normal	normal	notpresent	notpresent	140	49	0.5	150	4.9	15.7	47
396	42	70	1.025	0	0	normal	normal	notpresent	notpresent	75	31	1.2	141	3.5	16.5	54
397	12	80	1.02	0	0	normal	normal	notpresent	notpresent	100	26	0.6	137	4.4	15.8	49
398	17	60	1.025	0	0	normal	normal	notpresent	notpresent	114	50	1	135	4.9	14.2	51
399	58	80	1.025	0	0	normal	normal	notpresent	notpresent	131	18	1.1	141	3.5	15.8	53

400 rows × 25 columns

Figure 1: Raw Data Set

Cleaning the dataset involves:

- Removing null values
- Converting all the entries to a uniform data type, in this case it is float

```
[ ] data.head()
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	bu	sc	sod	pot	hemo	pcv	wbcc	rbcc	ht
3	48.0	70.0	1.005	4.0	0.0	2.0	1.0	2.0	1.0	117.0	56.0	3.8	111.0	2.5	11.2	32.0	6700.0	3.9	2.
5	60.0	90.0	1.015	3.0	0.0	0.0	0.0	1.0	1.0	74.0	25.0	1.1	142.0	3.2	12.2	39.0	7800.0	4.4	2.
9	53.0	90.0	1.020	2.0	0.0	1.0	1.0	2.0	1.0	70.0	107.0	7.2	114.0	3.7	9.5	29.0	12100.0	3.7	2.
11	63.0	70.0	1.010	3.0	0.0	1.0	1.0	2.0	1.0	380.0	60.0	2.7	131.0	4.2	10.8	32.0	4500.0	3.8	2.
12	68.0	70.0	1.015	3.0	1.0	0.0	2.0	2.0	1.0	208.0	72.0	2.1	138.0	5.8	9.7	28.0	12200.0	3.4	2.

Figure 2: After Cleaning the Data Set

For prediction we are using 2 classification algorithms

- Random forest algo: decision trees are created for the data samples. Then, prediction from each of them are obtained and best solution is chosen.
- Naïve bayes: This classification technique is based on Bayes' theorem. And we assume all predictors are independent of each other.

In this project, Gaussian Naïve Bayes classifier is used having the assumption that the data from each label is drawn from a simple Gaussian distribution.

```
[ ]  from sklearn.naive_bayes import GaussianNB
     from sklearn.ensemble import RandomForestClassifier

     models=[
             ('rf',RandomForestClassifier()),
             ('nb',GaussianNB())
     ]
```

```
[ ]  from sklearn.model_selection import train_test_split
     X_train, X_test,y_train,y_test=train_test_split(X,y,test_size=0.2)
     #train 80% data and test 20% data
     for i,j in models:
         clf=j
         clf.fit(X_train,y_train)
         acc=clf.score(X_test,y_test)
         print(i,acc)

     #printing accuracy of naive base and random forest

     rf 1.0
     nb 0.9512195121951219
```

Figure 3: After applying the Classification Algorithms

OUTPUT

- For this project, 80% of the total data set is trained and the rest 20% is tested using the classification algorithms.
- Random Forest Algorithm is 100% accurate whereas Gaussian Naïve Bayes classifier is 95.1295% accurate

PLOTTING CORRELATION BETWEEN VARIOUS PREDICTORS

```
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.metrics import roc_curve, auc, confusion_matrix, classification_report,accuracy_score
from sklearn.ensemble import RandomForestClassifier
import warnings
warnings.filterwarnings('ignore')
%matplotlib inline
```

Figure 4: Importing various libraries

```python
def auc_scorer(clf, X, y, model): # Helper function to plot the ROC curve
    if model=='RF':
        fpr, tpr, _ = roc_curve(y, clf.predict_proba(X)[:,1])
    elif model=='SVM':
        fpr, tpr, _ = roc_curve(y, clf.decision_function(X))
    roc_auc = auc(fpr, tpr)

    plt.figure()    # Plot the ROC curve
    plt.plot(fpr, tpr, label='ROC curve from '+model+' model (area = %0.3f)
    plt.plot([0, 1], [0, 1], 'k--')
    plt.xlim([0.0, 1.0])
    plt.ylim([0.0, 1.05])
    plt.xlabel('False Positive Rate')
    plt.ylabel('True Positive Rate')
    plt.title('ROC Curve')
    plt.legend(loc="lower right")
    plt.show()
```

```python
    return fpr,tpr,roc_auc

corr_df = data.corr()

# Generate a mask for the upper triangle
mask = np.zeros_like(corr_df, dtype=np.bool)
mask[np.triu_indices_from(mask)] = True

# Set up the matplotlib figure
f, ax = plt.subplots(figsize=(11, 9))

# Generate a custom diverging colormap
cmap = sns.diverging_palette(220, 10, as_cmap=True)

# Draw the heatmap with the mask and correct aspect ratio
sns.heatmap(corr_df, mask=mask, cmap=cmap, vmax=.3, center=0,
            square=True, linewidths=.5, cbar_kws={"shrink": .5})
plt.title('Correlations between different predictors')
plt.show()
```

Figure 5: Code to plot the correlation between various predictors

OUTPUT

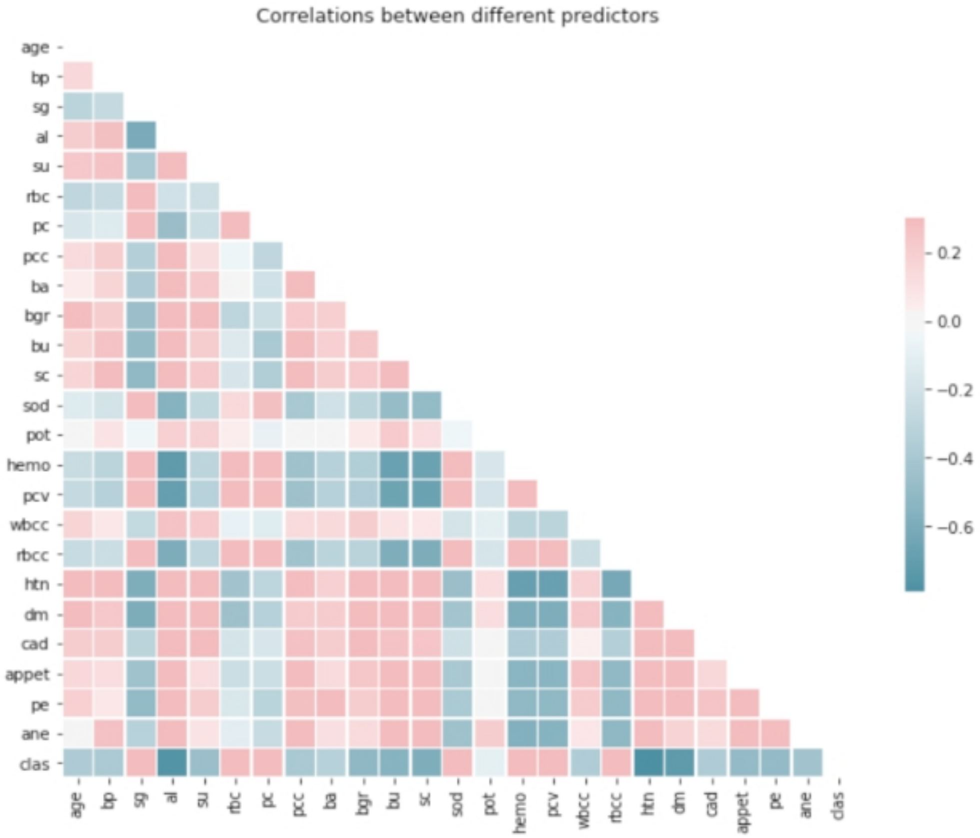


Figure 6: Correlation between various predictors

B) USING IMAGE PROCESSING AND DEEP LEARNING TO DETECT CHRONIC KIDNEY INFECTION

AVAILABLE DATA

The images that were part of our study, were scraped from Google Images, out of the five, three were labelled with the markings for the kidney stone. Here is an example:

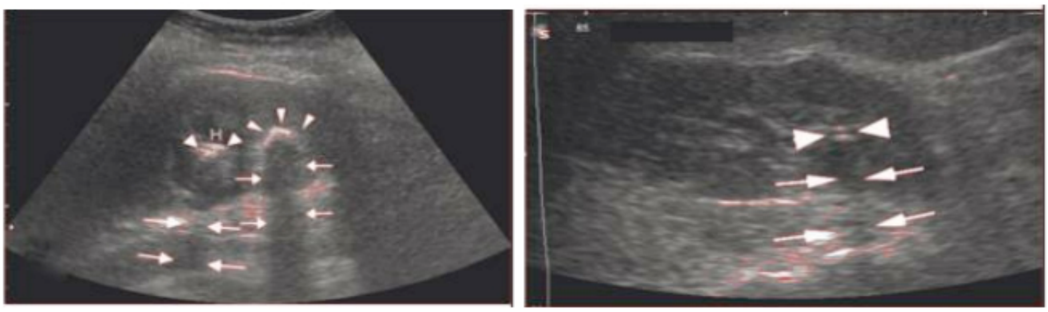
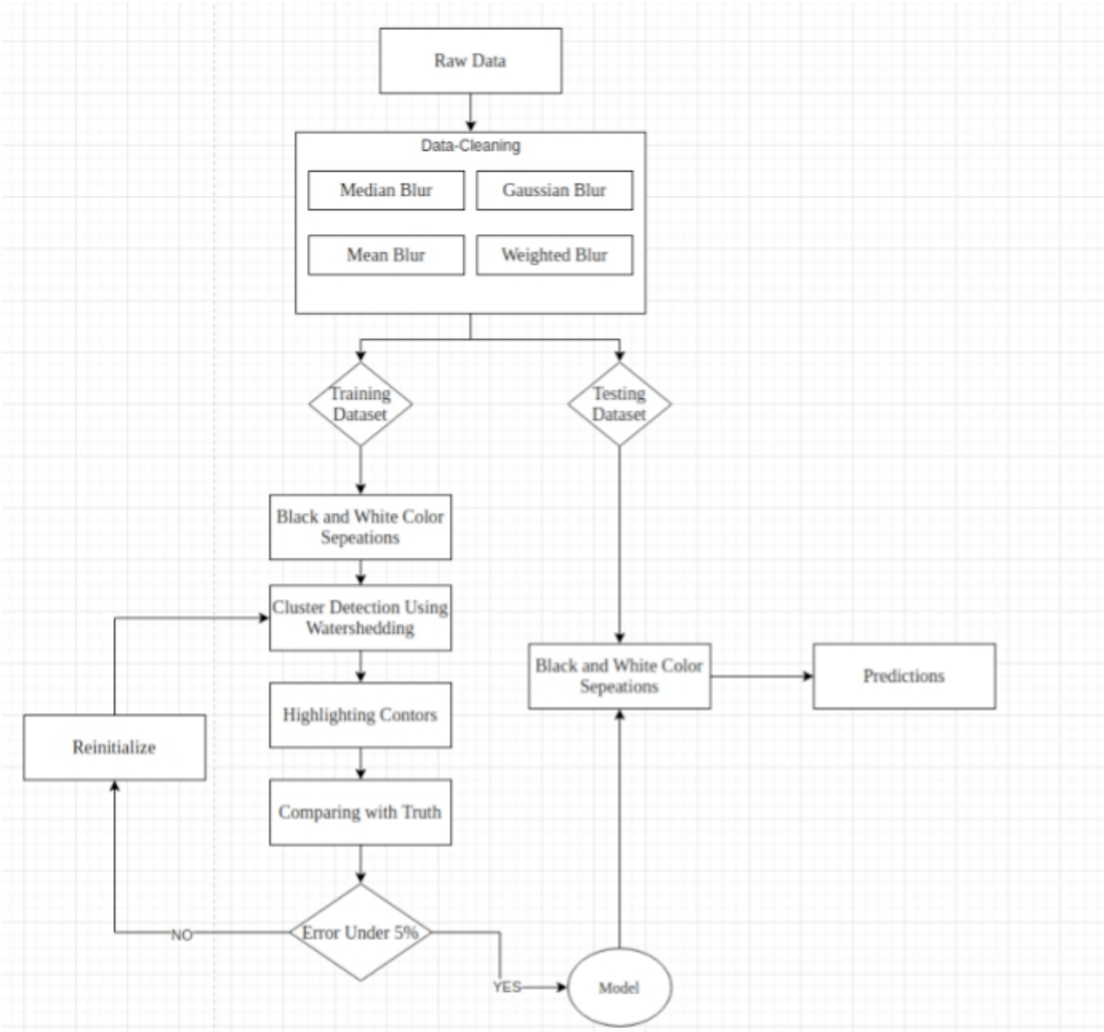


Figure 6 -The white markings were made for just highlighting the purpose

PROCESS DESCRIPTION



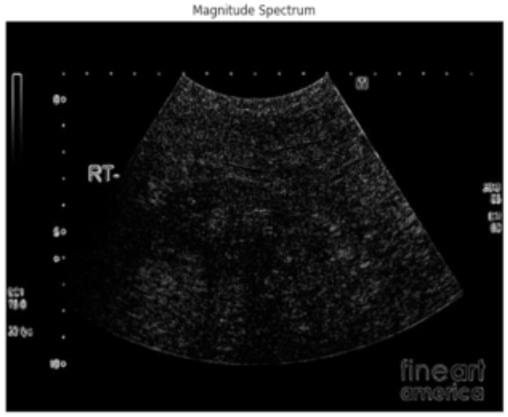
MEDIAN BLUR: Median filters are useful in reducing random noise, especially when the noise amplitude probability density has large tails and periodic patterns. The median filtering process is accomplished by sliding a window over the image. The filtered image is obtained by placing the median of the values in the input window, at the location of the centre of that window, at the output image.

GAUSSIAN BLUR: The Gaussian blur feature is obtained by blurring (smoothing) an image using a Gaussian function to reduce the noise level. It can be considered as a nonuniform low-pass filter that preserves low spatial frequency and reduces image noise and negligible details in an image.

AVERAGE BLUR: Blurs the Current Image by setting each pixel equal to the average pixel value of its specified box neighbourhood.

WEIGHTED AVERAGE BLUR: In the weighted average filter, we gave more weight to the centre value. Due to which the contribution of the centre becomes more than the rest of the values.

AFTER DATA CLEANING: After applying the specific blurs, we get the following output,

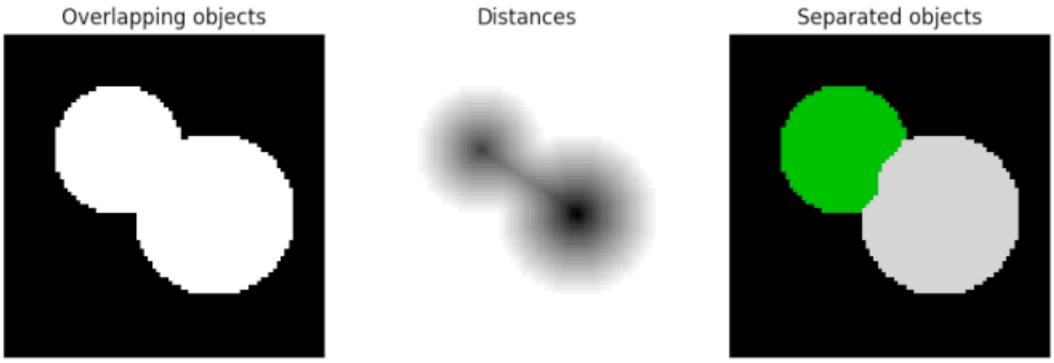


WATER SHEDDING

As we saw in the previous illustrations, kidney stones appear as small grey entities in the ultrasound. Hence, we need to separate the same from the background noise. This can be done with multiple techniques; we can use an algorithm called active contouring which is well known for segmenting elements from the background. But there is a problem here as the contours here, wouldn’t look as definitive as usual due to the fact that the background and the image of the kidney stones itself are too similar in colour, also even if we could differentiate there will still be the problem that the contours would we really close to each other, which makes them look like a single entity. To combat the closeness of the contours, we use a technique known as water shedding.

A watershed is a transformation defined on a grayscale image. The name refers metaphorically to a geological watershed, or drainage divide, which separates adjacent drainage basins. The watershed transformation treats the image it operates upon like a topographic map, with the brightness of each point representing its height, and finds the lines that run along the tops of ridges.

Water shedding can be visualized using the following diagram.



C) MATERIALS: IMPLANTABLE ARTIFICIAL KIDNEY

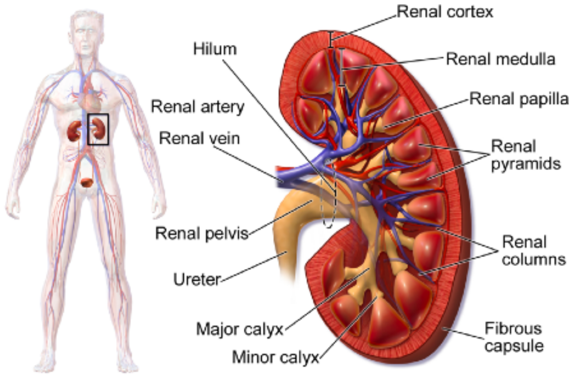
The implantable bioartificial kidney builds upon the existing extracorporeal Renal Assist Device (RAD), which is a bioartificial kidney that combines a membrane hemofilter and a bioreactor of human renal tubule cells to mimic many of the metabolic, endocrine, and immunological functions of a healthy kidney. While clinical trials confirmed that the RAD can safely treat acute renal failure in a critical care setting, adoption of the RAD for routine treatment of kidney failure patients is hampered by its labor-intensive and complex operation, large size, and high marginal cost. The ultimate goal of The Kidney Project is to apply microelectromechanical systems (MEMS) and nanotechnology to miniaturize the extracorporeal RAD into a surgically implantable, self-monitoring, and self-regulating bioartificial kidney. After a single surgery to establish a permanent blood connection, the bioartificial kidney processes blood continuously for 24 hours per day, which mitigates the inconveniences and morbidities associated with intermittent hemodialysis. There are several benefits to the implantable bioartificial kidney including the alleviation of the necessity of constant physician oversight and a heavy regimen of immunosuppressant drugs and medication.

MINIATURIZATION CHALLENGES

Early in our design process we identified two major obstacles to miniaturizing the RAD. size and pump requirements of modern dialysis units water volume needed for dialysis. At the same time, we were inspired by the success of hollow fiber polymer membranes used in treating renal failure extracorporeally. However, we noticed that the kidney’s natural fibers were uniform, elongated, slit-shaped structures, rather than the irregular and more cylindrically-shaped pores of polymer membranes. Our team has applied MEMS technology for the production of silicon nanopore membranes with slit-shaped pores that are tailored for implementation in a bioartificial kidney.

D) MODEL

In our model we have tried to replicate the design of an actual kidney .The size of the kidneys is measured mainly sonographically, although both CT and MRI scans also can be used to estimate renal size human kidney is 10-13 cm. In general, the left kidney is slightly



Kidney Anatomy

3. Results and Discussion

A) DETECTING CHRONIC KIDNEY DISEASE USING MACHINE LEARNING ALGORITHMS

For a given set of parameters of a patient, we try to predict if the patient is suffering from chronic kidney disease using the trained model

```
inp=[62,80,1.01,2,3,'normal','normal','notpresent','notpresent',423,53,1.8,111,2.5,9.6,31,7500,3.5,'no','yes','no

f_list = [] #converting to float
for item in inp:
    f_list.append(float(item))

print(f_list)


inp1=[f_list]

clf.predict(inp1)
lst_1=clf.predict(inp1).tolist()
print(lst_1)

if(lst_1[0])==0.0:
    print("Chronic kidney disease")
```

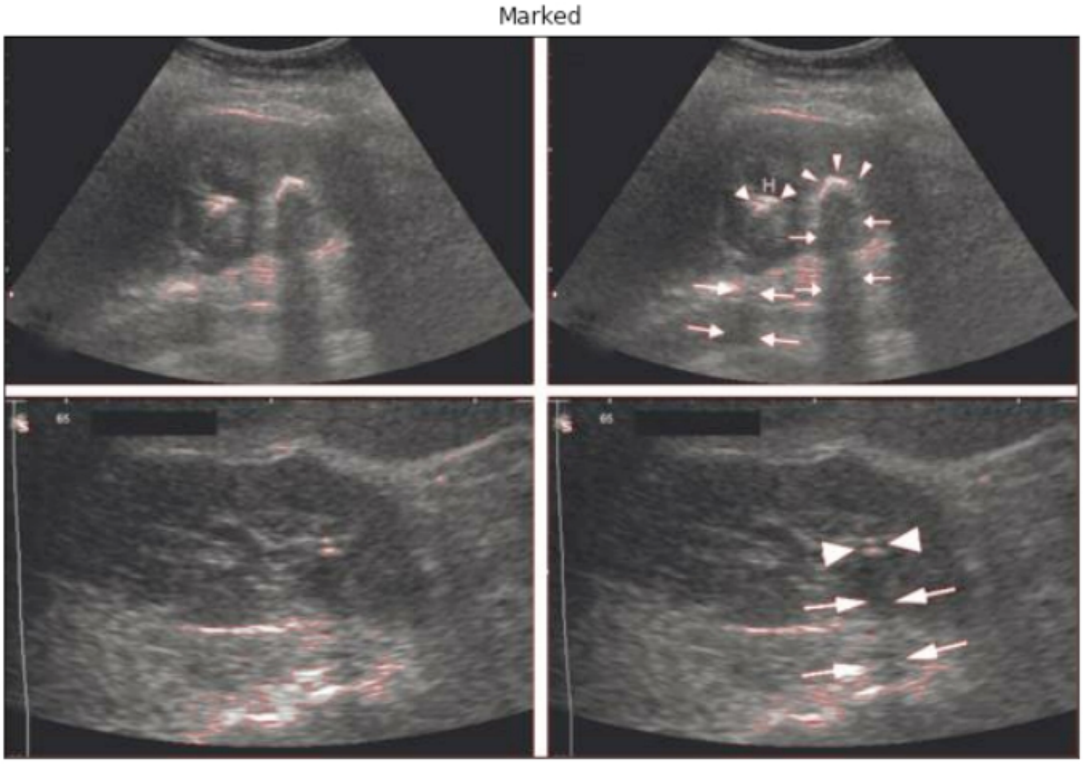
```
[62.0, 80.0, 1.01, 2.0, 3.0, 2.0, 2.0, 1.0, 1.0, 423.0, 53.0, 1.8, 111.0, 2.5, 9.6, 31.0, 7500.0, 3.5, 1.0, 2.0, 
[0.0]
Chronic kidney disease
```

Figure 7: Indicating the patient is suffering from chronic kidney disease

B) USING IMAGE PROCESSING AND DEEP LEARNING TO DETECT CHRONIC KIDNEY INFECTION

WATER SHEDDING

After applying water-shedding, train our neural network to learn to separate contours using water shedding. After a successful batch of training, this is the final result, as we can see the deep learning model was able to find the kidney stones, highlighted in red even with such a scarce amount of data.



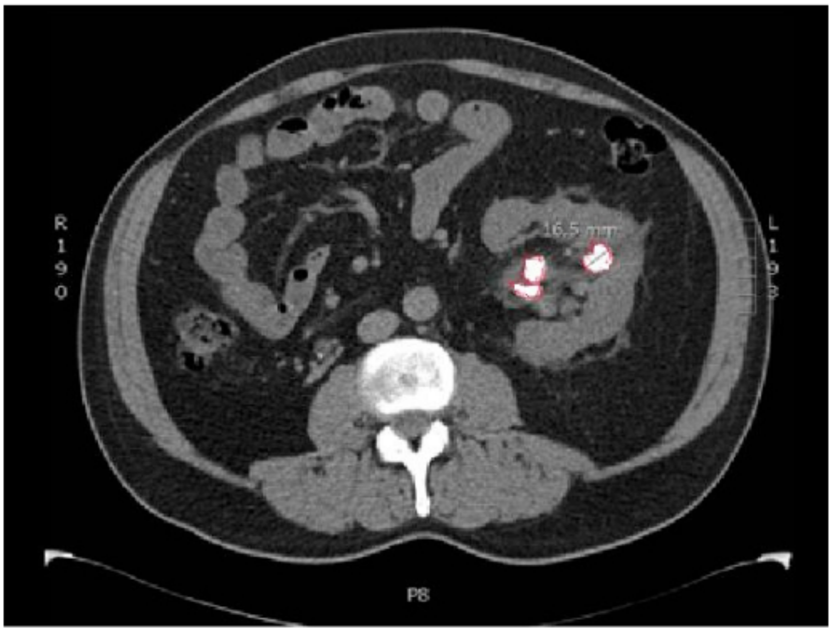
The confidence of our model predictions was about 87%.

CT SCAN IMAGES

After successful training of our model, we applied the same model to CT-images, the difference being the CT images are in a .gif format and are a continuous stream of images. Hence, we used OpenCV, to feed our model video data to get these final results. As the output was also in a .gif format we are adding some snapshots of our model predictions.

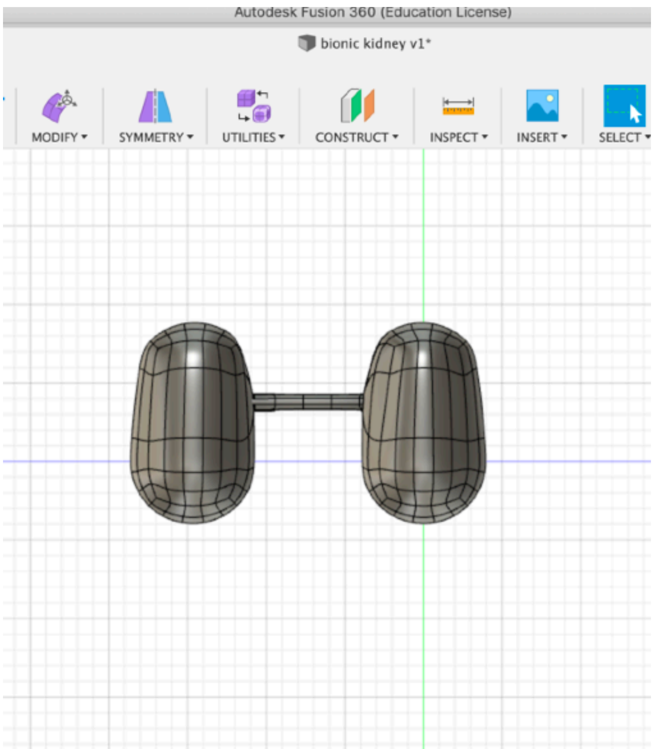
Input Image

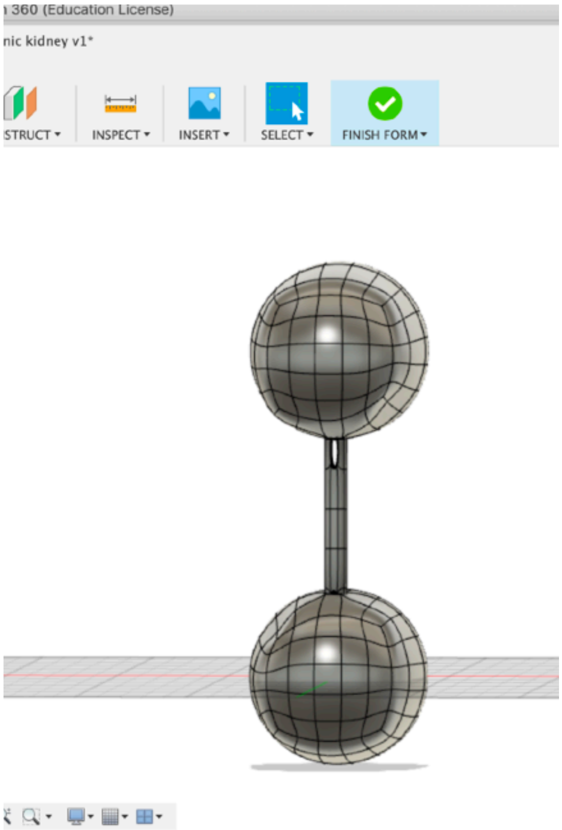
Output Image



Model Confidence: 73%, the model confidence is low because of the colour of the black bone.

C) MODEL





4. CONCLUSIONS

The novelty in our Approach

Due to the confidentiality of Medical Data, we were only able to find a small set of images for our inferences. Using a classical Deep Learning Approach i.e. training a model on a set of images and then applying the same model to testing data, wouldn't be feasible. We have devised a method that just requires only three images and still provides viable predictions. The method involves image processing techniques of contrast separation and the deep learning model learning contrast separations.Hospitals are often cost-sensitive and face complex operational problems specially when it comes to post care monitoring. So, came up with an idea where the patient can enter various parameters in the from the comfort of his house.These parameters are analysed using machine learning algorithms to predict if the prosthetic kidney is healthy or not.Hence, the patient is constantly monitored and can visit the hospital only if necessary. This project also reduces the pain of patients for waiting for a suitable match with the donor kidney blood group and hence it can be conferred that this idea is cost effective, patient and hospital friendly.

5. REFERENCES

1. Romagnani, Paola & Remuzzi, Giuseppe & Glassock, Richard & Levin, Adeera & Jager, Kitty & Tonelli, Marcello & Massy, Ziad & Wanner, Christoph & Anders, Hans-Joachim. (2017). Chronic kidney disease. *Nature Reviews Disease Primers* Vol. 3. 17088. 10.1038/nrdp.2017.88.

2. Chidambaranathan, Malathy & Mani, Gayathri. (2020). kidney stone detection with CT images using neural network. 10.37200/IJPR/V24I8/PR280269.

3. www.kaggle.com

4. https://rb.gy/nsvwuw

5. https://pharm.ucsf.edu/kidney