



**EDUCACIÓN**  
SECRETARÍA DE EDUCACIÓN PÚBLICA



**TECNOLÓGICO  
NACIONAL DE MÉXICO®**



# INSTITUTO TECNOLÓGICO DE LEÓN

División de Estudios de Posgrado e Investigación

“CLASIFICACIÓN DE VARIACIONES EN EL COLOR DE NEOPLASIAS  
UTILIZANDO LA TEXTURA”

TESIS

Que presenta:

ING. MARIELA CRUZ ÁNGELES

Para obtener el grado de:

MAESTRA EN CIENCIAS  
DE LA COMPUTACIÓN

Con la dirección de:

DR. RAÚL SANTIAGO MONTERO

Y Co Dirección de:

DR. ERNESTO BRIBIESCA CORREA

Revisores:

DRA. MARÍA DEL ROSARIO BALTAZAR FLORES  
DR. JUAN FRANCISCO MOSIÑO

León, Guanajuato.

Agosto 2021



Instituto Tecnológico de León

León, Guanajuato, 06/julio/2021

DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN  
OFICIO No. DEPI-101-2021

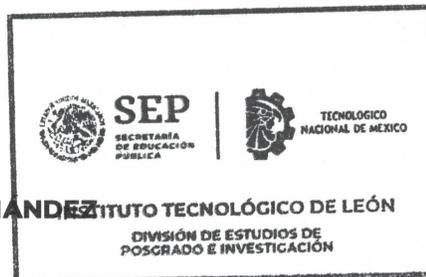
**ING. MARIELA CRUZ ÁNGELES  
PRESENTE**

De acuerdo al fallo emitido por la Comisión Revisora, integrada por los: Dr. Raúl Santiago Montero, Dr. Ernesto Bribiesca Correa, Dra. María del Rosario Baltazar Flores, Dr. Juan Francisco Mosiño, considerando que llena todos los requisitos establecidos en los Lineamientos Generales para la Operación del Posgrado del Tecnológico Nacional de México, se autoriza la impresión del trabajo de tesis titulado: "Clasificación de variaciones en el color de neoplasias utilizando la textura". Lo que hacemos de su conocimiento para los efectos y fines correspondientes.

**ATENTAMENTE**

Excelencia en Educación Tecnológica®  
Ciencia Tecnología y Libertad

**DR. DAVID ASAEL GUTIÉRREZ HERNÁNDEZ**  
**JEFE DE LA DEPI**



C.c.p. Expediente



RP/IL-072  
2017-04-10 - 2021-04-

Av. Tecnológico s/n Fracc. Industrial  
Julián de Obregón C.P 37290  
León, Gto. México Tel. 01 (477) 7105200,  
e-mail: tecleon@leon.tecnm.mx  
tecnm.mx | leon.tecnm.mx





León, Gto., a 9 de junio del 2021.

**C. ING. LUIS ROBERTO GALLEGOS MUÑOZ**  
**JEFE DE SERVICIOS ESCOLARES**  
**P R E S E N T E**

Por este medio hacemos de su conocimiento que la tesis titulada "**Clasificación de variaciones en el color de neoplasias utilizando la textura**", ha sido leída y aprobada por los miembros del Comité Tutorial para su evaluación por el jurado del acto de examen de grado al alumno (a) **C. Mariela Cruz Ángeles**, con número de control **M14200488** como parte de los requisitos para obtener el grado de Maestro(a) en Ciencias de la Computación (MCCOM-2011-05).

Sin otro particular por el momento, quedamos de Usted.

**ATENTAMENTE**  
**COMITÉ TUTORIAL**

  
Dr. Raúl Santiago Montero

**DIRECTOR**

  
Dr. Ernesto Bribiesca Correa

**CODIRECTOR ó REVISOR**

  
Dra. María del Rosario Baltazar Flores

Juan Francisco Mosiño

Dr. Juan Francisco Mosiño



## DECLARACION DE AUTENTICIDAD Y DE NO PLAGIO

Yo, Ing. Mariela Cruz Ángeles, identificada con No. Control M14200488, alumno (a) del programa de la **Maestría en Ciencias de la Computación**, autor (a) de la Tesis titulada: “Clasificación de variaciones en el color de neoplasias utilizando la textura” DECLARO QUE:

- 1.- El presente trabajo de investigación, tema de la tesis presentada para la obtención del título de **MAESTRO (A) EN CIENCIAS DE LA COMPUTACIÓN** es original, siendo resultado de mi trabajo personal, el cual no he copiado de otro trabajo de investigación, ni utilizado ideas, fórmulas, ni citas completas “stricto sensu”, así como ilustraciones, fotografías u otros materiales audiovisuales, obtenidas de cualquier tesis, obra, artículo, memoria, etc. en su versión digital o impresa.
- 2.- Declaro que el trabajo de investigación que pongo a consideración para evaluación no ha sido presentado anteriormente para obtener algún grado académico o título, ni ha sido publicado en sitio alguno.
- 3.- Declaro que las pruebas o experimentos derivados de esta investigación fueron realizados bajo el consentimiento de los involucrados y con fines estrictamente académicos conforme a criterios éticos de confidencialidad.

Soy consciente de que el hecho de no respetar los derechos de autor y hacer plagio, es objeto de sanciones universitarias y/o legales por lo que asumo cualquier responsabilidad que pudiera derivarse de irregularidades den la tesis, así como de los derechos sobre la obra presentada.

Asimismo, me hago responsable ante el Tecnológico Nacional de México/Instituto Tecnológico de León o terceros, de cualquier irregularidad o daño que pudiera ocasional por el incumplimiento de lo declarado.

De identificarse falsificación, plagio, fraude, o que el trabajo de investigación haya sido publicado anteriormente; asumo las consecuencias y sanciones que de mi acción se deriven, responsabilizándome por todas las cargas pecuniarias o legales que se deriven de ello sometiéndome a las normas establecidas en los Lineamientos y Disposiciones de la Operación de Estudios de Posgrado en el Tecnológico Nacional de México.

León, Guanajuato a 26 días del mes de julio de 2021.

Nombre y firma del autor

---

Ing. Mariela Cruz Ángeles

## ACUERDO PARA USO DE OBRA (TESIS DE GRADO)

A QUIEN CORRESPONDA

PRESENTE

Por medio del presente escrito, Ing. Mariela Cruz Ángeles, (en lo sucesivo el AUTOR) hace constar que es titular intelectual de la obra denominada: "Clasificación de variaciones en el color de neoplasias utilizando la textura", (en lo sucesivo la OBRA) en virtud de lo cual autoriza al Tecnológico Nacional de México/Instituto Tecnológico de León (en lo sucesivo TECN/IT León) para que efectúe resguardo físico y/o electrónico mediante copia digital o impresa para asegurar su disponibilidad, divulgación, comunicación pública, distribución, transmisión, reproducción, así como digitalización de la misma con fines académicos y sin fines de lucro como parte del Repositorio Institucional del TECN/ITLeón.

De igual manera, es deseo del AUTOR establecer que esta autorización es voluntaria y gratuita, y que de acuerdo a lo señalado en la Ley Federal del Derecho de Autor y la Ley de Propiedad Industrial el TECN/IT León cuenta con mi autorización para la utilización de la información antes señalada, estableciendo que se utilizará única y exclusivamente para los fines antes señalados. El AUTOR autoriza al TECN /IT León a utilizar la obra en los términos y condiciones aquí expresados, sin que ello implique se le conceda licencia o autorización alguna o algún tipo de derecho distinto al mencionada respecto a la "propiedad intelectual" de la misma OBRA; incluyendo todo tipo de derechos patrimoniales sobre obras y creaciones protegidas por derechos de autor y demás formas de propiedad intelectual reconocida o que lleguen a reconocer las leyes correspondientes. Al reutilizar, reproducir, transmitir y/o distribuir la OBRA se deberá reconocer y dar créditos de autoría de la obra intelectual en los términos especificados por el propio autor, y el no hacerlo implica el término de uso de esta licencia para los fines estipulados. Nada de esta licencia menoscaba o restringe los derechos patrimoniales y morales del AUTOR.

De la misma manera, se hace manifiesto que el contenido académico, literario, la edición y en general de cualquier parte de la OBRA son responsabilidad de AUTOR, por lo que se deslinda al (TECN/ITLeón) por cualquier violación a los derechos de autor y/o propiedad intelectual, así como cualquier responsabilidad relacionada con la misma frente a terceros. Finalmente, el AUTOR manifiesta que estará depositando la versión final de su documento de Tesis, OBRA, y cuenta con los derechos morales y patrimoniales correspondientes para otorgar la presente autorización de uso.

En la ciudad de León, del estado de Guanajuato, a los 26 veintiséis días del mes de julio de 2021.

Atentamente,

Nombre y firma autógrafa de EL AUTOR



---

Ing. Mariela Cruz Ángeles

# Dedicatoria

*Desde que tengo uso de razón, siempre quise hacer algo que pudiera beneficiar a muchas personas, soñaba con tener mucho dinero para ayudar a mi comunidad y a las personas que quería. Ahora que he crecido veo que la información científica es una manera de apoyar el trabajo de muchos investigadores y facilitar el desarrollo de sus proyectos, por eso quiero dedicar esta tesis a todas las personas que como yo, tengan el gusto por la investigación y las ganas de crear e innovar en favor de la comunidad. No te rindas, nunca sabes el número de vidas que la tuya puede impactar.*

# Agradecimientos

*Primeramente doy gracias a Dios por permitirme una vez más cumplir uno de mis sueños con éxito. En segundo lugar agradezco el apoyo incondicional de mi mamá Isa y mi hermanito Jesús, sin duda alguna, ustedes fueron el motor y apoyo principal para poder completar esta etapa de mi vida. Todo lo que soy es gracias a ustedes y para ustedes, los amo.*

*Agradezco también a todas esas personas que estuvieron ahí siempre dándome ánimos y apoyándome cada que lo necesité. Isis, Dany, Iván, Christian, gracias por compartir tantas aventuras, nuevas experiencias y conocimientos conmigo, sin duda alguna la maestría no hubiera sido lo mismo sin ustedes. Gus, Fede, gracias por ayudarme técnica y moralmente en este proceso. Sami, gracias por siempre estar, por apoyarme en mis decisiones, por acompañarme en un proceso tan importante como este y por creer en mí aún cuando ni yo lo hago.*

*Además agradezco a mi asesor y amigo, Dr. Raúl Santiago Montero por confiar en mí al darme la oportunidad de trabajar con él y motivarme a seguir investigando. A mi coasesor Dr. Ernesto Bribiesca, así como a mi comité evaluador, Dr. Juan Francisco Mosiño y Dra. María del Rosario Baltazar Flores, por sus importantes aportaciones para el desarrollo de esta tesis. Al Departamento de Estudios de Posgrado del TecNM-León y al Consejo Nacional de Ciencia y Tecnología por la beca otorgada en el periodo de estudios de maestría.*

GRACIAS.

# Índice general

Índice de tablas	VI
Índice de figuras	VII
Acrónimos	IX
<b>1. Introducción</b>	<b>1</b>
1.1. Problemática . . . . .	4
1.2. Justificación . . . . .	4
1.3. Hipótesis . . . . .	6
1.4. Objetivo General . . . . .	6
1.4.1. Objetivos Específicos . . . . .	6
1.5. Metodología . . . . .	7
1.6. Estructura de tesis . . . . .	9
<b>2. Marco teórico y antecedentes</b>	<b>11</b>
2.1. Cáncer de piel . . . . .	11
2.1.1. Melanoma . . . . .	11

2.1.2. Neoplasia . . . . .	12
2.2. Dermatoscopía . . . . .	13
2.2.1. Algoritmos de diagnóstico dermatoscópico . . . . .	14
2.2.2. Poder discriminante . . . . .	16
2.3. Imagen digital . . . . .	16
2.4. Modelos de color . . . . .	16
2.4.1. Modelo de color RGB . . . . .	17
2.5. Coeficientes de correlación 2D . . . . .	17
2.5.1. Textura . . . . .	18
2.5.2. Preprocesamiento de una imagen . . . . .	18
2.5.3. Procesamiento de una imagen . . . . .	19
2.5.4. Segmentación . . . . .	19
2.6. Extracción de características . . . . .	19
2.7. Reconocimiento de patrones . . . . .	20
2.8. Textura como descriptor . . . . .	21
2.9. Filtro de Gabor . . . . .	22
2.10. Aprendizaje automático . . . . .	22
2.10.1. Algoritmos de aprendizaje automático . . . . .	23
2.11. Cuantificación del color en neoplasias . . . . .	24
2.12. Textura como una característica . . . . .	27
2.13. Detección de melanoma a partir de imágenes dermatoscópicas . . . . .	27
2.14. Avances generales en la investigación . . . . .	28
2.15. Herramientas requeridas para el desarrollo del trabajo de investigación . .	31
2.15.1. GIMP . . . . .	31

2.15.2. ImageJ y WEKA . . . . .	32
2.15.3. Python . . . . .	32
2.15.4. OpenCV . . . . .	33
2.15.5. Características de la computadora . . . . .	35
<b>3. Descripción y experimentación</b>	<b>36</b>
3.1. Adquisición de instancias de prueba . . . . .	36
3.1.1. Base de datos PH2 . . . . .	36
3.1.2. Base de datos ISIC Archive . . . . .	38
3.2. Detección de la región de interés . . . . .	39
3.2.1. Segmentación manual . . . . .	39
3.2.2. Segmentación semi automática . . . . .	40
3.2.3. Segmentación automática . . . . .	41
3.2.4. Binarización . . . . .	43
3.2.5. Eliminación de ruido y recorte a partir de máscaras binarias . . . . .	44
3.3. Aplicación de filtro de Gabor . . . . .	46
3.4. Descripción . . . . .	47
<b>4. Análisis de resultados</b>	<b>49</b>
4.1. Análisis de resultados . . . . .	49
4.1.1. Coeficientes de correlación 2D . . . . .	49
4.1.2. Tiempo . . . . .	51
4.1.3. Aplicación de filtro de Gabor . . . . .	52
4.1.4. Generación de poblaciones . . . . .	55
4.1.5. Clasificador KNN . . . . .	56

<b>5. Conclusiones y trabajo futuro</b>	<b>57</b>
5.1. Conclusiones . . . . .	57
5.2. Recomendaciones y trabajo futuro . . . . .	58
<b>Referencias</b>	<b>59</b>
<b>A. Códigos programados</b>	<b>63</b>
<b>B. Gráficas de valores de intensidad clase Malignos</b>	<b>69</b>
<b>C. Gráficas de valores de intensidad clase Benignos</b>	<b>110</b>

# Índice de tablas

1. Acrónimos . . . . .	IX
2.1. Criterios dermatoscopicos . . . . .	15
2.2. Significado hispatológico de colores . . . . .	26
4.1. Promedio de coeficientes de correlación . . . . .	51
4.2. Promedio de tiempo de procesamiento en milisegundos . . . . .	52
4.3. Promedio de media y desviación estándar en las imágenes filtradas . . . . .	53
4.4. Valores ejemplo de desviación estándar . . . . .	54

# Índice de figuras

1.1. Metodología propuesta . . . . .	9
2.1. Sistema de Reconocimiento de Patrones, CITA . . . . .	20
2.2. Características usadas en el diagnóstico dermatoscópico . . . . .	30
2.3. Clasificadores usados en el diagnóstico dermatoscópico . . . . .	31
3.1. Base de datos PH2 . . . . .	38
3.2. Base de datos ISIC Archive . . . . .	39
3.3. Resultados de segmentación manual (a) Melanoma, (b) Nevo, (c) Benigno. . . . .	40
3.4. Algoritmo segmentación semi automática . . . . .	41
3.5. Algoritmo segmentación semi automática . . . . .	41
3.6. Etiquetado de clases en WEKA . . . . .	42
3.7. Características de entrenamiento y etiquetado de clases en WEKA . . . . .	43
3.8. Máscaras binarias segmentadas . . . . .	44
3.9. Proceso de eliminación de ruido en máscaras binarias . . . . .	45
3.10. Proceso de eliminación de ruido y recorte . . . . .	46
3.11. Ejemplo de Aplicación de un filtro de Gabor . . . . .	48
3.12. Ejemplo de imagen filtrada . . . . .	48

4.1. Ejemplo de dispersión de valores por clase . . . . .	55
4.2. Generación de poblaciones con una distribución normal . . . . .	56

# Acrónimos

La tabla 1 muestra el significado de los acrónimos utilizados en la presente investigación.

Tabla 1: Acrónimos

Acronimo	Significado	Español
ML	Machine Learning	Aprendizaje automático
DL	Deep Learning	Aprendizaje Profundo
RGB	Red Green Blue	Rojo Verde Azul
CMY	Cyan Magenta Yellow	Cian Magenta Amarillo
FIJI	Fiji Is Just Image	Fiji es solo imagen
WEKA	Waikato Environment for Knowledge Analysis	Entorno Waikato para el análisis del conocimiento
ROI	Region Of Interest	Región de Interés

# Capítulo 1

## Introducción

El melanoma es un tipo de cáncer de piel que se caracteriza por generar metástasis rápidamente cuando el tumor tiene contacto con la sangre. Las células sanguíneas contaminadas se esparcen en el cuerpo provocando la aparición de cáncer en otros órganos (Celebi y cols., 2007)(Goldstein y Goldstein, 2001). Si se detecta en etapas tempranas este tipo de neoplasia maligna, disminuye el riesgo de que afecte a otros órganos.

Dentro de los métodos de detección de melanoma más utilizados por los especialistas, sobresalen la regla de las 7 características y la regla ABCD. Estos procedimientos se soportan fundamentalmente en un análisis visual por parte del especialista (Zaballos, Carrera, Puig, y Malveyh, 2004). Sin embargo y de acuerdo con los datos proporcionados por Heinze-Martin, en el año 2018 se reportaba una tasa de 0.31 de médicos especialistas en México por cada 100,000 habitantes con una especialidad en oncología médica (Heinze-Martin, Olmedo-Canchola, Bazán-Miranda, Bernard-Fuentes, y Guízar-Sánchez, 2018) y no hay estadísticas de la de la proporción de especialistas en cáncer de piel. Este hecho o afecta claramente una detección temprana de melanoma.

La falta de especialistas en el área no es privativo de México, sino un fenómeno global. Como consecuencia se han desarrollado múltiples técnicas que ayuden a la detección de los diversos cánceres de piel, en sus orígenes se hacía uso de aceite sobre la piel y una lupa para poder examinar el área de pigmentación. Al paso de los años se ha implementado el uso de las herramientas tecnológicas para visualizar mejor las lesiones, de aquí nace el uso de la dermatoscopia; los dermatoscopios permiten el aumento de 10 veces de la piel (Peñafiel, 2014) el uso de luz integrada y un lente que pueda aumentar la visualización hace que el especialista tenga más herramientas de mejora en su trabajo de diagnóstico. Piccolo (Piccolo y cols., 2002) informa que médicos con 5 años de experiencia en detección de cáncer, con el uso del dermatoscopio, aumentaron de un 71 % a un 90 % de sensibilidad y especificidad en el diagnóstico. Además, de acuerdo con estudios realizados por Vestergaard (Vestergaard, Macaskill, Holt, y Menzies, 2008), la dermatoscopia es más precisa que el examen ocular sin más herramientas para el diagnóstico de melanoma cutáneo cuando se realiza en el entorno clínico.

La dermatoscopia permite visualizar con una imagen amplificada las características de la región de interés (ROI, por sus siglas en inglés), a partir de esa imagen se implementa un procesamiento digital para hacer un prediagnóstico de malignidad automático de melanoma. La mayoría de esos procesos automáticos se ven en la necesidad de realizar una segmentación de la ROI, es decir, separar la zona de piel no pigmentada de la sí pigmentada, para hacer un análisis de características, con esto realizan mediciones automáticas que le permitirán al especialista dar un diagnóstico más cuantitativo y menos cualitativo de la lesión (Cruz Mahecha, Vargas Martínez, y cols., 2018)(Rastgoo, Garcia, Morel, y Marzani, 2015)(Zaballos y cols., 2004). El examen dermatoscópico permite observar diversos colores que ayudan a elaborar el diagnóstico dermatológico y en ocasiones juega un papel

decisivo al brindar una orientación en el diagnóstico diferencial entre una lesión maligna y una indolente (Rodríguez-Cabral, Montoya-Sosa, Roldán-Marín, y Carlos-Ortega, 2014).

Describir la diferencia de color en neoplasias como un factor de presencia de cáncer es esencial para diagnósticos tempranos, los descriptores de textura pueden ser utilizados como medidas cuantitativas de los cambios de intensidad de los valores RGB dentro de una región en imágenes digitales.

En la presente investigación se propuso el análisis de la característica de color presente en imágenes que contienen neoplasias benignas y malignas, para cuantificar este descriptor primeramente se procedió a segmentar semi-automáticamente cada una de las imágenes de la base de datos, posteriormente se aplicó un filtro de Gabor con la finalidad de intensificar la textura presente, una vez aplicado el filtro se calcularon las medidas de dispersión y tendencia central de los valores de los píxeles que componen la región de interés, y con esos valores se generó un dataset que finalmente se utilizó para entrenar un clasificador KNN, se probó con validación cruzada a 8 pliegues, los resultados de clasificación entre neoplasias benignas y neoplasias malignas arrojaron un 71 % de efectividad, mostrando así que usando la textura como descriptor en un clasificador sí es posible diferenciar las clases.

**Aportaciones principales de esta tesis:** Como resultado de este proyecto de investigación se tiene: primero, que sí es posible cuantificar el color, y segundo, que esa cuantificación por sí sola da resultados por arriba del 70 % en la clasificación de neoplasias malignas y benignas.

## 1.1. Problemática

La regla ABCD es una de las más utilizadas para detección de malignidad en una neoplasia (Pathan, Prabhu, y Siddalingaswamy, 2018), por ello es una de las más implementadas en los algoritmos de inteligencia artificial. Diversas investigaciones (Nachbar y cols., 1994)(Johr, 2002)(Smaoui y Bessassi, 2013)(Ng, Fung, y Lee, 2005) reportan el uso de combinaciones de descriptores de asimetría, borde, color y diámetro para potenciar el poder discriminante al momento de evaluar la malignidad una neoplasia. El color ha sido comunmente trabajado con la cuantificación de la asimetría de color (Healsmith, Bourke, Osborne, y Graham-Brown, 1994), sin embargo (Celebi y cols., 2007) obtuvo un conjunto de 108 características calculando la desviación estándar y media de 6 espacios de color, para determinar la diferencia de color relativa entre la periferia interna, la periferia externa y la lesión, con lo cual se muestran indicios de una posible diferencia también entre clases de neoplasias (benignas y malignas).

La propuesta de investigación es cuantificar el color a través de un descriptor de textura como lo es el filtro de Gabor, con la finalidad de encontrar el ángulo que determine los cambios de tonalidad en la imagen. La problemática a resolver, pone en cuestión si la textura como descriptor de los cambios de tonalidad de color puede ser utilizada en procesos de clasificación de neoplasias.

## 1.2. Justificación

De acuerdo con la Universidad Nacional Autónoma de México en estudios realizados por la División de Estudios de Posgrado, Facultad de Medicina en la Ciudad de México,

el número de especialistas es inferior al recomendado internacionalmente e insuficiente para cubrir las necesidades en salud del país. En 2017 el número de habitantes en México ascendía a 123,518,272, y 1,784 de esos habitantes eran médicos que contaban con especialidad en dermatología y solo el 82.1 % de ellos contaban con su certificación vigente. La tasa de médicos especialistas en oncología médica en México por cada 100,000 habitantes es de 0.31. (Heinze-Martin y cols., 2018) La información que se tiene demuestra que no hay suficientes especialistas en dermatología como para dar tratamiento a todas las personas que sufren de cáncer de piel, el acceso a este tipo de servicios es solo para zonas urbanas por lo que no es accesible a personas de bajos recursos o ubicadas geográficamente en zonas rurales. Principalmente lo que se busca con el desarrollo de este proyecto es tener un sistema confiable y accesible que pueda ser de utilidad para la población en general y así detectar la malignidad de neoplasias en etapas tempranas. El melanoma maligno es uno de los cánceres en el mundo, que aumenta más rápidamente, con una incidencia estimada de 59.580% y un total estimado de 7770 muertes en los Estados Unidos solo en 2005. El diagnóstico temprano del melanoma es particularmente importante, ya que con una simple escisión se extrae, si se detecta temprano (Celebi y cols., 2007). El poder diseñar un clasificador de neoplasias permitirá realizar pronósticos masivos y confiables día a día, así se podrán filtrar los casos malignos de los benignos, para que las personas que realmente lo necesiten sean atendidas a tiempo y así contribuir a la reducción de la tasa de mortalidad en el país por este tipo de cáncer, para ello se propone en etapas a futuro el desarrollo de un dermatoscopio que pueda ser utilizado en cualquier dependencia de salud para hacer pronosticación de malignidad de cáncer de piel. A pesar de que se ha demostrado que la variación de color es uno de los indicadores más importantes para la detección de cánceres de piel, no ha sido tan común el uso de éste en el desarrollo

de dermatoscopios, sin embargo, en este proyecto se pretende utilizar la textura como descriptor primario para hacer la clasificación.

### **1.3. Hipótesis**

Si existe un mayor grado de dispersión de los componentes RGB, de una región digital que contiene la imagen de una neoplasia maligna en comparación con una imagen de una neoplasia benigna, entonces será mayor el poder discriminante para determinar si la neoplasia es maligna.

### **1.4. Objetivo General**

Evaluar la capacidad discriminante utilizando la variación de los componentes RGB a través de descriptores de textura dentro de una base de datos de imágenes digitales que contienen neoplasias y que pueden ser benignas o malignas.

#### **1.4.1. Objetivos Específicos**

- Analizar la variación de los componentes RGB y textura en imágenes que contienen neoplasias malignas y benignas.
- Generar, diseñar e implementar experimentos para medir el grado de discriminación que se tiene en un clasificador que use como descriptor a la textura.
- Implementar los módulos de procesamiento, descripción, clasificación y validación estadística de los datos obtenidos en la etapa de experimentación.

## 1.5. Metodología

La metodología establecida para el presente proyecto de investigación se basa en un modelo experimental y cuantitativo que, incluye las etapas que a continuación se describen.

1. **Estado del arte y marco teórico:** Se realizó una búsqueda de información en artículos relacionados con el tema de investigación propuesto, además de consultar el significado de los conceptos básicos de la investigación para generar el estado del arte y el marco conceptual.
2. **Adquisición de las instancias de prueba:** Con base en el estado del arte se analizó qué bases de datos presentaban las características necesarias para el entrenamiento del clasificador, es decir una base de datos etiquetada y comúnmente trabajada para entrenar un clasificador de neoplasias, además las bases de datos a trabajar debían permitir un equilibrio de clases, es decir el mismo número de elementos para la clase maligna y benigna.
3. **Detección de región de interés:** Se propusieron 3 métodos, el primero consiste en realizar una segmentación manual para aislar la zona de piel pigmentada de la no pigmentada, esto para tener una segmentación ideal que sirviera como referencia para realizar la segmentación de manera automática. El segundo método consiste en utilizar un conjunto de clasificadores para aislar los pixeles que forman parte de la región de interés de manera automática. El tercer método propuesto se basa en una segmentación semi-automática de la ROI, utilizando filtros y operaciones morfológicas en el procesamiento de las imágenes.

Para elegir el método a utilizar se obtuvieron coeficientes de correlación entre la imagen segmentada manualmente y las segmentadas semi-automática y automáticamente. La segmentación que tuviera un promedio más cercano a 1, sería la elegida.

4. **Descripción y Experimentación:** El objetivo de esta etapa es potenciar la textura como descriptor a utilizar, aplicando filtros de Gabor. Se propuso un experimento de variación de los parámetros de tamaño de máscara y ángulo del filtro de Gabor, se probaron las configuraciones en las bases de datos y se obtuvieron promedios de las medidas de dispersión y tendencia central de los valores de los píxeles de cada una de las clases una vez aplicado el filtro. Para elegir la mejor configuración del filtro de Gabor en este experimento, se analizó qué configuración determinaba una diferencia mayor entre clases de sus valores estadísticos.
5. **Análisis de datos:** Se generaron tablas que muestran los resultados de los experimentos realizados, en cada caso se analizó el comportamiento por clase de la característica que se estaba evaluando, este análisis permitió generar las conclusiones y evaluar la hipótesis planteada.

En la Figura. 1.1 se muestran gráficamente las etapas de la metodología a desarrollar.



Figura 1.1: Metodología propuesta

## 1.6. Estructura de tesis

1. **Introducción:** En este capítulo se muestra el problema a atacar, la hipótesis planteada y los objetivos a cumplir, así como la metodología y la descripción de la misma para poder desarrollar el proyecto de investigación.
2. **Marco teórico y antecedentes:** En este capítulo se describen los conceptos importantes involucrados en esta tesis, también se muestra el estado del arte de la presente investigación.
3. **Descripción y Experimentación:** Aquí se puede ver el desarrollo de las etapas de adquisición de instancias de prueba, segmentación de la región de interés y la descripción de la textura. Además de un análisis de la segmentación que se eligió.
4. **Análisis de resultados:** Se reportan los resultados obtenidos en cada experimento realizado, a través de tablas e imágenes que ayudan a visualizar a profundidad los resultados.

5. **Conclusiones y trabajo futuro:** Finalmente se dan a conocer las conclusiones del trabajo, así como las posibles actividades que se podrían realizar a futuro como parte de la continuidad del presente proyecto de investigación.

# Capítulo 2

## Marco teórico y antecedentes

### 2.1. Cáncer de piel

La piel es el órgano más extenso del cuerpo humano, lo protege y es por ello que puede ser afectada por un gran número de patologías, entre ellas el cáncer; éste puede definirse como la multiplicación rápida, desordenada y agresiva de células anormales, principalmente de la capa más superficial de la piel, la dermis (Palencia, s.f.). El cáncer de piel se revela generalmente con la generación de un tumor y según sea su crecimiento, hacia el interior o el exterior de la dermis, representa un grado diferente de amenaza.

#### 2.1.1. Melanoma

El melanoma es un tumor derivado de la proliferación de melanocitos atípicos con o sin la capacidad de producir pigmento, caracterizado por su gran capacidad de dar metástasis. Se presenta principalmente en la piel, pero se pueden presentar en otros órganos (Goldstein y Goldstein, 2001). El melanoma maligno es un tipo de cáncer de piel que se caracteriza

por su gran agresividad y resistencia al tratamiento.

En las últimas décadas, la comunidad de investigadores ha propuesto varios algoritmos computarizados de análisis de lesiones para la detección de melanoma. Estos algoritmos se enfocan principalmente en generar hiperplanos para diferenciar el melanoma de las lesiones benignas (Rastgoo y cols., 2015). Una vez que los algoritmos son capaces de detectar la presencia de una neoplasia maligna se implementan en dermatoscopios para su uso no invasivo en pacientes que presentan pigmentaciones atípicas en la piel.

El melanoma, el más letal de todos los cánceres de piel, se origina en los melanocitos, las células que dan el color a la piel. Al igual que todos los tipos de cáncer de piel, se caracteriza por ser una enfermedad silenciosa, usualmente indolora que puede originarse en piel sana o a partir de un lunar u otra lesión ya existente que va cambiando su forma, tamaño, color o textura. (Araujo Moncayo, Capera Tole, Plazas Montaña, y cols., s.f.)

La incidencia del cáncer de piel ha aumentado a nivel mundial desde 1979, sobre todo el de tipo melanoma. La exposición prolongada a la radiación solar y a su componente ultravioleta sin ninguna precaución es un factor de gran influencia en la mayoría de los casos, incluso se ha visto un incremento de incidencias en personas jóvenes debido al uso de dispositivos de bronceado artificial y a la falta de protección al exponerse al sol. (Cruz Mahecha y cols., 2018)

### **2.1.2. Neoplasia**

De acuerdo con el diccionario de cáncer que proporciona el Instituto Nacional de Cáncer de los Institutos Nacionales de la Salud de EE. UU., una neoplasia es definida como la masa anormal de tejido que aparece cuando las células se multiplican más de

lo debido o no se destruyen en el momento apropiado. Las neoplasias son benignas (no cancerosas) o malignas (cancerosas).

## 2.2. Dermatoscopía

A pesar de que las lesiones en la piel son visibles directamente, existen muchas estructuras de la misma que no lo son. Con el avance de la tecnología y su aplicación en la medicina, hoy en día se hace uso del dermatoscopio para observar con más detalle la lesión pigmentada. La dermatoscopía o Microscopia de Epiluminiscencia (ELM), es una técnica que mediante un sistema de amplificación e iluminación permite eliminar el reflejo y la reflexión de la superficie de la piel, permitiendo una mejor visualización en vivo de la estructura y sub estructura de las lesiones pigmentadas, cuando se usa por dermatólogos expertos, permite reconocer características morfológicas que no son visibles a simple vista.

Estudios indican que la dermatoscopía logra una mejoría del 20-25 % en la precisión diagnóstica del melanoma, implicando así, una reducción de biopsias innecesarias y una mejor oportunidad para el seguimiento y documentación de las lesiones existentes. Los instrumentos actualmente usados para la realización de una exploración dermatoscópica son el dermatoscopio manual, el estereomicroscopio quirúrgico binocular y portátil, el videodermatoscopio y los equipos de dermatoscopía digital. El dermatoscopio usualmente utilizado tiene un aumento de 10x que es suficiente para una evaluación de las lesiones pigmentadas de la piel en la rutina diaria. (Zaballos y cols., 2004)

### 2.2.1. Algoritmos de diagnóstico dermatoscópico

Existen diferentes métodos que permiten distinguir parámetros y patrones que se asocian con neoplasias malignas, como lo son la regla ABCD, control de 7 puntos y el análisis de patrones. La regla de dermatoscopia ABCD se basa en los criterios de asimetría (A), borde (B), color (C) y estructura diferencial (D), que presente la neoplasia, ésta regla mejoró la precisión diagnóstica cuando se aplicó retrospectivamente a diapositivas clínicas (Nachbar y cols., 1994). La regla ABCD se puede aprender fácilmente y calcular rápidamente, y ha demostrado ser confiable.

El control de 7 puntos es un método de diagnóstico ideado para ayudar a la clasificación de melanomas por examinadores menos experimentados, requiere la identificación de siete criterios dermatológicos jerarquizados. A las características más frecuentemente relacionadas con el melanoma, llamadas criterios mayores, se les asigna un puntaje de 2, por otro lado, a las características menos asociadas al melanoma, llamadas criterios menores, se les asigna un puntaje de 1. Mediante la simple adición de las puntuaciones individuales, una puntuación total de 3 o más permite la clasificación de melanoma con una sensibilidad del 95 % y una especificidad del 75 %. (Zaballos y cols., 2004). Las características se muestran en la tabla número 2.1

Tabla 2.1: Criterios dermatoscópicos

Criterio dermatoscópico	Puntuación máxima
CRITERIOS MAYORES	
Retículo pigmentado atípico	2
Velo azul-blanquecino	2
Patrón vascular atípico	2
CRITERIOS MENORES	
Proyecciones irregulares	1
Puntos/Glóbulos irregulares	1
Manchas de pigmento irregulares	1
Estructuras asociadas a la regresión	1
PUNTUACIÓN TOTAL	
Menor a 3	Benigna
Mayor o igual a 3	Maligna

El análisis de patrones es la técnica que médicos con más experiencia y conocimiento usan. Las lesiones melanocíticas se caracterizan por tener unos pocos patrones globales que cubren la mayor parte de la lesión, permitiendo una categorización preliminar rápida, y unos patrones locales, estructuras individuales presentes en distintas regiones de la lesión; sin embargo, se pueden observar características locales adicionales que representan pistas de diagnóstico útiles. Además, se pueden presentar características de confusión, es decir, criterios dermoscópicos que con poca frecuencia están presentes dentro de una categoría de diagnóstico dando lugar a una clasificación equivocada (Zaballos y cols., 2004).

### **2.2.2. Poder discriminante**

Es un término derivado de la bioquímica y la interpretación de resultados de los estudios clínicos como parte de la teoría semiológica, refiere a la habilidad del ser humano para distinguir entre un resultado maligno o benigno para poder diagnosticar enfermedades. (Homatidis y Konstantareas, 1981)

## **2.3. Imagen digital**

Una imagen se define como una función bidimensional,  $f(x, y)$ , donde  $x$  e  $y$  son coordenadas espaciales en un plano, y la amplitud de  $F$  en cualquier conjunto de coordenadas  $(x, y)$ , es la intensidad o nivel de gris de la imagen en ese punto. Cuando  $X$ ,  $Y$  y el valor de la amplitud de  $F$  son todas finitas y cantidades discretas, llamamos a la imagen una imagen digital. (Gonzalez, Woods, y cols., 2002)

Una imagen digital es una matriz, cuyas filas y columnas identifican un punto de la imagen y el valor de ese elemento de la matriz indica el nivel de gris de ese punto. Una imagen digital está compuesta de un número finito de elementos, los cuales cada uno tiene su localización particular y valor, dichos elementos son llamados píxel. El píxel es el término más usado para denotar los elementos de una imagen digital. (Gonzalo, de la CruzJesus, y cols., 2008)

## **2.4. Modelos de color**

Los modelos de color especifican los colores de una forma normalizada y aceptada genéricamente. Un modelo de color expresa la cantidad máxima de datos de color, que se

pueden almacenar en un determinado formato de archivo gráfico. Su propósito es el de facilitar la especificación de los colores utilizando algún estándar. Los modelos de color están orientados al hardware. Los modelos de color son: RGB (rojo, verde, azul) para monitores y cámaras, CMY (cyan, magenta, amarillo) para impresoras a color.

### **2.4.1. Modelo de color RGB**

Trabaja con tres canales Rojo, Verde y Azul, construida por 8 bits/píxel por canal (24 bits en total). Es un modelo de color aditivo, es decir, cuanto más rojo, verde y azul se agregue, más se parecerá el color al blanco. El valor de cada canal (rojo, verde, o azul) puede ir desde 0 (sin color) hasta 255 (color con la máxima saturación).

## **2.5. Coeficientes de correlación 2D**

El coeficiente de Correlación en 2D se ha utilizado para detectar similitudes entre señales que están en 2D, que comúnmente se guardan en matrices, éstas representan dos imágenes del mismo tamaño, por lo que se puede decir que ésta métrica sirve para medir el rendimiento para la calidad de imagen (Ramadan, 2017). Para éste trabajo se utilizará el coeficiente de correlación para medir el rendimiento de cada imagen segmentada manualmente con su respectiva segmentación automática y semi automática obtenida de cada uno de los clasificadores. Una vez dados los elementos de entrada se hace el cálculo del coeficiente de acuerdo con la ecuación 1, donde:  $r$  es el coeficiente de correlación,  $A$  y  $B$  son las matrices a comparar,  $m$  y  $n$  son los índices en  $x,y$  de las matrices,  $\bar{A}$  es el promedio de ambas dimensiones de la matriz  $A$ , y  $\bar{B}$  es el promedio de las dos dimensiones de la matriz  $B$ . (Ramadan, 2017).

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (B_{mn} - \bar{B})^2)}} \quad (2.1)$$

### 2.5.1. Textura

Textura es una propiedad que representa la superficie o la estructura de un objeto, desde un punto de vista de textura en imágenes se puede definir textura como algo que consiste en elementos mutuamente relacionados; por lo tanto, se puede considerar un grupo de píxeles y la textura descrita dependerá en gran medida del número considerado de píxeles. La textura consiste en primitivas de textura o elementos de textura, a veces llamados texels. (Sonka, Hlavac, y Boyle, 2014)

Textura es la variación de datos en escalas más pequeñas que la escala de interés. Es importante por dos razones principales: puede representar ruido en un sistema de visión automática y la textura puede ser una señal importante en el reconocimiento de objetos, ya que nos dice algo sobre el material del que está hecho el objeto (Petrou y García, 2006).

### 2.5.2. Preprocesamiento de una imagen

Preprocesamiento es el nombre utilizado para operaciones en imágenes en el nivel más bajo de abstracción, tanto la entrada como la salida son imágenes de intensidad. Estas imágenes icónicas suelen ser del mismo tipo que los datos originales capturados por el sensor, con una imagen de intensidad generalmente representada por una matriz o matrices de valores de función de imagen. El preprocesamiento no aumenta el contenido de información de la imagen. (Sonka y cols., 2014)

Realizar un preprocesamiento en imágenes para su acondicionamiento y la extracción

de características requiere un esquema de trabajo donde los elementos que afecten la obtención de la región de interés y su posterior procesamiento sean removidos de la imagen. (Martínez, Funes, Rosales-Silva, y Ramos-Arredondo, 2016)

### **2.5.3. Procesamiento de una imagen**

El campo del procesamiento de imágenes digitales se refiere al procesamiento de imágenes digitales por medio de una computadora digital. Tenga en cuenta que una imagen digital está compuesta por un número finito de elementos, cada uno de los cuales tiene una ubicación y un valor particular, comúnmente llamados píxeles. (Gonzalez y cols., 2002)

### **2.5.4. Segmentación**

La segmentación busca dividir una imagen en un subconjunto de píxeles, con la finalidad de que cada subimagen contenga cada uno de los objetos representados en la imagen. Existen casos en los que los objetos se solapan entre ellos y el proceso de segmentación es más complejo. El objetivo de la etapa de segmentación de las imágenes dermatoscópicas utilizadas en esta investigación es el separar la zona de piel pigmentada de la no pigmentada, llamando así a la zona con coloración la región de interés.

## **2.6. Extracción de características**

Comunmente después de haber realizado un preprocesamiento a una imagen y de haber segmentado la región de interés, se procede a la etapa de extracción de características. Un extractor de rasgos o características aplica operadores sobre una imagen segmentada o no segmentada, para identificar posiciones de rasgos que ayudarán a la formación de

hipótesis sobre la presencia de un objeto dado. Entre más características se tengan del objeto será mejor su representación. Sossa, (Azuela, 2013) expone que para que un rasgo descriptor sea útil para el reconocimiento de un objeto, éste debe poseer un conjunto de propiedades deseables como: poder de discriminación, fiabilidad, incorrelación, rapidez y economía.

## 2.7. Reconocimiento de patrones

El objetivo del reconocimiento de patrones es construir sistemas que puedan clasificar los objetos de interés del mundo real en una de varias clases en función de mediciones; el reconocimiento de patrones es el estudio de cómo las máquinas pueden ir observando el ambiente para aprender a distinguir patrones de interés de un fondo y realizar decisiones razonables sobre las categorías de los mismos.

Graficamente, el proceso de reconocimiento automático se puede observar en la figura, 2.1



Figura 2.1: Sistema de Reconocimiento de Patrones, CITA

Tal como se aprecia en la figura 2.1, una vez que se extraen los rasgos, ahora la representación del objeto está compuesta por un patrón de características únicas del objeto y corresponden al banco de descripciones que identifican al objeto de entre otros.

## 2.8. Textura como descriptor

Textura es la variación de datos en escalas pequeñas dentro de una imagen, permite distinguir los objetos que hay al rededor, por lo que no puede ser ignorada por ningún sistema automático de visión. (Petrou y García, 2006)

Varios autores han intentado definir cualitativamente la textura. Pickett (Pickett, 1970) afirma que "la textura se utiliza para describir variaciones de matrices bidimensionales; Los elementos y las reglas de espaciado o disposición pueden manipularse arbitrariamente, siempre que permanezca una repetitividad característica". Hawkins (Hawkins, 1970) ha proporcionado una descripción más detallada de la textura: "La noción de textura parece depender de tres ingredientes: (1) algún 'orden' local se repite en una región que es grande en comparación con el tamaño del orden, (2) el orden consiste en la disposición no aleatoria de partes elementales y (3) las partes son entidades aproximadamente uniformes que tienen aproximadamente las mismas dimensiones en todas partes dentro de la región texturizada".

La textura es una característica fundamental presente en las imágenes, juega un papel muy importante tanto en visión por computadora como en reconocimiento de patrones. El análisis de la textura en las imágenes ha sido ampliamente trabajado en diferentes áreas de investigación, así mismo su análisis ha sido clasificado en tres categorías: estadístico, basado en modelos, basado en filtrado. (Pérez Carrasco, Serrano Gotarredona, y Acha Piñero, 2009)

## 2.9. Filtro de Gabor

Los filtros de Gabor han sido empleados en multitud de aplicaciones de procesamiento de imágenes, entre ellas clasificación y segmentación de texturas, reconocimiento de imágenes y objetos.

Se han utilizado filtros de Gabor combinado con el uso de redes neuronales para la clasificación de lesiones en la piel, o para segmentar una imagen en diferentes regiones de texturas similares basadas en una especificada (Rangel Cortes, 2019)

Aplicando filtros de Gabor en diferentes escalas y orientaciones, los patrones de textura pueden ser eficientemente descritos en el dominio frecuencial y localizados en el dominio espacial. (Pérez Carrasco y cols., 2009)

Un filtro de Gabor bidimensional es una red sinusoidal de campo complejo que está modulada por una función gaussiana bidimensional en el dominio espacial (Pérez Carrasco y cols., 2009). Los filtros Gabor tienen orientación sintonizable y bandas de paso de frecuencia radial y frecuencias centrales sintonizables.

## 2.10. Aprendizaje automático

El aprendizaje se refiere a un amplio espectro de situaciones en las cuales el aprendiz incrementa su conocimiento o sus habilidades para cumplir una tarea. El aprendizaje aplica inferencias a determinada información para construir una representación apropiada de algún aspecto relevante de la realidad o de algún proceso. (Samuel, 1959)

El aprendizaje automático, es una rama de la Inteligencia Artificial que se encarga de generar algoritmos que tienen la capacidad de aprender y no tener que programarlos de

manera explícita. El desarrollador no tendrá que sentarse a programar por horas tomando en cuenta todos los escenarios posibles ni todas las excepciones posibles. Lo único que hay que hacer es alimentar el algoritmo con un volumen gigantesco de datos para que el algoritmo aprenda y sepa qué hacer en cada uno de estos casos. (Pereira, Mitchell, y Botvinick, 2009)

Hay dos tipos de aprendizajes: el supervisado y el no supervisado. El primer tipo de aprendizaje se refiere a cuando cuando se entrena un algoritmo de Machine Learning dándole las preguntas (características) y las respuestas (etiquetas). Así en un futuro el algoritmo pueda hacer una predicción conociendo las características. El aprendizaje no supervisado se espera que el algoritmo diga a qué grupo pertenece el elemento en estudio. El algoritmo encuentra patrones en los datos que se le proporcionan y los clasifica en grupos. Luego compara los nuevos datos y los ubica en uno de los grupos y es así como puede predecir de que se trata. (Villalonga Gutiérrez y cols., 2019)

### **2.10.1. Algoritmos de aprendizaje automático**

Un clasificador es una función que toma los valores de ejemplo (predictores o variables independientes) en varias características para predecir la clase a la que pertenece (la variable dependiente). Los clasificadores de aprendizaje automático (ML), se usan popularmente en la predicción de patrones, dependiendo del conjunto de datos disponible. (Pereira y cols., 2009)

El aprendizaje supervisado tiene el objetivo de inferir una función o mapeo de los datos de entrenamiento que están etiquetados. Los datos de entrenamiento consisten en el vector de entrada  $x$  y el vector de salida  $y$  de etiquetas. Los datos etiquetados manualmente son

un recurso vital para el aprendizaje supervisado (Mohammed, Khan, y Bashier, 2016).

## 2.11. Cuantificación del color en neoplasias

Los principios de la dermatoscopia basaban sus investigaciones en ciertas características que se presentaban dentro de una neoplasia para determinar su malignidad, de aquí se derivan: simetría en el borde, forma y transición de cuatro colores (blanco, café claro, café oscuro y gris-azul), estructuras y áreas homogéneas, las cuales suman puntos para determinar su malignidad. (Iyatomi, 2010)

Pietro (Rubegni y cols., 2002) muestra como resultado de sus diversas investigaciones relacionadas con lesiones de piel que, una vez digitalizado y parametrizado las diversas características presentes en las imágenes de lesiones de piel, algunas características poseen mayor poder discriminante que otras para determinar su malignidad. Las características más importantes son categorizadas de la siguiente manera: geometría, color, textura e islas de color (grupos de color dentro de la lesión). Dentro de la geometría se considera el área, el máximo y mínimo de diámetro, radio, varianza de simetría de contorno, circularidad, fractibilidad de bordes y elipsoidalidad. La variación de colores como rojo, verde y azul dentro de la lesión. Las variables de textura fueron el contraste medio y la entropía de la lesión, así como el contraste y la entropía fractal. Las islas de colores fueron presentes principalmente en las regiones más oscuras de la región, compuestas de colores verde, rojo, azul-gris y sus diferentes combinaciones y transiciones.

Una imagen en color se representa como una matriz de píxeles, donde cada píxel contiene componentes numéricos (generalmente compuesto de tres elementos) que define un color según un espacio de color. Existe una variedad de espacios de color hechos para

diferentes aplicaciones (como visualización, impresión, coincidencia de color y retransmisiones televisivas). Una vez que se especifica el espacio de color, las características de color pueden extraerse de imágenes o regiones de imágenes.

En 2016, a través del survey “Incorporating Colour Information for Computer-Aided Diagnosis of Melanoma from Dermoscopy Images: A Retrospective Survey and Critical Analysis”, (Madooei y Drew, 2016) se presentó una evaluación de las investigaciones que se han llevado a cabo en los últimos años y que se han basado en la información que proporciona el color en melanomas. Algunos de los estudios más relevantes muestran lo siguiente:

Algunos estudios utilizaron valores de píxeles directamente como características de color, mientras que la mayoría de los otros emplearon momentos de color e histogramas de color. Estas son características primitivas o de bajo nivel que, por ejemplo, parametrizan la distribución del valor de color en una imagen. Una tendencia creciente en los últimos años es el desarrollo de características visuales de alto nivel (clínicamente significativas) como la asimetría del color, la variación del color, la clasificación del color y la detección del color.

Entre otros estudios, hay aquellos que se centraron en la “cuantización del color” y aquellos destinados a la “segmentación del color”. La cuantificación del color tiene como objetivo reducir el número de colores. Las imágenes en color verdadero generalmente contienen miles de colores, lo que hace que su visualización, almacenamiento, transmisión y procesamiento sean problemáticos. Por esta razón, la cuantización del color se usa comúnmente como un paso de preprocesamiento para diversas tareas de procesamiento de gráficos e imágenes. La mayoría de los métodos de cuantificación son esencialmente basados en algoritmos de agrupamiento de datos. (Madooei y Drew, 2016)

En el artículo “Principios básicos de dermatoscopia”, entre otras cuestiones, se analizó la importancia del color en las neoplasias, concluyendo así que: “a mayor número de colores más sospecha de malignidad”. (Rodríguez-Cabral y cols., 2014) Se da a conocer una tabla con los colores que puede presentar la neoplasia y su significado histopatológico, tal como se muestra en la tabla 2.2

Tabla 2.2: Significado hispatológico de colores

Color	Significado histopatológico
Negro	Se debe al depósito de melanina en el estrato córneo y epidermis superficial. De igual manera, observamos negro cuando la melanina está presente en grandes cantidades en varias capas.
Marrón (claro y oscuro)	Generalmente se debe a que la localización del pigmento melánico se encuentra en la unión dermoepidérmica.
Gris y azul-gris	Se debe al pigmento melánico o melanófagos en la dermis superficial.
Azul acero	Lo observamos cuando la pigmentación se localiza a mayor profundidad, dermis reticular, porque las porciones de luz visible con longitud de onda más corta (espectro azul-violeta) están más dispersas.
Rojo	Se asocia con incremento en el número o la dilatación de vasos sanguíneos por inflamación, traumatismo o neovascularización.
Blanco	Suele observarse en casos de regresión, cicatrización o ambas. Debido a que la regresión puede observarse en nevos y en melanoma, las lesiones que muestren una combinación de tonalidades azul blanquecino, blanco-grisáceo o ambas de preferencia deben ser extirpadas y estudiadas en términos histopatológicos, sobre todo en casos con regresión mayor a 10 % de la superficie de la lesión.
Amarillo	Corresponde a material sebáceo o a zonas de hiperqueratosis.
Naranja	Son zonas de erosión o ulceración que pueden abarcar toda la lesión o sólo un segmento.

## 2.12. Textura como una característica

En 2015 se publicó el artículo (Rastgoo y cols., 2015) “Automatic differentiation of melanoma from dysplastic nevi”, en él se muestran los resultados de una investigación que propone un framework automático de diferenciación entre melanoma y nevos displásicos (que son 2 tipos de neoplasias diferentes), durante la investigación se utilizó la regla clínica “ABCD”, la evaluación reveló el potencial de la textura para la diferenciación de melanoma y nevos displásicos, el framework alcanzó la mayor sensibilidad de 98 %.

La variación de color y el histograma de color es una característica conocida que se ha utilizado ampliamente en el pasado para la detección de las características del melanoma. Para esta investigación el descriptor contiene, la media y varianza de nueve canales de color (R, G, B H, S, V L, A, B) e histograma de canales R, G y B. Cada histograma para cada canal está construido con 42 contenedores, lo que lleva a un tamaño de descriptor final de  $(9 \times 2) + (42 \times 3) = 144$ . (Rastgoo y cols., 2015)

## 2.13. Detección de melanoma a partir de imágenes dermatoscópicas

En la tesis (Cruz Mahecha y cols., 2018) “Detección de Melanomas a partir de imágenes dermatoscópicas”, desarrollada en la Universidad Distrital Francisco José de Caldas en Bogotá, Colombia, se presentó un modelo basado en la segmentación semántica capaz de alcanzar la precisión de un dermatólogo experto en un 92.96 % pero que tarda en promedio 416.5 segundos en ejecutarse, en comparación, se presentó un modelo basado en la segmentación por umbral óptimo capaz de alcanzar una precisión del 90.63 % de un

dermatólogo experto y cuyo tiempo de ejecución se reduce a 28.98 segundos.

Las pruebas se realizaron sobre una computadora con sistema operativo Windows 10, 16Gb de memoria RAM y 2Tb de disco duro. En este trabajo los puntos de interés contenían información tanto de color como de forma, características representativas para la clasificación de lesiones pigmentadas en la piel. (Cruz Mahecha y cols., 2018)

## 2.14. Avances generales en la investigación

En esta sección se dan a conocer los avances generales en las áreas de estudio relacionadas con la presente investigación, se presentan los avances en las áreas de: preprocesamiento de imágenes, segmentación de lesiones de piel, extracción de características, selección de características y clasificación.

- **Preprocesamiento de imágenes:** para facilitar la detección de bordes en las imágenes de lesiones de piel se han estado aplicando diferentes técnicas de filtrado, conversiones de color, detección y remoción de bello y reducción de ruido. El canal azul del modelo de color RGB es el que más se ha usado en el preprocesamiento, debido a que es el canal que predomina en las lesiones, aunque también el uso de la escala de grises ha sido efectivo en algunos casos. La variación de iluminación en la captura de las imágenes también ha sido un aspecto importante a considerar y una cuestión a mejorar en esta etapa. (Pathan y cols., 2018)
- **Segmentación de lesiones de piel:** La segmentación alude a la división de una imagen en distintas regiones que contienen píxeles con atributos afines. Las técnicas de segmentación se clasifican como técnicas de segmentación de bajo y alto nivel pa-

ra obtener una perspectiva con respecto a los enfoques de segmentación de lesiones. Las técnicas de segmentación de bajo nivel son enfoques convencionales que incluyen métodos que son computacionalmente más simples, más rápidos, y requieren procesamiento posterior. Técnicas de segmentación de alto nivel incluyen a aquellos enfoques que integran la segmentación de bajo nivel para construir sofisticados algoritmos de segmentación y evitan posprocesamiento. (Mendonca y cols., 2007) (Celebi y cols., 2007) (Garnavi y cols., 2010)

- **Extracción de características:** La extracción de características dermatoscópicas ha sido usada con diferentes enfoques, ya que depende del tipo de regla de análisis que se use para conocer las características a tomar en cuenta. La regla más usada es la ABCD, consistente en el análisis de 4 criterios de Asimetría, Borde, Color y Diferencia estructural. El método Menzies, la regla de los 7 puntos, el algoritmo CASH, características de forma, textura y color, entre otros métodos se han empleado para la extracción de características. De acuerdo con valores estadísticos reportados por Sameena (Pathan y cols., 2018) la característica más usada es la del color, tal como se muestra en la gráfica 2.2.

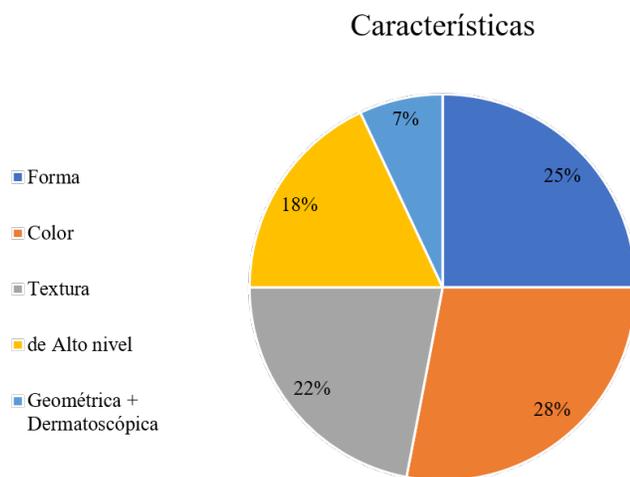


Figura 2.2: Características usadas en el diagnóstico dermatoscópico

- **Selección de características:** La selección de características es un punto importante, puesto que así se aumenta el poder discriminante, además que elimina la redundancia, la irrelevancia y simplifica el cálculo cuando se emplea para clasificar. La optimalidad del subconjunto de características se cuantifica mediante un criterio de evaluación. Es común auxiliarse de algoritmos evolutivos, redes neuronales, genéticos, o cualquier otro algoritmo de optimización que ayude a la evaluación de las características y muestre cuáles de ellas presentan un poder discriminante mayor.
- **Clasificación:** Diversos clasificadores se han empleado, siendo el más común el ANN, en la gráfica 2.3 se muestran los clasificadores más usados de acuerdo con la información proporcionada por Sameena (Pathan y cols., 2018)

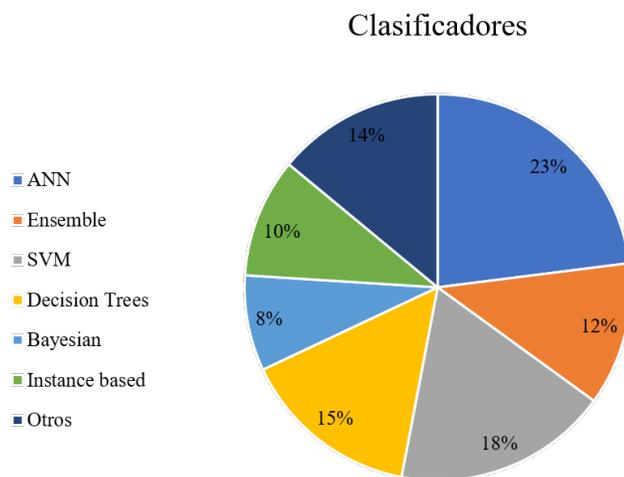


Figura 2.3: Clasificadores usados en el diagnóstico dermatoscópico

## 2.15. Herramientas requeridas para el desarrollo del trabajo de investigación

### 2.15.1. GIMP

GIMP (GNU Image Manipulation Program) es un editor de imágenes multiplataforma disponible para GNU / Linux, OS X, Windows y más sistemas operativos. Es software gratuito, puede cambiar su código fuente y distribuir sus cambios. Se utilizó para el procedimiento de segmentación manual. El Selector varita mágica que proporciona GIMP permite hacer una selección por regiones circundantes, se utilizó para hacer la selección por los colores presentes en las imágenes; además su herramienta Selector de colores de fondo y primer plano permitió dejar la imagen con un fondo color blanco.

### **2.15.2. ImageJ y WEKA**

Para realizar la segmentación semi automática se utilizó ImageJ, el cual es un programa de código abierto escrito en Java que permite añadir extensiones de software para añadir funcionalidades a el procesamiento de las imágenes, una de las más conocidas es WEKA. (Rueden y cols., 2017).

Trainable WEKA Segmentation (TWS), es una herramienta de aprendizaje automático comúnmente utilizada para entrenar clasificadores y segmentar automáticamente. Para la segmentación de la imagen TWS transforma el problema de segmentación en un problema de clasificación de píxeles. (Arganda-Carreras y cols., 2017). En este trabajo ésta herramienta se usó para pobrar algoritmos supervisados de clasificación para aislar la región de interés.

### **2.15.3. Python**

A finales de 1980, Guido Van Rossum trabajaba en Los Países Bajos en CWI (Centrum voor Wiskunde en Informatica) en el equipo del sistema operativo Amoeba, en 1989 creó la primera versión del lenguaje Python, llamada así en honor a los Monty Python de los que era fan. Python está inspirado en los lenguajes ABC (este último está inspirado en Algol y está diseñado para tener éxito en BASIC, Pascal y Awk). (Chazallet, 2016) Este lenguaje de programación tiene diversas ventajas con respecto a otros lenguajes, a continuación, se enumeran algunos principios enunciados por Guido van Rossum:

1. Simple e intuitivo.
2. Open source, todo el mundo puede contribuir.
3. Comprensible, tan sencillo de entender como leer inglés.
4. Apropiado para las tareas diarias.
5. Permite tiempos de desarrollo cortos.

#### **2.15.4. OpenCV**

OpenCV (Open Source Computer Vision Library) es una biblioteca de software de visión artificial y aprendizaje automático de código abierto. OpenCV se creó para proporcionar una infraestructura común para aplicaciones de visión por computadora y para acelerar el uso de la percepción de la máquina en los productos comerciales. Al ser un producto con licencia BSD, OpenCV facilita que las empresas utilicen y modifiquen el código. (Bradski y Kaehler, 2008)

La biblioteca tiene más de 2500 algoritmos optimizados, que incluyen un conjunto completo de algoritmos de aprendizaje automático y visión por computadora clásicos y de última generación. Estos algoritmos se pueden utilizar para detectar y reconocer rostros, identificar objetos, clasificar acciones humanas en videos, rastrear los movimientos de la cámara, rastrear objetos en movimiento, extraer modelos 3D de objetos, producir nubes de puntos 3D a partir de cámaras estéreo, unir imágenes para producir una alta resolución. imagen de una escena completa, buscar imágenes similares de una base de datos de imágenes, eliminar ojos rojos de imágenes tomadas con flash, seguir los movimientos oculares, reconocer paisajes y establecer marcadores para superponerlos con realidad au-

mentada, etc. OpenCV tiene más de 47 mil personas de usuarios comunidad y un número estimado de descargas superior a 18 millones. La biblioteca se utiliza ampliamente en empresas, grupos de investigación y organismos gubernamentales.

Junto con empresas bien establecidas como Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota que emplean la biblioteca, hay muchas nuevas empresas como Applied Minds, VideoSurf y Zeitera, que hacen un uso extensivo de OpenCV. Los usos desplegados de OpenCV abarcan desde unir imágenes de streetview juntas, detectar intrusiones en videos de vigilancia en Israel, monitorear equipos de minas en China, ayudar a los robots a navegar y recoger objetos en Willow Garage, detectar accidentes por ahogamiento en piscinas en Europa, ejecutar arte interactivo en España y Nueva York, revisando las pistas de aterrizaje en busca de escombros en Turquía, inspeccionando las etiquetas de los productos en las fábricas de todo el mundo hasta la detección rápida de rostros en Japón.

Tiene interfaces C ++, Python, Java y MATLAB y es compatible con Windows, Linux, Android y Mac OS. OpenCV se inclina principalmente hacia aplicaciones de visión en tiempo real y aprovecha las instrucciones MMX y SSE cuando están disponibles. Actualmente se están desarrollando activamente interfaces CUDA y OpenCL con todas las funciones. Hay más de 500 algoritmos y aproximadamente 10 veces más funciones que componen o admiten esos algoritmos. OpenCV está escrito de forma nativa en C ++ y tiene una interfaz con plantilla que funciona a la perfección con los contenedores STL.

### **2.15.5. Características de la computadora**

Todo el trabajo que se presenta en éste artículo fue realizado en una computadora con las siguientes características: sistema operativo Windows 10 Home de 64 bits, procesador Intel (R) Core (TM) i5-7200U CPU 2.50GHz, memoria RAM de 4.00 GB

# Capítulo 3

## Descripción y experimentación

### 3.1. Adquisición de instancias de prueba

Las bases de datos utilizadas en el desarrollo de la presente investigación se describen a continuación. En términos generales se tomaron 200 imágenes de la clase malignos de la base de datos PH2 y 200 imágenes de la clase benignos de la base de datos ISIC Archive.

#### 3.1.1. Base de datos PH2

La creación de la base de datos PH2 fue posible gracias a una colaboración conjunta entre la Universidade do Porto y la Universidade de Lisboa en conjunto con el servicio de Dermatología del Hospital Pedro Hispano en Matosinhos, Portugal. La base de datos PH2 se creó principalmente para poner a disposición un conjunto de datos común que puede utilizarse para la evaluación del rendimiento de diferentes sistemas de diagnóstico asistido por computadora de imágenes dermatoscópicas. (Mendonca, Celebi, Mendonca, y Marques, 2015)

La base de datos consta actualmente de 200 imágenes dermatoscópicas junto con las anotaciones médicas correspondientes, que comprenden 80 nevos comunes, 80 nevos atípicos y 40 melanomas malignos. Las imágenes dermatoscópicas se adquirieron cuidadosamente utilizando un aumento de  $20 \times$  en condiciones sin cambios. Son imágenes en color RGB de 8 bits con una resolución de  $768 \times 560$  píxeles. El conjunto de imágenes disponibles en la base de datos PH2 se seleccionó con algunas restricciones, con respecto a su calidad, resolución y características dermatoscópicas, de modo que sean lo suficientemente adecuadas como para usarse en una base de datos de referencia dermatoscópica. (Mendonca y cols., 2015)

Las características dermatoscópicas disponibles en la base de datos PH2 son: asimetría, colores, red de pigmentos, puntos, glóbulos, rayas, áreas de regresión y velo azul blanquecino. Estas características dermatoscópicas se seleccionaron por dos razones principales. Primero, este conjunto de criterios dermatoscópicos corresponde a aquellas características que se usan comúnmente por dermatólogos para realizar un diagnóstico clínico. Además, este conjunto de características comprende la mayoría de las características dermatoscópicas que deben evaluarse en los procedimientos de diagnóstico médico más utilizados. (Mendonca y cols., 2015) En la Figura. 3.1 se muestran algunas de las imágenes que forman parte de la base de datos PH2.

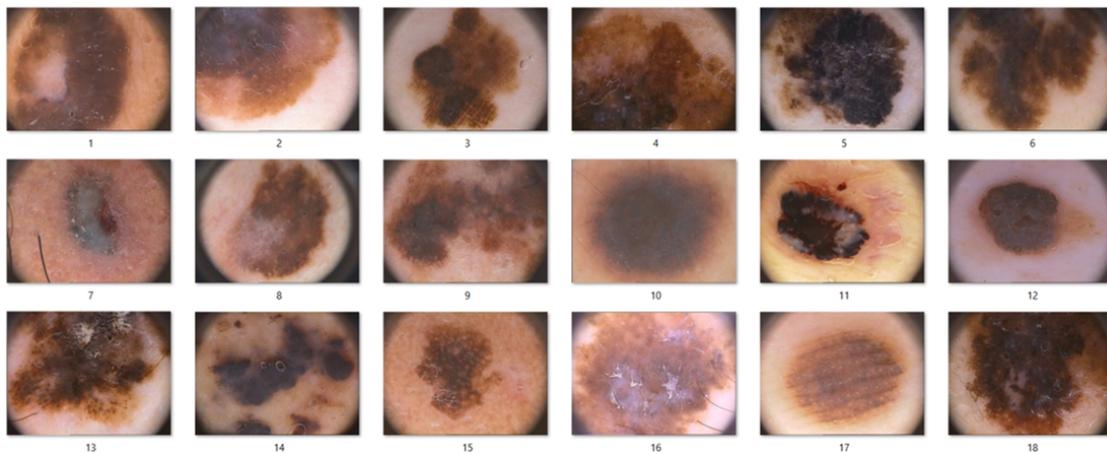


Figura 3.1: Base de datos PH2

### 3.1.2. Base de datos ISIC Archive

The International Skin Imaging Collaboration: Melanoma Project es una asociación entre la academia y la industria diseñada para facilitar la obtención de imágenes digitales de la piel para ayudar a reducir la mortalidad por melanoma. ISIC cuenta con un sitio web que permite tener acceso público a un banco de imágenes de la piel para probar y validar los estándares propuestos para la detección de melanoma. Este archivo sirve como recurso público de imágenes para la enseñanza y para el desarrollo y prueba de sistemas de diagnóstico automatizados. (Lopez, Giro-i Nieto, Burdick, y Marques, 2017) En la figura 3.2 se muestran algunas de las imágenes que forman parte de la base de datos ISIC Archive, cabe mencionar que son imágenes de neoplasias benignas las que se muestran, sin embargo el sitio web contiene neoplasias tanto benignas como malignas.

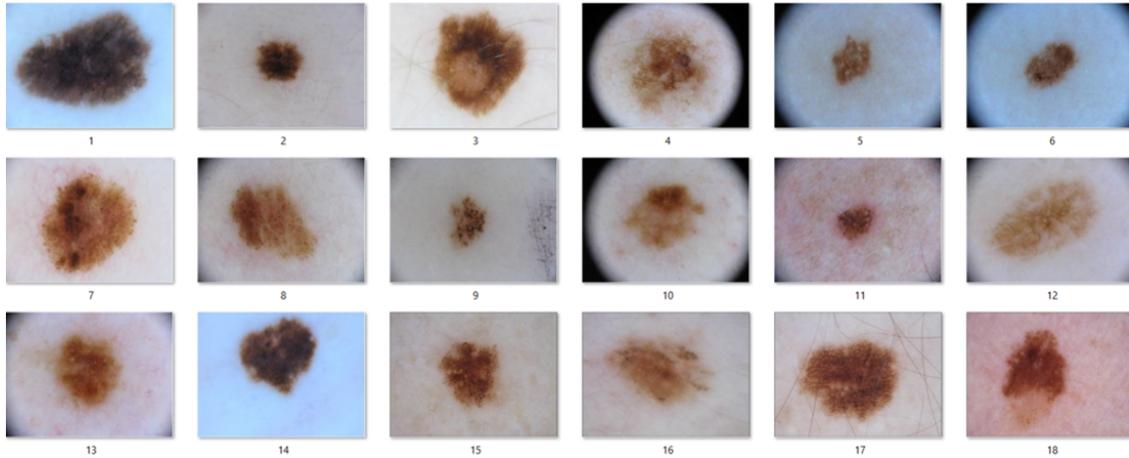


Figura 3.2: Base de datos ISIC Archive

## 3.2. Detección de la región de interés

Para llevar a cabo la etapa de detección de región de interés, se probaron dos tipos de segmentación, los procedimientos y resultados de segmentación se describen en las siguientes subsecciones.

### 3.2.1. Segmentación manual

El primer método a implementar es una segmentación manual utilizando el software GIMP, aquí simplemente se fue haciendo uso de las herramientas de edición de imágenes para dejar las imágenes con la ROI y un fondo de imagen en color blanco, tal como se puede observar en la figura 3.3. Esta segmentación se aplicó a todas las imágenes de las bases de datos.

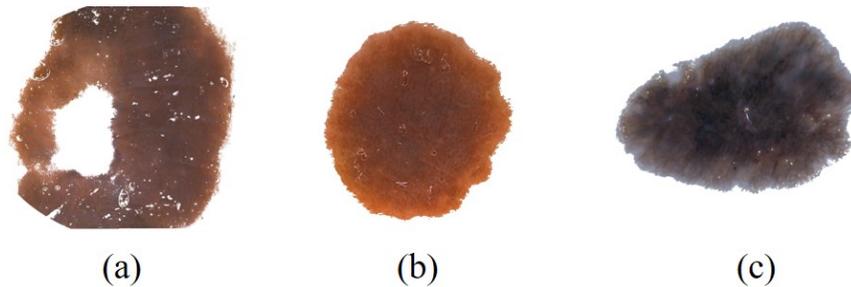


Figura 3.3: Resultados de segmentación manual (a) Melanoma, (b) Nevo, (c) Benigno.

### 3.2.2. Segmentación semi automática

El segundo procedimiento consiste en hacer un preprocesamiento de la imagen aplicando operaciones morfológicas y filtros; para este procedimiento se programó un algoritmo en Python 3, haciendo uso de la librería OpenCV. Tal como se muestra en la figuras 3.4 y 3.5, el procesamiento de la imagen consiste en lo siguiente: se carga la imagen original y se pasa a escala de grises, a esa misma imagen se le aplica un filtro gaussiano con tamaño de ventana 13x13 y desviación estándar de 1.5, una vez aplicado el primer filtro gaussiano se aplica uno nuevo cambiando el valor de la desviación estándar por 1.0, se toma la imagen filtrada y se aplica erosión 5 veces con tamaño de kernel 3x3, a esa imagen erosionada se le aplica dilatación en 1 iteración y finalmente se binariza la imagen para obtener la máscara de recorte que posteriormente es usada para obtener la imagen en RGB con fondo blanco en las mismas condiciones que la imagen segmentada manualmente.

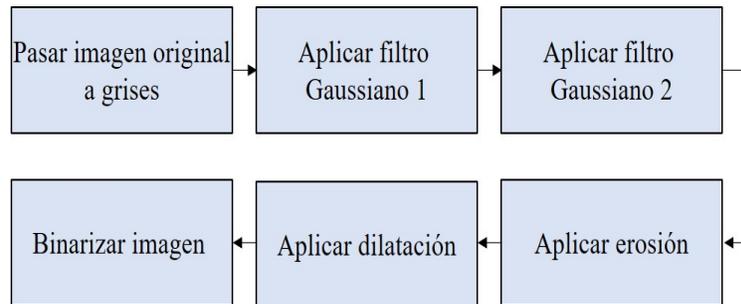


Figura 3.4: Algoritmo segmentación semi automática

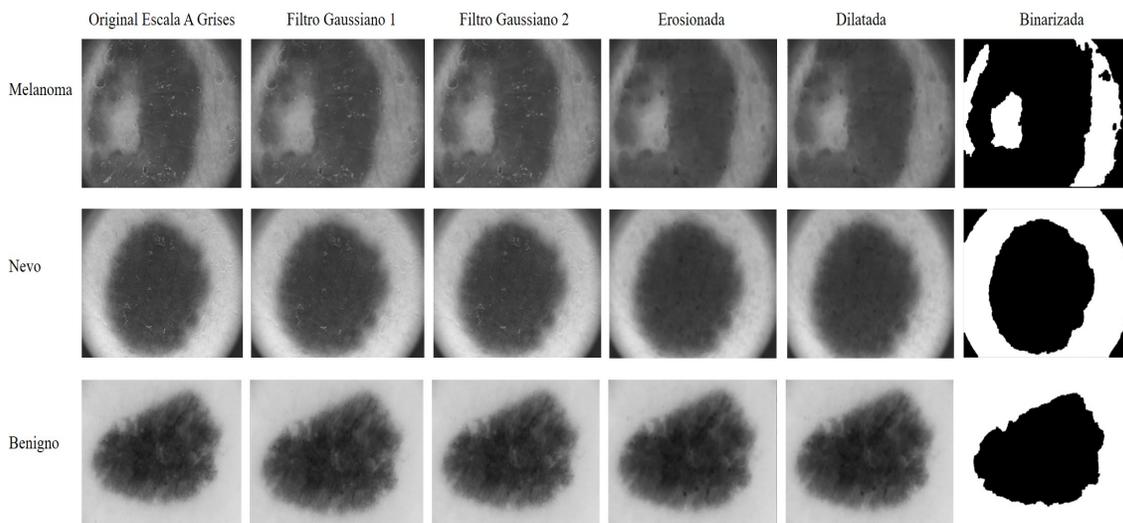


Figura 3.5: Algoritmo segmentación semi automática

### 3.2.3. Segmentación automática

En el tercer procedimiento se utilizó el ejecutable de ImageJ con su extensión de Weka Trainable segmentation, primeramente se realizó el etiquetado de clases manualmente como se muestra en la figura 3.6 en color rojo se etiquetó el área de la imagen que contiene la región de interés y en color verde se etiquetó el área que solo es piel no pigmentada; una vez identificadas las clases a segmentar se procede a establecer las características

de entrenamiento y el tipo de algoritmo de clasificación a utilizar figura 3.7, una vez configurado el algoritmo se procede a entrenarlo y aplicarlo al banco de imágenes. En total se aplicaron 13 de los algoritmos de segmentación que WEKA proporciona, éstos fueron aplicados a todas las imágenes de la base de datos.

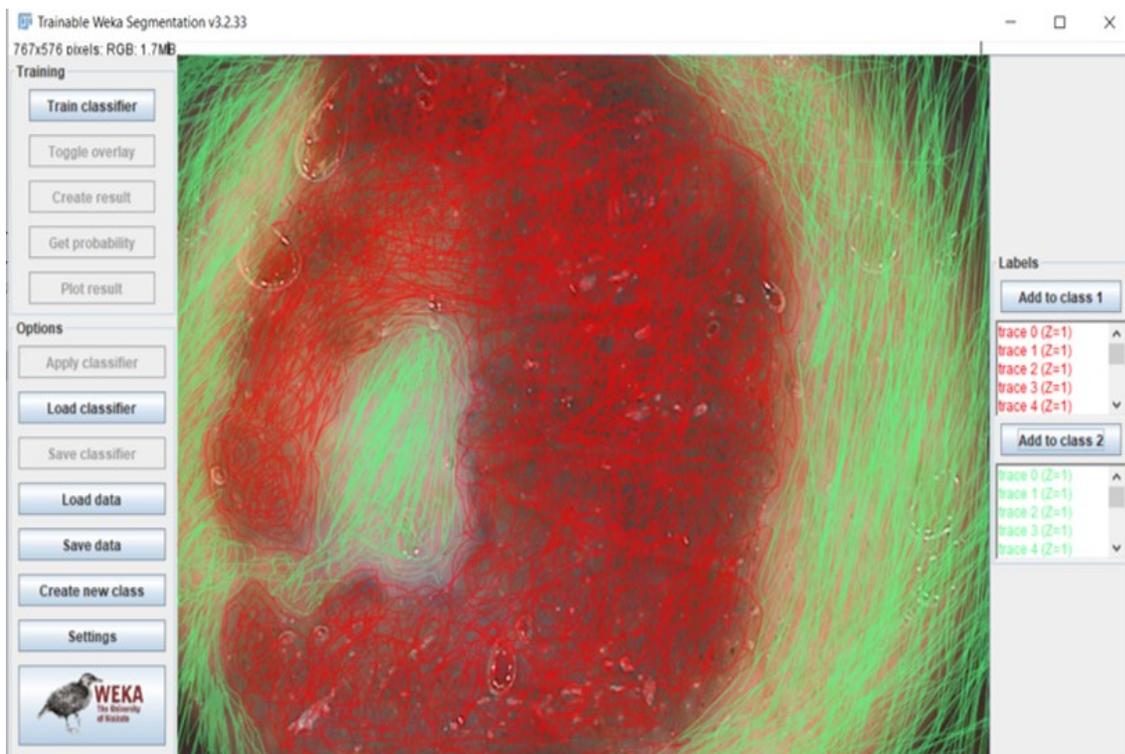


Figura 3.6: Etiquetado de clases en WEKA

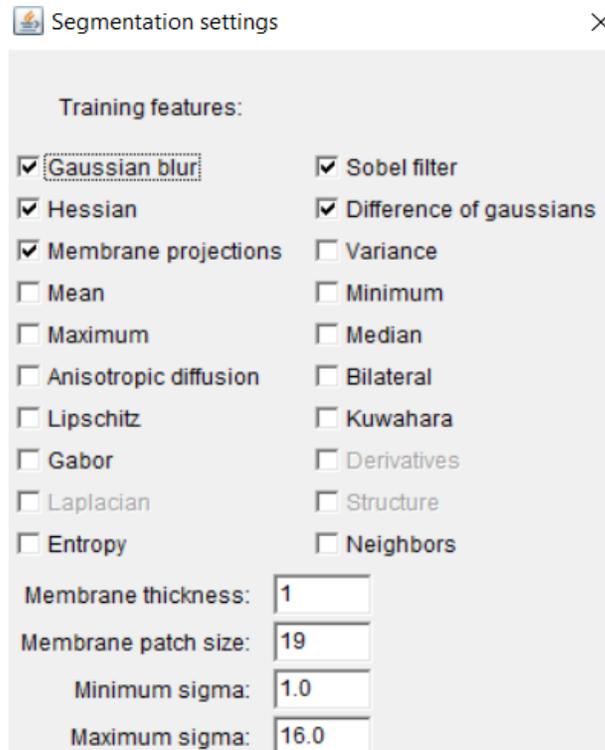


Figura 3.7: Características de entrenamiento y etiquetado de clases en WEKA

### 3.2.4. Binarización

Una vez aplicadas las técnicas de segmentación, se procedió a realizar la binarización de las imágenes resultado, esto con la finalidad de generar unas máscaras binarias que posteriormente se utilizarían para recortar las imágenes originales. El resultado de las máscaras binarias se observa en la figura 3.8.

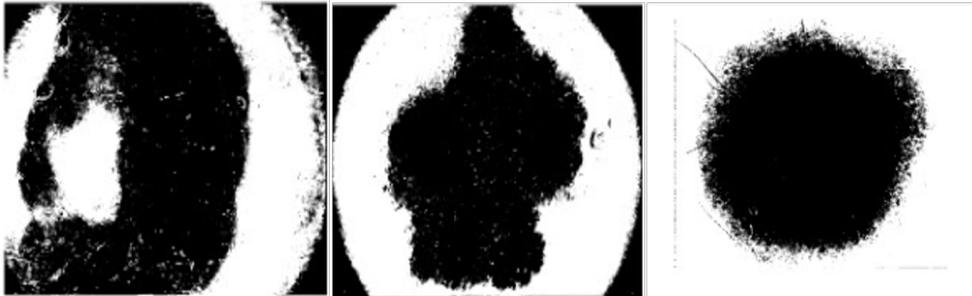


Figura 3.8: Máscaras binarias segmentadas

### 3.2.5. Eliminación de ruido y recorte a partir de máscaras binarias

Después de obtener las máscaras binarias de las diferentes técnicas de segmentación se notó que debido al ángulo de captura y la ausencia de luz en los bordes de la imagen, se estaban teniendo pixeles que no formaban parte de la ROI, definidos como ruido, lo que afectaría notoriamente la segmentación. Para eliminar el ruido se crearon dos máscaras en forma de elipse, una vertical y una horizontal, así se procedería a realizar un recorte con esas máscaras en las imágenes binarias antes de hacer el recorte en las imágenes originales. El proceso de la eliminación de ruido se muestra en las imágenes de la figura 3.9.

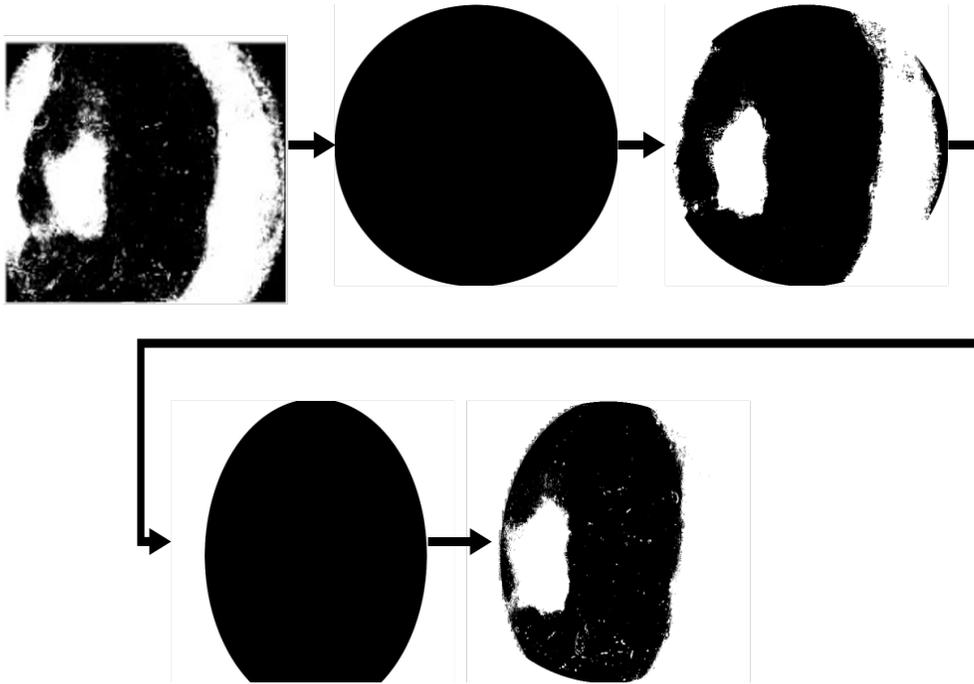


Figura 3.9: Proceso de eliminación de ruido en máscaras binarias

Una vez que se eliminó la mayor área posible de ruido se procedió a realizar el recorte de la imagen original, el proceso se realizó con ambos tipos de imágenes las segmentadas automática y semi-automáticamente. Los resultados se muestran en la figura 3.10, así como también se muestran ejemplos de las imágenes segmentadas manualmente, para permitir un análisis visual de los resultados.



Figura 3.10: Proceso de eliminación de ruido y recorte

Una vez completada esta etapa, la base de datos está en condiciones para darle continuidad al proceso de detección de malignidad en neoplasias.

### 3.3. Aplicación de filtro de Gabor

Tal como se revisó en el estado del arte, el aplicar un filtro de Gabor en las imágenes permite potenciar la textura, sin embargo, no hay una configuración universal del filtro de Gabor sino que el valor de cada uno de sus parámetros depende de las condiciones de las imágenes a las que se les desea aplicar el filtro. Para esta etapa de la metodología, el experimento consiste en lo siguiente: se toman 200 imágenes de la clase de neoplasias

benignas y 200 imágenes de la clase de neoplasias malignas. Se debe aplicar el filtro de Gabor a ambas clases, para ello se propone variar solo dos parámetros del filtro de Gabor, los cuales son el tamaño de ventana y el ángulo ( $\theta$ ). El objetivo es tomar los estadísticos de los píxeles que componen la imagen después de haberles aplicado el filtro de Gabor para ver el comportamiento de sus valores de intensidad, es decir, se va a analizar la textura en las clases para ver si existen diferencias entre una clase y otra. Se programó el filtro y posteriormente se realizó el cálculo de la media y desviación estándar de cada una de las imágenes, al final se tomó un promedio de cada uno de los parámetros por clase, y se analizaron los resultados.

### **3.4. Descripción**

La descripción de la textura se puede llevar a cabo de 3 formas: estadística, basada en modelos y basada en filtrado [15], en este trabajo se combinó la parte de filtrado y estadística. Una vez que se obtuvieron las imágenes segmentadas se debía aplicar el filtro de Gabor porque el descriptor principal para el clasificador es la textura, se propuso realizar un recorrido en la configuración del filtro cambiando sus parámetros de ángulo y tamaño de ventana. Básicamente el recorrido que se realiza es el que se muestra en la figura número 3.11, la ventana del filtro de Gabor recorre la imagen y va modificando el valor de los píxeles. En la figura 3.12 se muestra la imagen una vez que se ha filtrado.

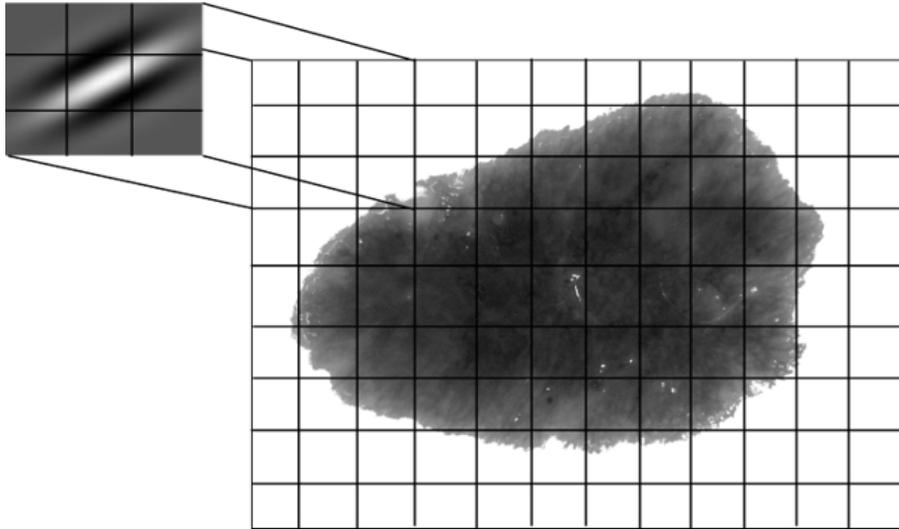


Figura 3.11: Ejemplo de Aplicación de un filtro de Gabor

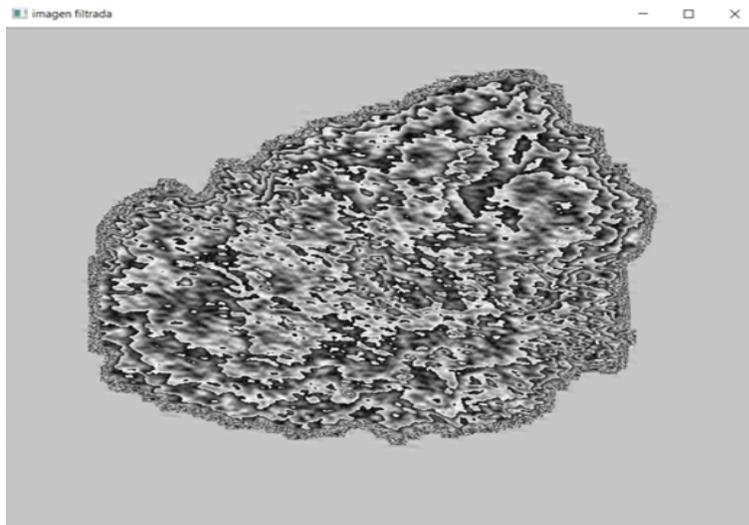


Figura 3.12: Ejemplo de imagen filtrada

# Capítulo 4

## Análisis de resultados

### 4.1. Análisis de resultados

A continuación se muestran los resultados obtenidos en las diferentes etapas de experimentación, en la primera parte de las sub secciones de éste capítulo se reportan los coeficientes de correlación obtenidos en las diferentes segmentaciones que se programaron. En la siguiente subsección se muestran los resultados de los tiempos utilizados en la ejecución de los algoritmos de segmentación. Por último se muestran los resultados obtenidos a partir de la aplicación del filtro de Gabor en las imágenes, así como su configuración más efectiva en su aplicación en un clasificador KNN.

#### 4.1.1. Coeficientes de correlación 2D

Teniendo los resultados de segmentación (automática y semi automática) se debía elegir una de las técnicas de segmentación para poder dar continuidad a la investigación. Para saber qué técnica dio mejores resultados se hizo el cálculo de coeficientes de correlación

en 2 dimensiones para cada una de las imágenes, tomando como referencia las imágenes segmentadas manualmente, ya que ésta segmentación se consideró como la ideal.

Resultado del cálculo de coeficientes se obtuvo un coeficiente de correlación 2D para cada una de las imágenes, así que se tomaron los coeficientes de cada tipo de imagen (melanoma, nevo, benigno) y se obtuvo un promedio para mostrar los resultados generales por tipo de segmentación y por tipo de imagen. Se tomaron en cuenta 3 clases para ver el comportamiento de la segmentación en cada uno de los tipos de imágenes, diferenciando así melanomas de nevos aunque ambas clases pertenecen al tipo de neoplasias malignas.

Los resultados son los siguientes: en la Tabla 4.1 se muestran los resultados de los promedios de coeficientes de correlación en 2D; en la columna 1 se muestran los 13 algoritmos de segmentación automática aplicada, en la columna 2 el promedio de los coeficientes de correlación obtenidos en melanomas, en la columna 3 el promedio de coeficientes de correlación obtenidos en nevos y en la columna 4 el promedio obtenido en benignos. También en la última fila de la tabla se muestran los promedios de los coeficientes de correlación de la segmentación semi automática. El promedio de coeficiente de correlación más cercano a 1 indica qué algoritmo de segmentación resultó ser mejor, además de que se resaltan visualmente para ayudar a identificarlos.

Tabla 4.1: Promedio de coeficientes de correlación

Algoritmo de segmentación automática	Melanomas	Nevos	Benignos
attribute selected classifier	0.7642	0.8326	0.9204
bayes net	0.7782	0.6782	<b>0.9557</b>
bftree	0.7613	0.8612	0.8932
classification via clustering	0.7633	0.8519	0.4826
fast random forest	0.7725	0.8492	0.8999
filtered classifier	0.7444	0.8052	0.8476
forest pa	0.7468	0.8398	0.9201
mlp Classifier	0.7762	0.8469	0.9134
multiclass classifier	0.7790	0.8518	0.8810
multiLayer perceptron	0.7630	0.8340	0.9231
naive bayes	0.7679	0.8241	0.8558
randomizable filtered	0.7561	0.7811	0.9544
simple logistic	0.7579	0.8574	0.9316
<b>segmentación semi automática</b>	<b>0.7891</b>	<b>0.8757</b>	0.9412

Tal como se muestra en la tabla 4.1 la segmentación semi-automática demostró tener un mejor desempeño en comparación con los diferentes algoritmos de segmentación automática. Por ésta razón se eligió esa segmentación para poder continuar con las demás etapas.

#### 4.1.2. Tiempo

Además de evaluar el comportamiento de cada algoritmo con el cálculo de los coeficientes de correlación, se tomó una muestra de 40 imágenes de cada clase y se tomó el tiempo empleado en cada una de las técnicas de segmentación. Primeramente se obtuvo el tiempo de procesamiento por imagen para después obtener un promedio del tiempo en milisegundos. Es importante mencionar que en el caso de los algoritmos de segmentación automática no se tomó en cuenta el tiempo de etiquetado de clases ni entrenamiento

del algoritmo de segmentación únicamente el tiempo empleado en la segmentación. Los resultados son reportados en la tabla 4.2

Tabla 4.2: Promedio de tiempo de procesamiento en milisegundos

Algoritmo de segmentación automática	Melanomas	Nevos	Benignos
attribute selected classifier	1572	1017	2608
bayes net	2730	7377	25207
bftree	1217	390	575
classification via clustering	1619	1721	1098
fast random forest	1161	2662	2762
filtered classifier	3467	4239	6577
forest pa	4916	827	1832
mlp Classifier	2240	3528	6614
multiclass classifier	1932	8763	11507
multiLayer perceptron	2740	20824	60768
naive bayes	4920	24292	30195
randomizable filtered	29079	30860	43180
simple logistic	2110	1248	6874
<b>segmentación semi automática</b>	<b>5</b>	<b>6</b>	<b>12</b>

Tal como se muestran los resultados en la tabla 4.2, la segmentación automática se logró realizar en menos tiempo en comparación con los algoritmos de segmentación automática.

### 4.1.3. Aplicación de filtro de Gabor

Se aplicó el filtro de Gabor a las dos clases de imágenes y se obtuvieron los valores de los pixeles de cada una de ellas para obtener su media y desviación estándar, por último se tomó un promedio de sus parámetros y los resultados se muestran en la tabla 4.3: en las columnas 1 y 2 se muestran los tamaños de ventana y ángulos que se configuraron, en la tercera y cuarta columna se pueden ver los promedios de los estadísticos de la clase

benignos, y en las últimas dos columnas se observan los estadísticos de la clase malignos.

Tabla 4.3: Promedio de media y desviación estándar en las imágenes filtradas

Tamaño máscara	Ángulo	Benignos		Malignos	
		Media	Desv. Std	Media	Desv. Std
<b>3x3</b>	<b>60</b>	<b>185.51</b>	<b>40.49</b>	<b>180.15</b>	<b>46.75</b>
	70	187.92	41.27	182.37	47.49
	80	183.93	40.00	178.67	46.25
	100	187.12	41.00	181.63	47.25
	140	184.71	40.24	179.41	46.50
6x6	60	224.23	57.44	216.40	62.70
	70	120.24	32.86	121.02	37.29
	80	86.43	37.90	89.93	42.21
	100	225.80	59.62	218.38	63.91
	140	135.31	32.93	134.80	37.28
9x9	60	68.93	43.80	73.97	48.30
	70	49.43	50.48	56.08	55.11
	80	222.42	63.10	217.04	63.94
	100	134.47	33.41	133.98	37.70
	140	181.30	43.04	176.78	47.65

Observando los resultados obtenidos en la tabla 4.3, se puede concluir que la desviación estándar es el parámetro que marca una notable diferencia entre clases. Se calculó una diferencia entre los valores obtenidos en la clase benignos y en la clase malignos y se pudo ver que la configuración de tamaño de ventana 3x3 con ángulo de 60°, fue la que marcó una diferencia mayor entre clases, así que se tomó esa configuración para utilizarla en un clasificador KNN, éste arrojó un 71.74% de efectividad en la clasificación, únicamente tomando en cuenta los valores de dispersión para entrenarlo. En la tabla número 4.4, se presentan los valores de dispersión de 5 imágenes de la clase benignos y 5 imágenes

de la clase malignos para hacer notar la variación que hay en sus parámetros entre una y otra clase, además en la figura número 4.1 se muestra gráficamente la variación de los valores de los píxeles en cada clase de imágenes, en la gráfica de la clase benignos se puede observar que el comportamiento es casi lineal, puesto que no hay mucha variación entre los valores de los píxeles y estos tienden a un comportamiento uniforme, por otro lado, al observar la gráfica de la imagen de la clase malignos, se puede ver que hay valores altos y bajos constantemente, ya que hay más variación entre los valores de los píxeles porque hay más variación en los colores presentes en la imagen. Además en la sección de Anexos, se integra una tabla con las gráficas de todas las imágenes de la base de datos.

Tabla 4.4: Valores ejemplo de desviación estándar

Benignos	Malignos
12.52	75.23
15.23	55.7
30.82	49.21
22.54	72.25
17.84	71.75

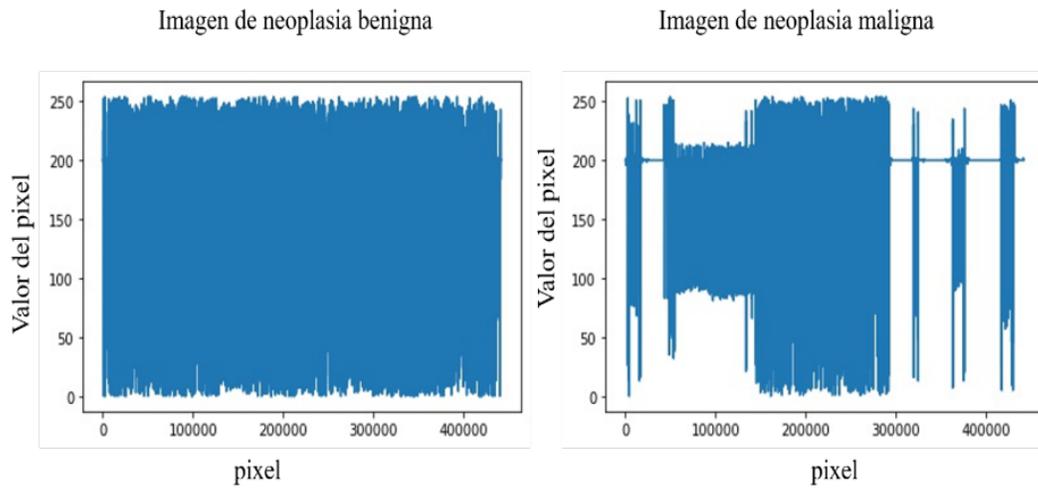


Figura 4.1: Ejemplo de dispersión de valores por clase

#### 4.1.4. Generación de poblaciones

Los valores de media y desviación estándar mostrados en la tabla 4.3, que marcaron más diferencia entre clases se tomaron para hacer un experimento de generación de poblaciones con una distribución normal, se simularon dos poblaciones, una para la clase malignos y otra para la clase benignos. Este experimento se llevó a cabo con la finalidad de ver el comportamiento de las poblaciones gráficamente, los resultados se muestran en la figura 4.2, tal como se puede ver no hay mucha diferencia entre una y otra, puesto que sus medidas de tendencia central son cercanas, esto ocasiona que se sobrepongan una sobre otra.

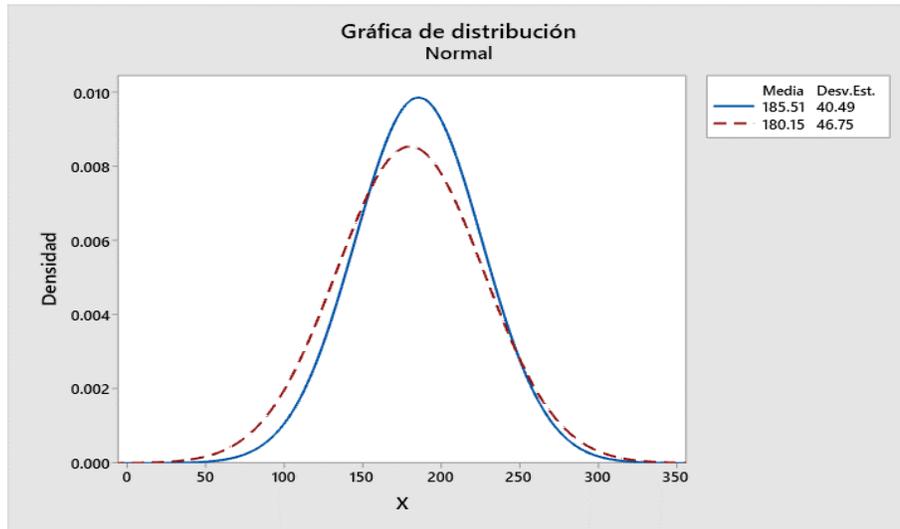


Figura 4.2: Generación de poblaciones con una distribución normal

#### 4.1.5. Clasificador KNN

Para poder cuantificar la efectividad de los valores de dispersión entre clases, se propuso programar un clasificador KNN. Para este experimento, primeramente se generó el dataset para alimentar el entrenamiento del clasificador, así que se pasaron los datos de media, desviación estándar y la etiqueta de la clase (maligno/benigno). Una vez generada la base de datos se procedió a utilizar un método de validación cruzada en 8 pliegues, los resultados arrojaron un 71 % de efectividad en la clasificación.

# Capítulo 5

## Conclusiones y trabajo futuro

### 5.1. Conclusiones

Sí es posible la cuantificación de color en neoplasias, en esta investigación fue a través de sus medidas de dispersión y tendencia central; con esa cuantificación se generó un dataset que sirvió para entrenar un clasificador con más del 70 por ciento de efectividad, con lo que se puede ver claramente que usando solo la textura como descriptor en un clasificador sí se tiene el poder discriminante para detectar la malignidad o benignidad de la neoplasia. Si al color se le asocia el uso de algún otro descriptor se obtendrá un porcentaje de clasificación superior. Es importante destacar que en este trabajo no se están utilizando bancos de filtro de Gabor, sino un solo filtro de Gabor con una configuración específica.

De acuerdo con los resultados mostrados en la tabla 4.1: promedios de coeficientes de correlación, la mejor segmentación para las neoplasias benignas es la segmentación automática utilizando el algoritmo Bayes Net, mientras que para los malignos (melanomas

y nevus) es la segmentación semi automática. Esto se debe a la variación de color presente en cada una de las clases.

En la tabla 4.3: Promedio de media y desviación estándar, se muestra que la clase de neoplasias malignas tiene en promedio una desviación estándar de 46.75, mientras que la de benignas en promedio tiene una desviación estándar de 40.49, esto sucede porque las neoplasias malignas presentan una variedad de colores mayor.

## **5.2. Recomendaciones y trabajo futuro**

En el trabajo a futuro se puede programar un algoritmo evolutivo para encontrar la configuración adecuada de los parámetros del filtro de Gabor o el banco de filtros de Gabor que genere un porcentaje de clasificación mayor. También se puede experimentar con la combinación de los demás descriptores de asimetría, borde y diámetro, en conjunto con el color para aumentar la efectividad de clasificación. Un segundo algoritmo evolutivo podría programarse para ver qué combinación de descriptores devuelve mejores resultados.

Por último, se recomienda aumentar el tamaño de la base de datos experimental para poder generar nuevos valores de dispersión y tendencia central que permitan generar una nueva población desde una muestra grande, esto para analizar el comportamiento de ambas poblaciones al tratar de separarlas.

# Referencias

- Araujo Moncayo, D. A., Capera Tole, A. M., Plazas Montaña, J. J., y cols. (s.f.). Cáncer en piel por exposición a radiación ultravioleta solar en trabajadores de construcción en el distrito capital.
- Arganda-Carreras, I., Kaynig, V., Rueden, C., Eliceiri, K. W., Schindelin, J., Cardona, A., y Sebastian Seung, H. (2017). Trainable weka segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics*, 33(15), 2424–2426.
- Armengot-Carbó, M., Martínez-Roig, V., y Bort, G. P. (2020). Prioridad de derivación de los casos de melanoma desde atención primaria. *Piel*, 35(4), 225–229.
- Azuela, J. H. S. (2013). *Visión artificial: rasgos descriptores para el reconocimiento de objetos*. Ra-Ma.
- Bradski, G., y Kaehler, A. (2008). *Learning opencv: Computer vision with the opencv library*. O'Reilly Media, Inc.
- Celebi, M. E., Kingravi, H. A., Uddin, B., Iyatomi, H., Aslandogan, Y. A., Stoecker, W. V., y Moss, R. H. (2007). A methodological approach to the classification of dermoscopy images. *Computerized Medical imaging and graphics*, 31(6), 362–373.
- Celebi, M. E., Wen, Q., Iyatomi, H., Shimizu, K., Zhou, H., y Schaefer, G. (2015). A state-of-the-art survey on lesion border detection in dermoscopy images. *Dermoscopy image analysis*, 10, 97–129.
- Chazallet, S. (2016). *Python 3: los fundamentos del lenguaje*. Ediciones ENI.
- Cruz Mahecha, M. F., Vargas Martínez, M. F., y cols. (2018). Detección de melanomas a partir de imágenes dermatoscópicas.
- Garnavi, R., Aldeen, M., Celebi, M. E., Bhuiyan, A., Dolianitis, C., y Varigos, G. (2010). Automatic segmentation of dermoscopy images using histogram thresholding on optimal color channels. *International Journal of Medicine and Medical Sciences*, 1(2), 126–134.
- Goldstein, B. G., y Goldstein, A. O. (2001). Diagnosis and management of malignant melanoma. *American family physician*, 63(7), 1359.
- Gonzalez, R. C., Woods, R. E., y cols. (2002). *Digital image processing*. Prentice hall Upper Saddle River, NJ.

- Gonzalo, P. M., de la CruzJesus, M., y cols. (2008). *Vision por computador. imagenes digitales y aplicaciones*. (Inf. Téc.).
- Hawkins, J. K. (1970). Textural properties for pattern recognition. *Picture processing and psychopictorics*, 347–370.
- HealSmith, M., Bourke, J., Osborne, J., y Graham-Brown, R. (1994). An evaluation of the revised seven-point checklist for the early diagnosis of cutaneous malignant melanoma. *British Journal of Dermatology*, 130(1), 48–50.
- Heinze-Martin, G., Olmedo-Canchola, V. H., Bazán-Miranda, G., Bernard-Fuentes, N. A., y Guízar-Sánchez, D. P. (2018). Medical specialists in Mexico. *Gaceta Médica de México*, 154(3), 342–351.
- Homatidis, S., y Konstantareas, M. M. (1981). Assessment of hyperactivity: Isolating measures of high discriminant ability. *Journal of Consulting and Clinical Psychology*, 49(4), 533.
- Iyatomi, H. (2010). Computer-based diagnosis of pigmented skin lesions. *book: Campolo D*, ed. *New developments in biomedical engineering*, 183–200.
- Johr, R. H. (2002). Dermoscopy: alternative melanocytic algorithms—the abcd rule of dermatoscopy, menzies scoring method, and 7-point checklist. *Clinics in dermatology*, 20(3), 240–247.
- Lopez, A. R., Giro-i Nieto, X., Burdick, J., y Marques, O. (2017). Skin lesion classification from dermoscopic images using deep learning techniques. En *2017 13th iasted international conference on biomedical engineering (biomed)* (pp. 49–54).
- Madooei, A., y Drew, M. S. (2016). Incorporating colour information for computer-aided diagnosis of melanoma from dermoscopy images: A retrospective survey and critical analysis. *International journal of biomedical imaging*, 2016.
- Martínez, M. A. C., Funes, F. J. G., Rosales-Silva, A., y Ramos-Arredondo, R. I. (2016). Preprocesamiento de imágenes dermatoscópicas para extracción de características. *Res. Comput. Sci.*, 114, 59–70.
- Mendonca, T., Celebi, M., Mendonca, T., y Marques, J. (2015). Ph2: A public database for the analysis of dermoscopic images. *Dermoscopy image analysis*.
- Mendonca, T., Marcal, A. R., Vieira, A., Nascimento, J. C., Silveira, M., Marques, J. S., y Rozeira, J. (2007). Comparison of segmentation methods for automatic diagnosis of dermoscopy images. En *2007 29th annual international conference of the IEEE Engineering in Medicine and Biology Society* (pp. 6572–6575).
- Mindru, F., Moons, T., y Van Gool, L. (1999). Recognizing color patterns irrespective of viewpoint and illumination. En *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (cat. no pr00149)* (Vol. 1, pp. 368–373).
- Mohammed, M., Khan, M. B., y Bashier, E. B. M. (2016). *Machine learning: algorithms and applications*. Crc Press.
- Nachbar, F., Stolz, W., Merkle, T., Cognetta, A. B., Vogt, T., Landthaler, M., ... Ple-

- wig, G. (1994). The abcd rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions. *Journal of the American Academy of Dermatology*, 30(4), 551–559.
- Ng, V. T., Fung, B. Y., y Lee, T. K. (2005). Determining the asymmetry of skin lesion with fuzzy borders. *Computers in biology and medicine*, 35(2), 103–120.
- Palencia, A. B. (s.f.). Anatomía y fisiología de la piel.
- Pathan, S., Prabhu, K. G., y Siddalingaswamy, P. (2018). Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—a review. *Biomedical Signal Processing and Control*, 39, 237–262.
- Peñafiel, M. J. M. (2014). Melanoma maligno cutáneo en una mujer indígena del municipio de alto baudó, chocó, colombia. *Médicas UIS*, 27(1).
- Pereira, F., Mitchell, T., y Botvinick, M. (2009). Machine learning classifiers and fmri: a tutorial overview. *Neuroimage*, 45(1), S199–S209.
- Pérez Carrasco, J. A., Serrano Gotarredona, M. d. C., y Acha Piñero, B. (2009). Clasificación de lesiones de piel basada en filtros de gabor y color. En *Congreso anual de la sociedad española de ingeniería biomédica (caseib)(2009)*, p 125-128.
- Petrou, M. M., y García, P. (2006). *Image processing: dealing with texture*. John Wiley & Sons.
- Piccolo, D., Ferrari, A., Peris, K., Daidone, R., Ruggeri, B., y Chimenti, S. (2002). Dermoscopic diagnosis by a trained clinician vs. a clinician with minimal dermoscopy training vs. computer-aided diagnosis of 341 pigmented skin lesions: a comparative study. *British Journal of Dermatology*, 147(3), 481–486.
- Pickett, R. M. (1970). Visual analysis of texture in the detection and recognition of objects. *Picture processing and psychopictorics*, 289–308.
- Pratt, W. K. (2013). *Introduction to digital image processing*. CRC press.
- Ramadan, Z. M. (2017). Using entropy and 2-d correlation coefficient as measuring indices for impulsive noise reduction techniques. *International Journal of Applied Engineering Research*, 12(21), 11101-11106..
- Rangel Cortes, J. J. (2019). Sistema para pre-diagnosticar el cáncer de piel tipo melanoma.
- Rastgoo, M., Garcia, R., Morel, O., y Marzani, F. (2015). Automatic differentiation of melanoma from dysplastic nevi. *Computerized Medical Imaging and Graphics*, 43, 44–52.
- Rodríguez-Cabral, R., Montoya-Sosa, G., Roldán-Marín, R., y Carlos-Ortega, B. (2014). Principios básicos de dermatoscopia. *Dermatología Revista Mexicana*, 58(3), 300–304.
- Rubegni, P., Cevenini, G., Burroni, M., Perotti, R., Dell’Eva, G., Sbano, P., . . . others (2002). Automated diagnosis of pigmented skin lesions. *International Journal of Cancer*, 101(6), 576–580.

- Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., y Eliceiri, K. W. (2017). Imagej2: Imagej for the next generation of scientific image data. *BMC bioinformatics*, 18(1), 1–26.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, 3(3), 210–229.
- Sancen-Plaza, A., Santiago-Montero, R., Sossa, H., Perez-Pinal, F. J., Martinez-Nolasco, J. J., y Padilla-Medina, J. A. (2018). Quantitative evaluation of binary digital region asymmetry with application to skin lesion detection. *BMC medical informatics and decision making*, 18(1), 1–11.
- Sandoval Serrano, L. J., y cols. (2018). Algoritmos de aprendizaje automático para análisis y predicción de datos. *Revista Tecnológica; no. 11*.
- Santiago-Montero, R., y Gaucín, A. M. O. (2015). Discriminación de lesiones de piel mediante la descripción cuantitativa del borde y asimetría utilizando el concepto de compacidad.
- Smaoui, N., y Bessassi, S. (2013). A developed system for melanoma diagnosis. *International Journal of Computer Vision and Signal Processing*, 3(1), 10–17.
- Sonka, M., Hlavac, V., y Boyle, R. (2014). *Image processing, analysis, and machine vision*. Cengage Learning.
- Vestergaard, M., Macaskill, P., Holt, P., y Menzies, S. (2008). Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting. *British Journal of Dermatology*, 159(3), 669–676.
- Villalonga Gutiérrez, E. P., y cols. (2019). Inteligencia artificial y el internet de las cosas: estudio y estado actual en empresas y consumidores.
- Zaballos, P., Carrera, C., Puig, S., y Malveyh, J. (2004). Criterios dermatoscópicos para el diagnóstico del melanoma. *Med Cutan Iber Lat Am*, 32(1), 3–17.

# Apéndice A

## Códigos programados

# CÓDIGO DE SEGMENTACIÓN SEMI-AUTOMÁTICA EN PYTHON

```
1 #SE IMPORTAN LAS LIBRERIAS A UTILIZAR
2 import matplotlib.pyplot as plt
3 import numpy as np
4 import cv2
5 from PIL import Image
6 #SE CARGAN LAS IMAGENES DE LA BASE DE DATOS
7 imagenes=np.array(['1.jpg', '2.jpg', '3.jpg', ..., '200.jpg'])
8 #SE CARGA LA MASCARA DE ELIPSE PARA ELIMINAR RUIDO
9 mascaraElipse=cv2.imread('mascaraRecorte.jpeg', 0)
10 #SE BINARIZA LA MASCARA
11 _, imagenBinaria=cv2.threshold(mascaraElipse, 210, 255, cv2.
    THRESH_BINARY_INV)
12 #SE CARGA LA OTRA MASCARA PARA ELIMINAR RUIDO
13 mascaraVertical=cv2.imread('mascaraVertical.jpeg', 0)
14 #SE ESTABLECE EL TAMANO DE LA IMAGEN PARA NORMALIZAR LA BASE DE
    DATOS
15 width = 767
16 height = 576
17 #SE INICIA UN CICLO EN EL QUE SE VA A OBTENER LA SEGMENTACION DE LA
    BASE DE DATOS
18 for x in range (imagenes.size): #INDICA QUE EL CICLO SE VA A
    EJECUTAR PARA CADA ELEMENTO DEL ARRAY DE LAS IMAGENES
19     imagenOriginal=cv2.imread(imagenes[x], 0)#SE CARGA LA IMAGEN
20     blur = cv2.GaussianBlur(imagenOriginal, (13,13), 1.5) #SE APLICA
    EL PRIMER FILTRO GAUSSIANO
21     blur = cv2.GaussianBlur(blur, (13,13), 1.0)# SE APLICA EL SEGUNDO
    FILTRO GAUSSIANO
22     Kernel = np.ones((3,3),np.uint8) #SE CREA EL KERNEL PARA
    APLICAR OPERACIONES MORFOLOGICAS
23     img_Erosion = cv2.erode(blur, Kernel, iterations=5) #SE APLICA
    EROSION EN 5 ITERACIONES
24     img_Dilatada = cv2.dilate(img_Erosion, Kernel, iterations = 1) #
    SE APLICA DILATACION EN 1 ITERACION
25     _, binarizada = cv2.threshold(img_Dilatada, 120, 255, cv2.
    THRESH_BINARY) #SE BINARIZA LA IMAGEN RESULTANTE
26     dim = (width, height) #SE ESTABLECEN LAS DIMENSIONES
27     resize1 = cv2.resize(imagenBinaria, dim, interpolation = cv2.
    INTER_AREA)# SE NORMALIZA LA MASCARA
28     resize2= cv2.resize(binarizada, dim, interpolation = cv2.
    INTER_AREA)#SE NORMALIZA LA IMAGEN RESULTANTE
29     binariaarray = np.array(resize1)#SE CREA UN ARREGLO DEL TAMANO
    DE LA MASCARA
30     originaarray = np.array(resize2)#SE CREA UN ARREGLO DEL TAMANO
```

```
DE LA IMAGEN RESULTANTE
31 #SE INDICA QUE DONDE LA MASCARA SEA NEGRA (VALOR 0), SE PONGA
UN VALOR BLANCO (255) EN EL ARREGLO DE LA
32 #IMAGEN RESULTANTE, PARA DEJAR EL FONDO COLOR BLANCO
33 originaarray[binariaarray==0]=255
34 newImage1 = Image.fromarray(originaarray)#SE GUARDA EL ARREGLO
COMO IMAGEN
35 mascaraBinaria=originaarray#SE CREA UNA VARIABLE DE LA MASCARA
BINARIA CON LOS VALORES DEL ARREGLO ORIGINAL
36 imagenOriginal1=cv2.imread(imagenes[x])#SE INVOCA NUEVAMENTE A
LA IMAGEN ORIGINAL
37 imagenOriginal1 = cv2.resize(imagenOriginal1, dim,
interpolation = cv2.INTER_AREA)#SE NORMALIZA LA IMAGEN ORIGINAL
38 #SE INDICA QUE DONDE LA MASCARA TENGA VALORES DE 255, SE
APLIQUEN VALORES DE 255 EN LA IMAGEN ORIGINAL
39 #SI LA CONDICION NO SE CUMPLE, SE MANTIENE EL VALOR ORIGINAL
DEL PIXEL DE LA IMAGEN ORIGINAL
40 imagenOriginal1[mascaraBinaria==255]=255
41 #SE INDICA QUE DONDE LA MASCARA VERTICAL TENGA VALORES DE 255,
SE APLIQUEN VALORES DE 255 EN LA IMAGEN ORIGINAL
42 #SI LA CONDICION NO SE CUMPLE, SE MANTIENE EL VALOR ORIGINAL
DEL PIXEL DE LA IMAGEN ORIGINAL
43 imagenOriginal1[mascaraVertical==255]=255
44 cv2.imwrite((imagenes[x]),imagenOriginal1)#SE GUARDA LA IMAGEN
RESULTANTE, ES DECIR LA IMAGEN SEGMENTADA
```

## CÓDIGO DE APLICACIÓN DE FILTRO DE GABOR EN PYTHON

```
1 #SE IMPORTAN LAS LIBRERIAS A UTILIZAR
2 import cv2
3 import numpy as np
4 from skimage.filters import gabor
5 #SE IMPORTA LA BASE DE DATOS DE IMAGENES
6 imagenes=np.array(['1.jpg', '2.jpg', '3.jpg', ..., '200.jpg'])
7 #SE CREAN DOS ARREGLOS PARA GUARDAR LOS VALORES DE MEDIA Y
  DESVIACION DE CADA IMAGEN
8 medias=[]
9 desviaciones=[]
10 #SE INICIA UN CICLO FOR PARA APLICAR EL FILTRO A TODAS LAS IMAGENES
11 for x in range (imagenes.size):#INDICA QUE EL CICLO SE VA A
  EJECUTAR PARA CADA ELEMENTO DEL ARRAY DE LAS IMAGENES
12     img = cv2.imread(imagenes[x])#SE CARGA LA IMAGEN
13     img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)#LA IMAGEN SE PASA A
  ESCALA DE GRISES
14     #SE INDICA LA CONFIGURACION DEL FILTRO DE GABOR
15     filt_real, filtered_img = gabor(img, frequency=9, theta=80,
  bandwidth=3, sigma_x=None, sigma_y=None,
16                                     n_stds=0, offset=0, mode='wrap',
  cval=0)
17     #SE EXCLUYEN LOS VALORES DE 255 EN LA IMAGEN PARA TOMAR LOS
  ESTADISTICOS SOLO DE LOS PIXELES QUE CONFORMAN LA ROI
18     pixelValues=np.where(filtered_img!=255, filtered_img, np.nan)
19     pixelValues = ~np.isnan(pixelValues)
20     #SE CALCULA LA MEDIA DE LOS VALORES OBTENIDOS Y SE GUARDA EN EL
  ARREGLO DE MEDIAS
21     medias.append(np.mean(pixelValues))
22     #SE CALCULA LA DESV. STD. DE LOS VALORES OBTENIDOS Y SE GUARDA
  EN EL ARREGLO DE DESVIACIONES
23     desviaciones.append(np.std(pixelValues))
24 #SE OBTIENE UN PROMEDIO DE TODAS LAS MEDIAS OBTENIDAS DE LAS
  IMAGENES
25 promedioMedias=np.mean(medias)
26 #SE OBTIENE UN PROMEDIO DE TODAS LAS DESVIACIONES OBTENIDAS
27 promedioDesviaciones=np.mean(desviaciones)
28 #SE IMPRIMEN LOS VALORES OBTENIDOS EN CADA VARIABLE
29 print(promedioMedias)
30 print(promedioDesviaciones)
```

# CÓDIGO DE APLICACIÓN DE FILTRO DE GABOR, GENERACIÓN DE DATASET Y CLASIFICADOR KNN EN PYTHON

```
1 #SE IMPORTAN LAS LIBRERIAS A UTILIZAR
2 import numpy as np
3 import cv2
4 import pandas as pd
5 from sklearn.neighbors import KNeighborsClassifier
6 from sklearn.model_selection import cross_val_score
7 from scipy import ndimage as ndi
8 #SE CREA UN KERNEL CON LA CONFIGURACION DEL FILTRO DE GABOR
9 g_kernel = cv2.getGaborKernel((3, 3), 1.0, 60, 1.0, 0.5, 0, ktype=
   cv2.CV_32F)
10 #SE CARGA LA BASE DE DATOS DE LA CLASE BENIGNOS
11 imagesB=np.array(['1.jpg', '2.jpg', '3.jpg', ..., '200.jpg'])
12 #SE CARGA LA BASE DE DATOS DE LA CLASE MALIGNOS
13 imagesM=np.array(['1.bmp', '2.bmp', '3.bmp', ..., '200.bmp'])
14 dataSet=[] #SE CREA UNA VARIABLE PARA GENERAR EL DATASET
15 #SE INICIA EL CICLO FOR PARA LA CLASE DE MALIGNOS
16 for x in range (imagesM.size): #SE INDICA QUE SE APLICA A TODAS LAS
   IMAGENES
17     img = cv2.imread(imagesM[x])#SE CARGA LA IMAGEN
18     img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)#SE PASA A ESCALA DE
   GRISES
19     filtered_img = ndi.convolve(img, g_kernel, mode='wrap')#SE
   APLICA FILTRO DE GABOR
20     pixelValues=np.where(filtered_img!=255, filtered_img, np.nan)#
   SE ELIGEN LOS VALORES DE LA ROI PARA OBTENER ESTADISTICOS
21     mascara = ~np.isnan(pixelValues)
22     mean=round((np.mean(pixelValues[mascara])),2)#SE CALCULA LA
   MEDIA Y SE REDONDEA
23     std=round((np.std(pixelValues[mascara])),2)#SE CALCULA LA
   DESVIACION Y SE REDONDEA
24     dataSet.append([mean,std,'malign'])# SE AGREGA LA ETIQUETA DE
   LA CLASE Y SE INSERTA AL DATASET
25 for x in range (imagesB.size):#SE INDICA QUE SE APLICA A TODAS LAS
   IMAGENES
26     img = cv2.imread(imagesB[x])#SE CARGA LA IMAGEN
27     img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)#SE PASA A ESCALA DE
   GRISES
28     filtered_img = ndi.convolve(img, g_kernel, mode='wrap')#SE
   APLICA FILTRO DE GABOR
29     pixelValues=np.where(filtered_img!=255, filtered_img, np.nan)#
   SE ELIGEN LOS VALORES DE LA ROI PARA OBTENER ESTADISTICOS
```

```

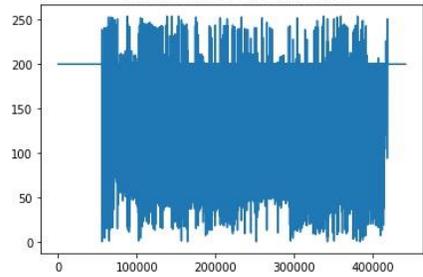
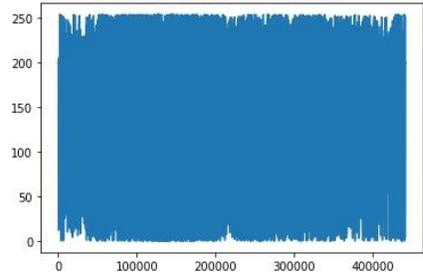
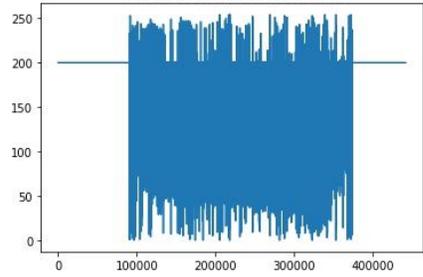
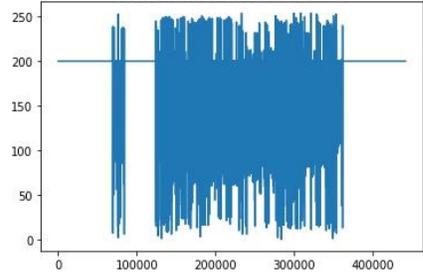
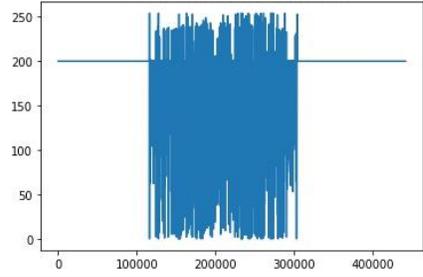
30     mascara = ~np.isnan(pixelValues)
31     mean=round((np.mean(pixelValues[mascara])),2)#SE CALCULA LA
        MEDIA Y SE REDONDEA
32     std=round((np.std(pixelValues[mascara])),2)#SE CALCULA LA
        DESVIACION Y SE REDONDEA
33     dataSet.append([mean,std,'benign'])# SE AGREGA LA ETIQUETA DE
        LA CLASE Y SE INSERTA AL DATASET
34     lista=np.array(dataSet)#SE CONVIERTE EN ARRAY EL DATASET
35     dataset=pd.DataFrame(lista)#SE INDICA EL TIPO DE DATO DEL
        DATASET
36 X = dataset.iloc[:, :-1].values#SE CREA LA VARIABLE CON LOS DATOS
        DE LA MEDIA
37 Y = dataset.iloc[:, 2].values#SE CREA LA VARIABLE CON LOS DATOS DE
        LA DESVIACION
38 KNN = KNeighborsClassifier(n_neighbors=1)#SE INICIA EL CLASIFICADOR
39 scores = cross_val_score(KNN, X, Y, cv=5)#SE APLICA VALIDACION
        CRUZADA
40 accuracy=(scores.mean() * 100)#SE CALCULA LA EFECTIVIDAD
41 print("Resultados de Clasificador KNN")#SE MUESTRAN LOS RESULTADOS
42 print(accuracy)

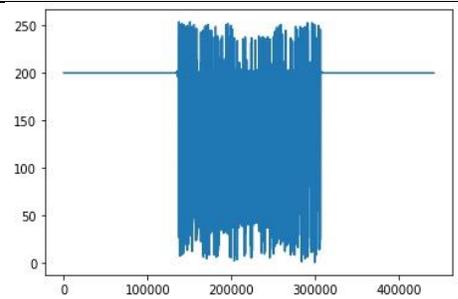
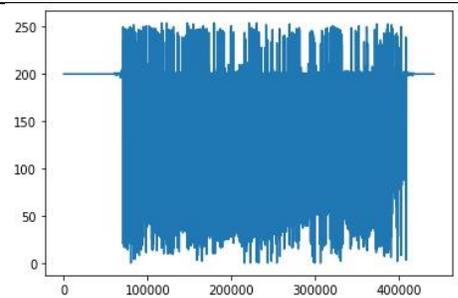
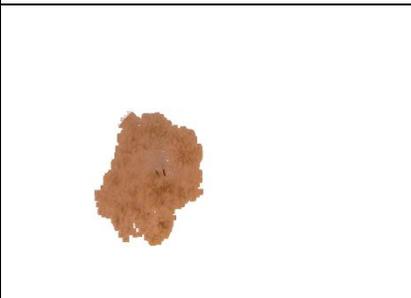
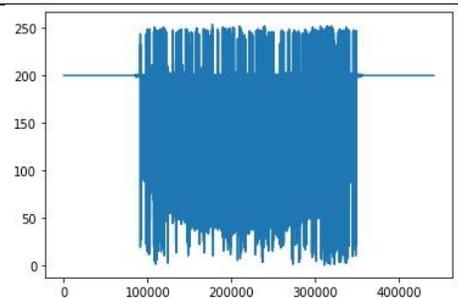
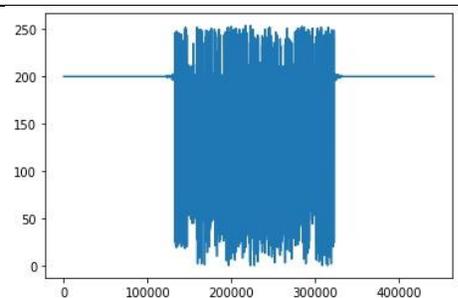
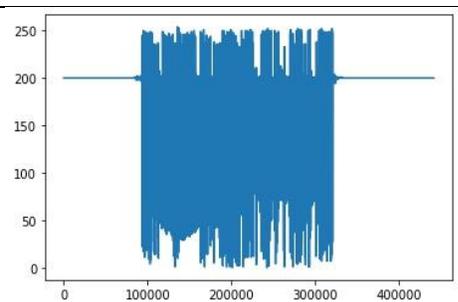
```

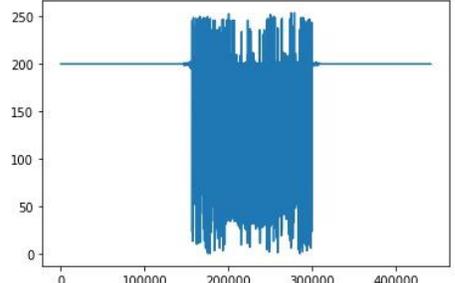
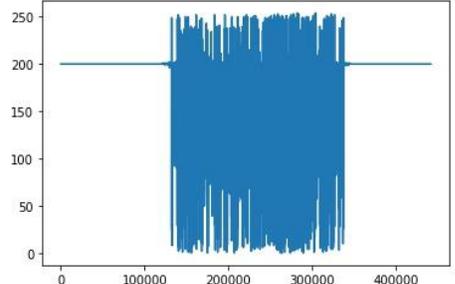
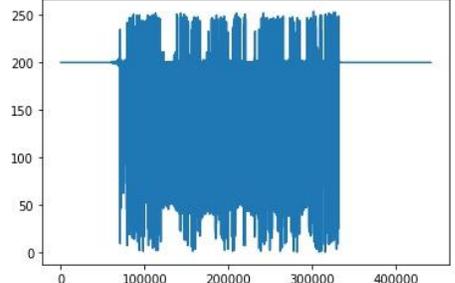
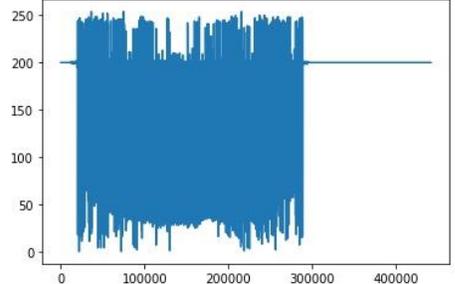
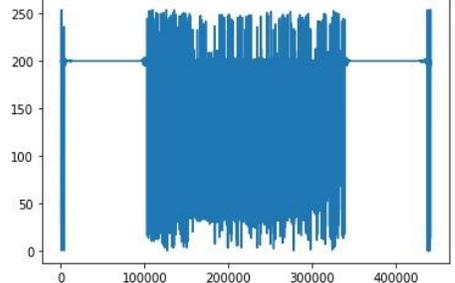
## Apéndice B

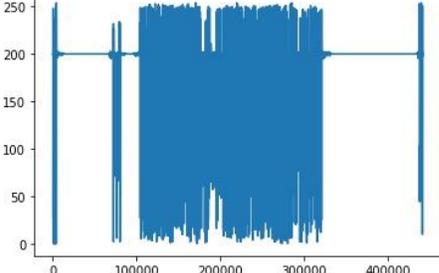
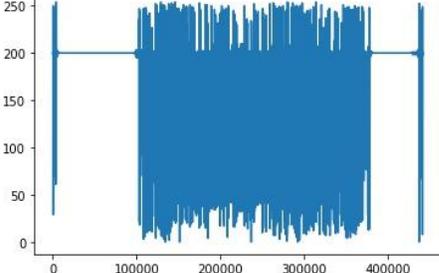
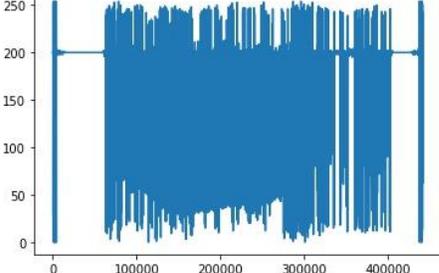
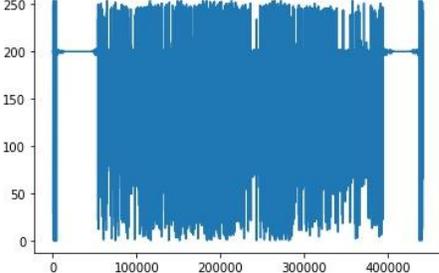
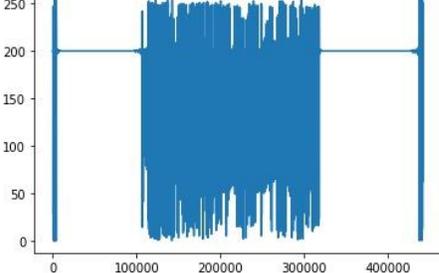
### Gráficas de valores de intensidad clase Malignos

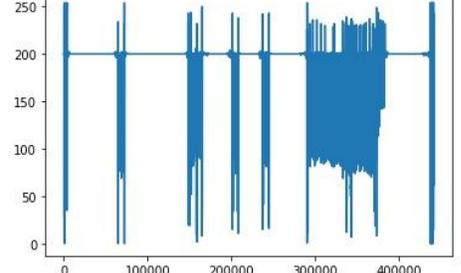
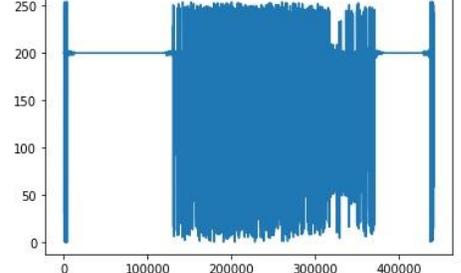
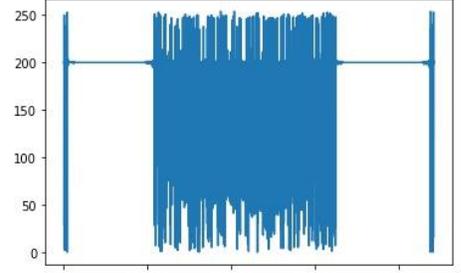
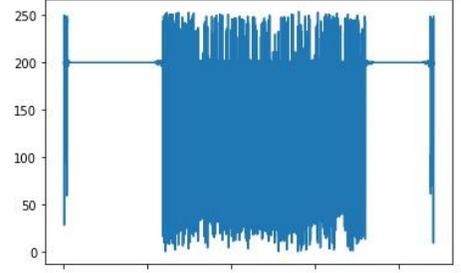
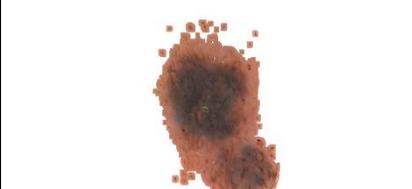
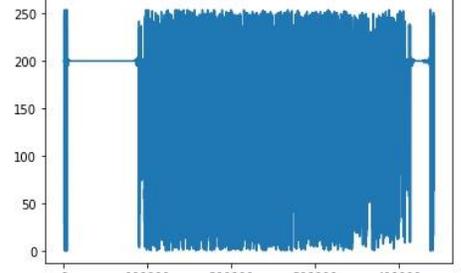
NEOPLASIAS MALIGNAS

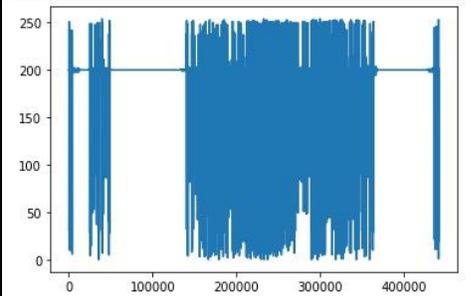
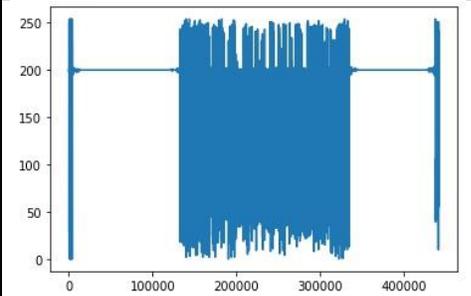
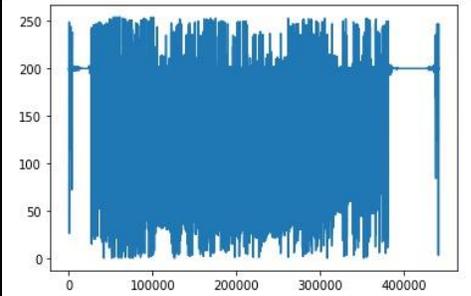
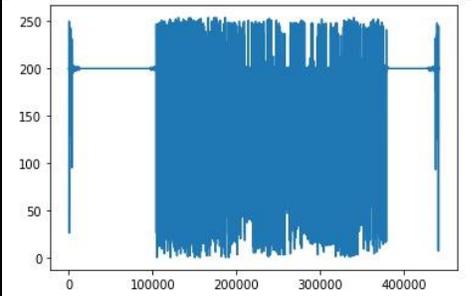
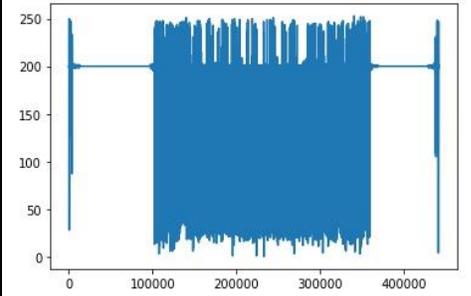
No.	Imagen	Gráfica	Media	Desviación
1		<p>Media: 165.05 Desv. sdttd: 56.08</p> 	165.05	56.08
2		<p>Media: 174.63 Desv. sdttd: 47.18</p> 	174.63	47.18
3		<p>Media: 177.01 Desv. sdttd: 48.28</p> 	177.01	48.28
4		<p>Media: 189.89 Desv. sdttd: 30.82</p> 	189.89	30.82
5		<p>Media: 194.95 Desv. sdttd: 22.54</p> 	194.95	22.54

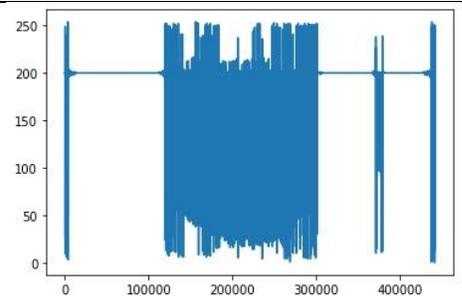
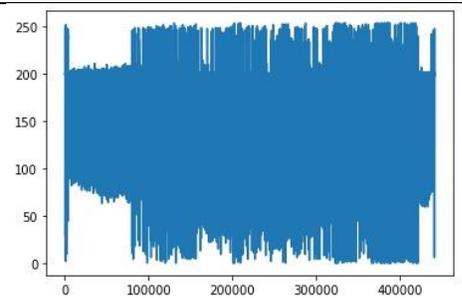
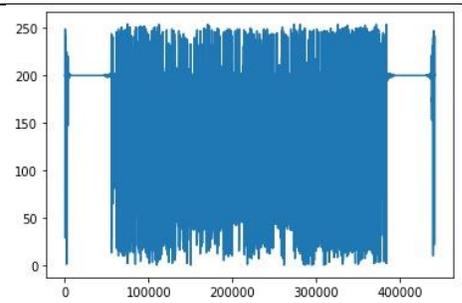
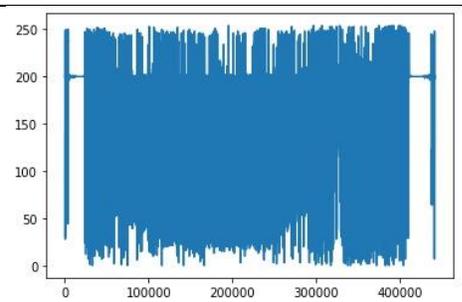
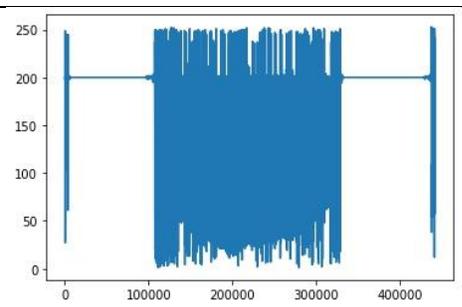
6			188.93	32.78
7			186.67	33.38
8			185.55	35.15
9			175.98	43.47
10			169.71	54.06

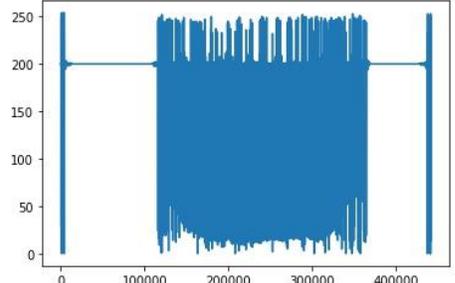
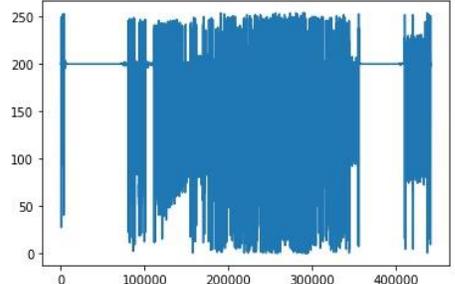
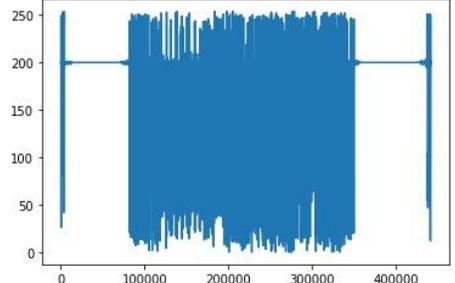
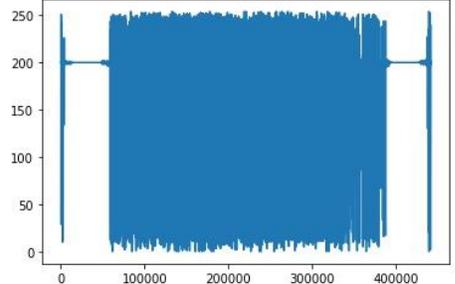
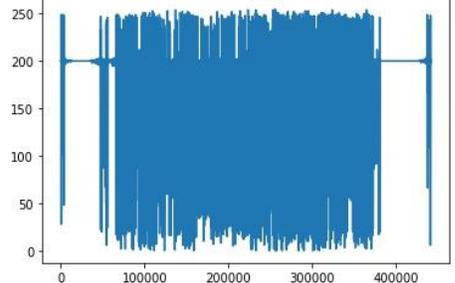
11			179.9	43.84
12			144.11	64.38
13			190.33	31.7
14			189.8	32.95
15			185.41	39.31

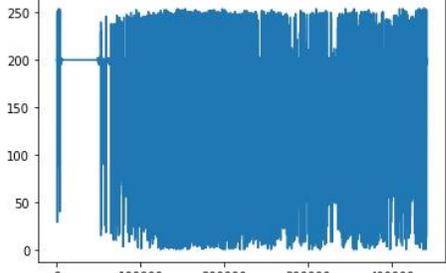
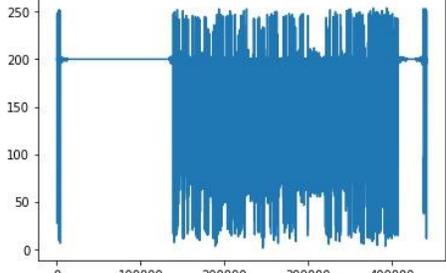
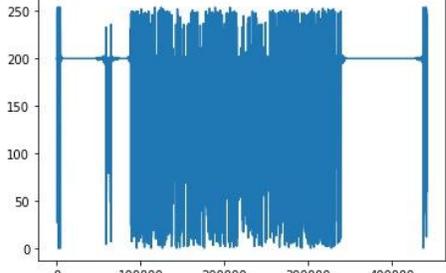
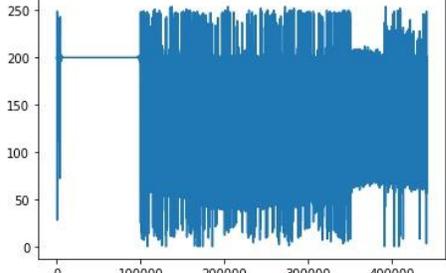
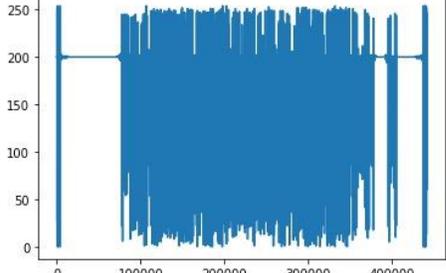
16			177.77	49.37
17			167.15	51.01
18			160.71	48
19			191.08	30.79
20			169.29	52.7

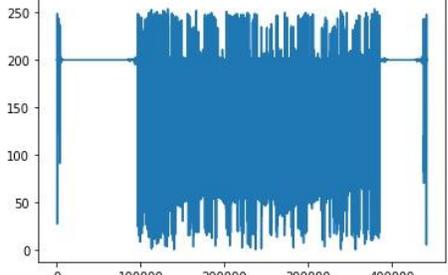
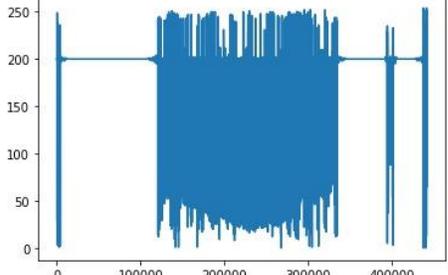
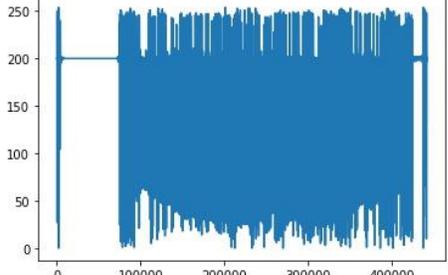
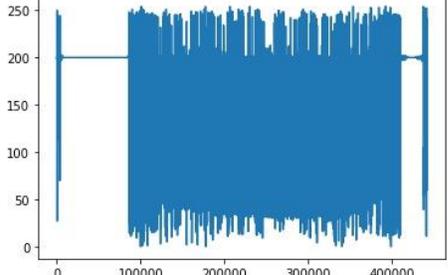
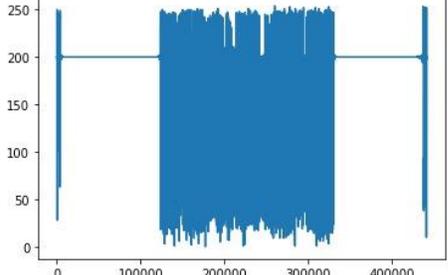
21			177.08	44.31
22			154.33	54.69
23			127.84	63.66
24			172.75	45.89
25			173.8	50.15

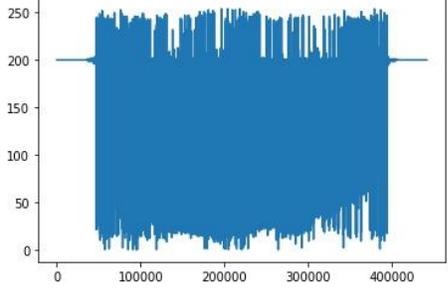
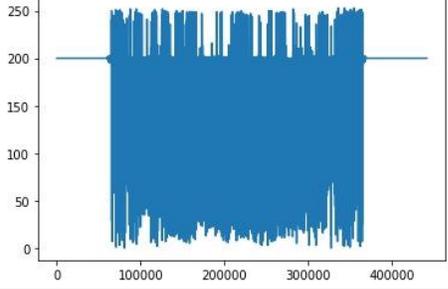
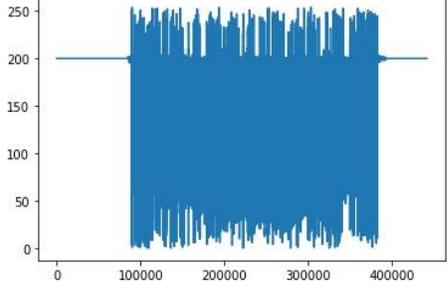
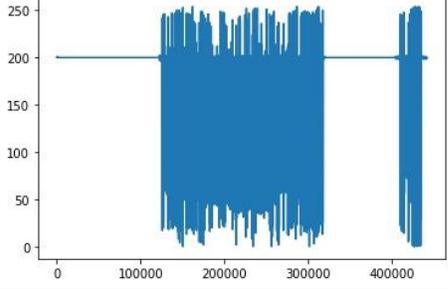
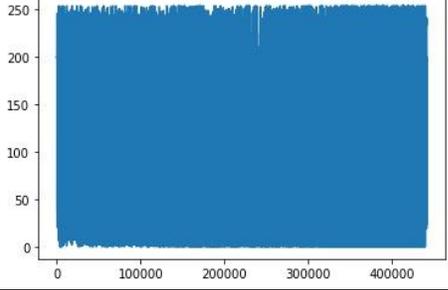
26			198.24	12.52
27			154.63	54.2
28			195.54	20.12
29			185.71	39.88
30			192.79	28.09

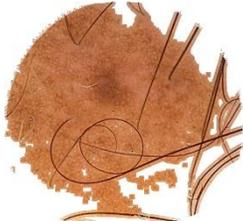
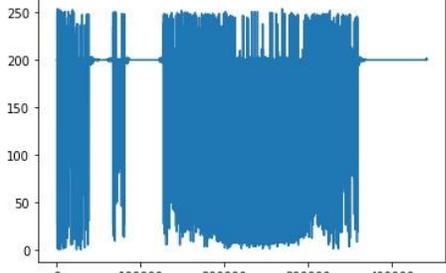
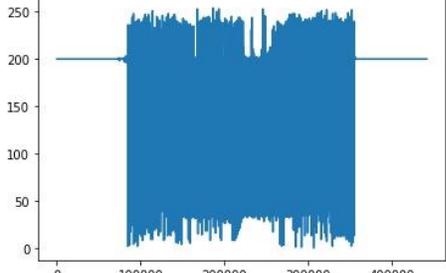
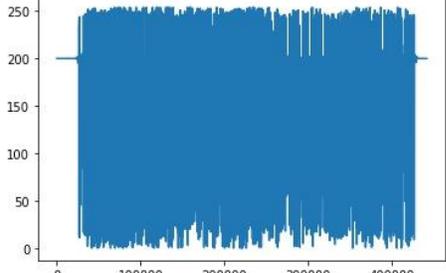
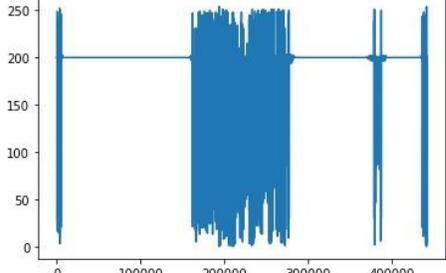
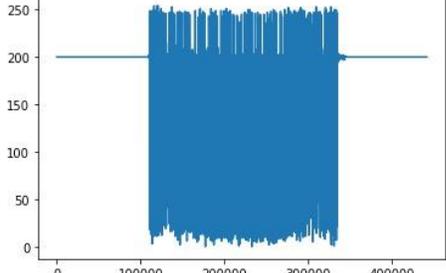
31			187.91	35.64
32			173.12	50.26
33			161.36	60.03
34			168.63	55.07
35			134.32	61.38

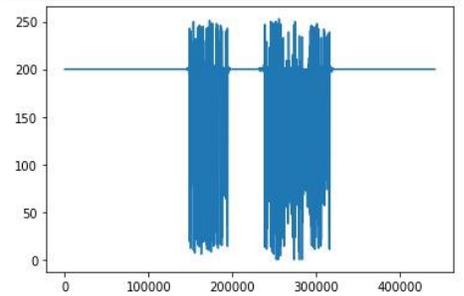
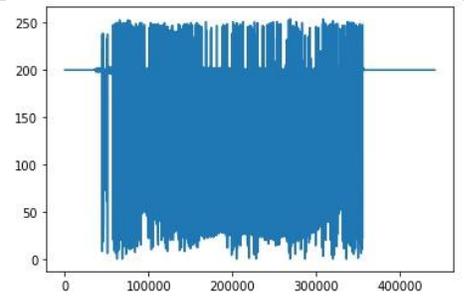
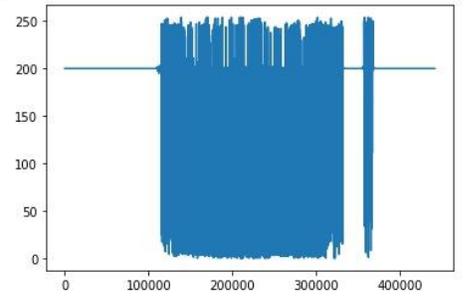
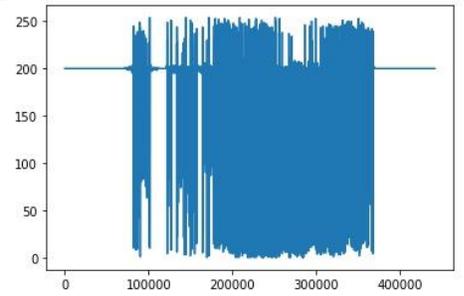
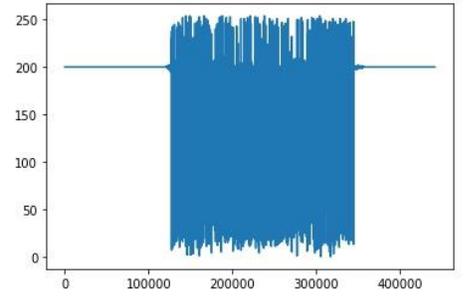
36			194.74	22.15
37			161.64	57.9
38			169.64	50.32
39			169.69	53.33
40			181.2	41.91

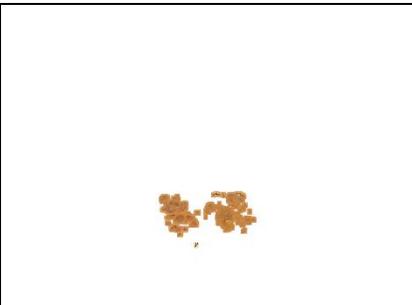
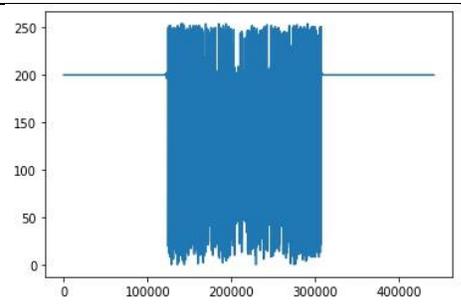
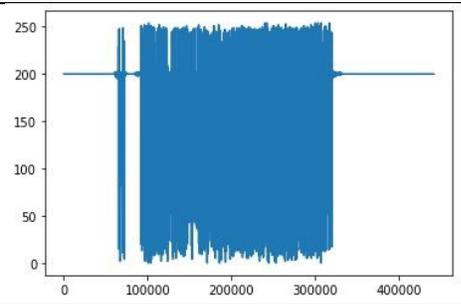
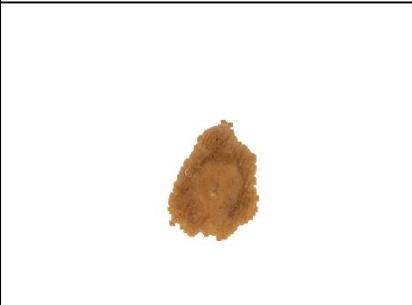
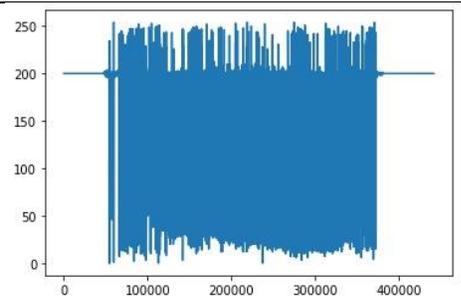
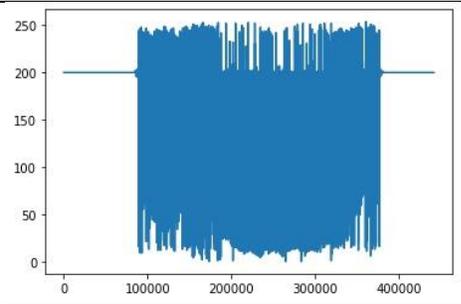
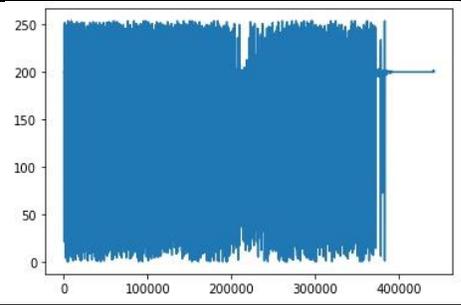
41			141.87	70.2
42			136.04	65.58
43			153.48	58.21
44			133.18	70.14
45			111.32	70.42

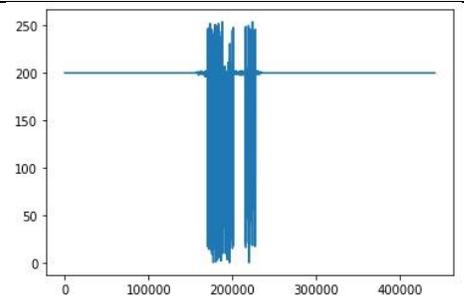
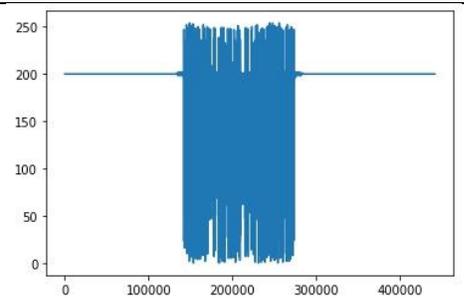
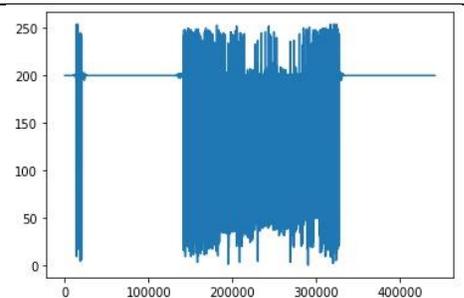
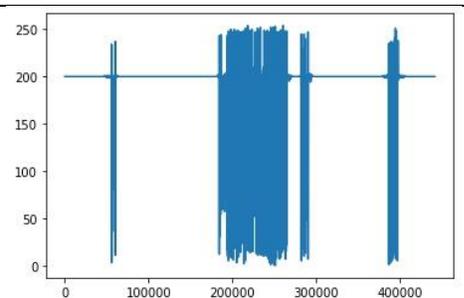
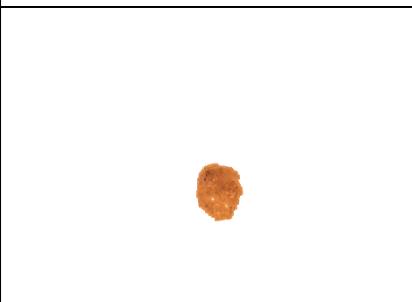
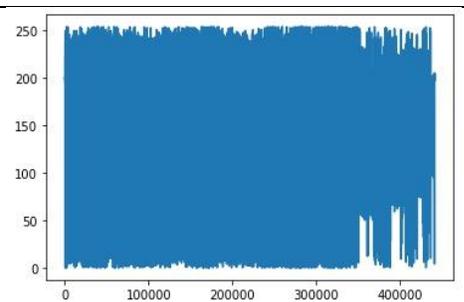
46			110.99	76.88
47			169.64	56.08
48			172.33	47.68
49			174.3	50.94
50			120.72	72.59

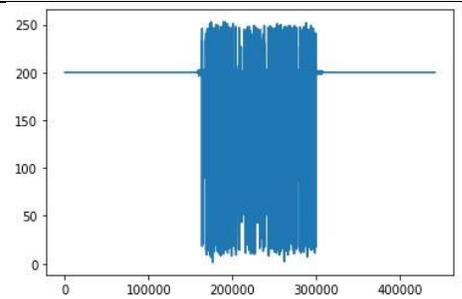
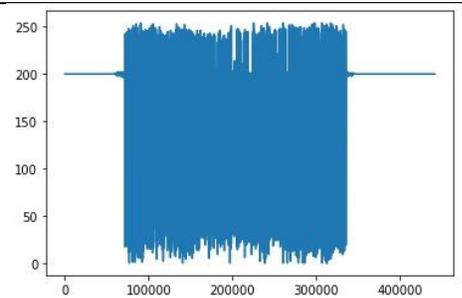
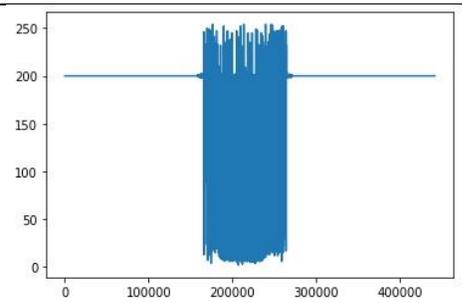
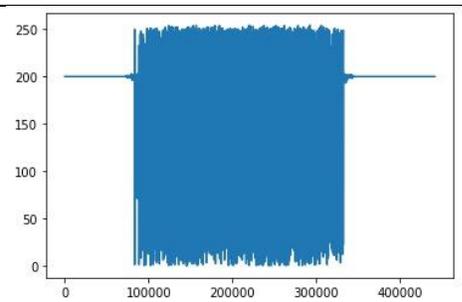
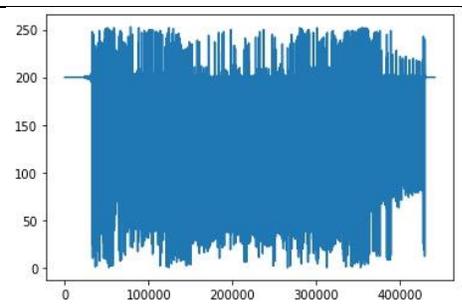
51			144.4	58.22
52			150.06	61.41
53			135.99	64.14
54			122.89	61.02
55			195.75	21.44

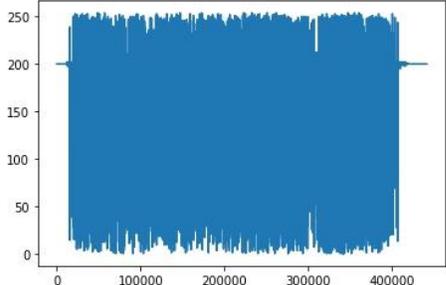
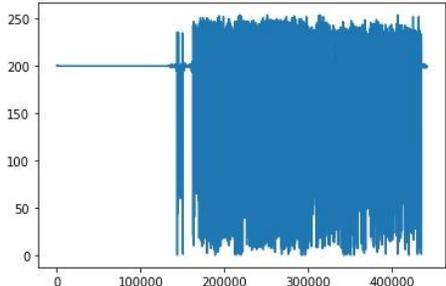
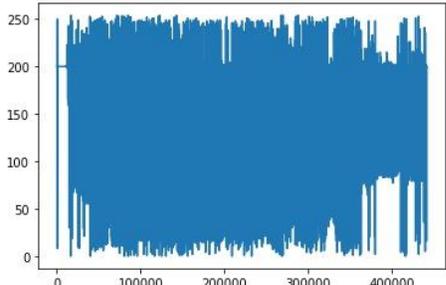
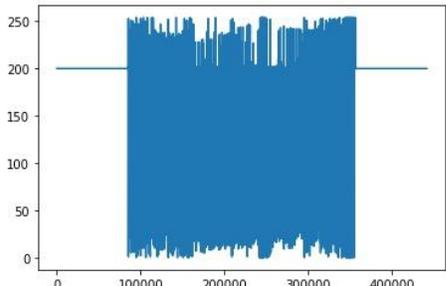
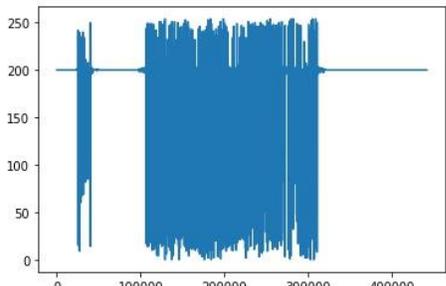
56			150.88	51.43
57			192.78	25.3
58			193.48	24.03
59			192.21	26.16
60			196.51	17.31

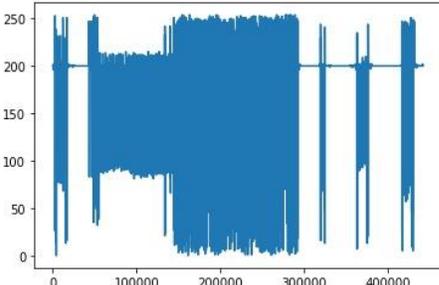
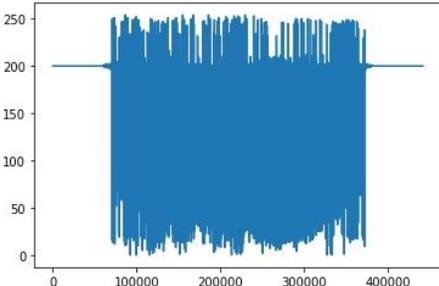
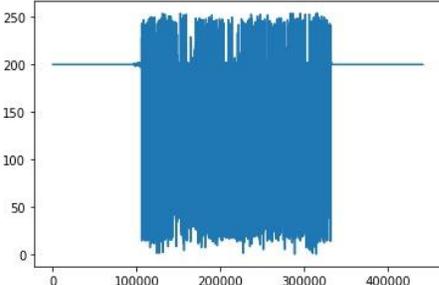
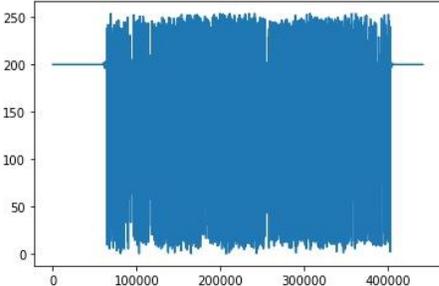
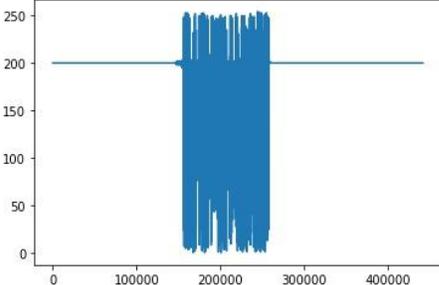
61			195.03	21.08
62			182.03	42.65
63			172.75	47.91
64			186.59	35.04
65			185.41	37.33

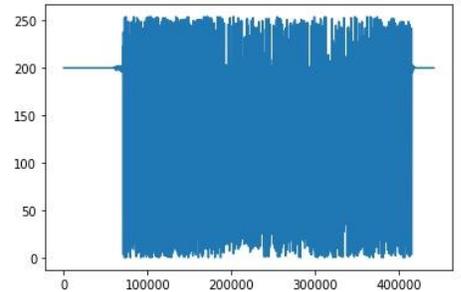
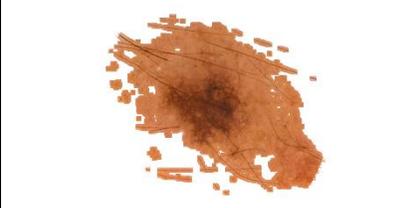
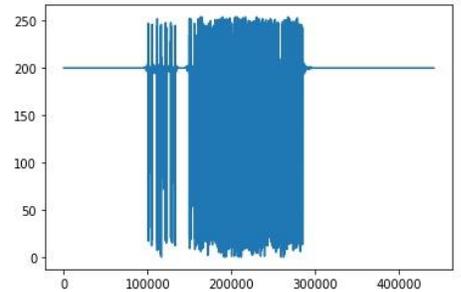
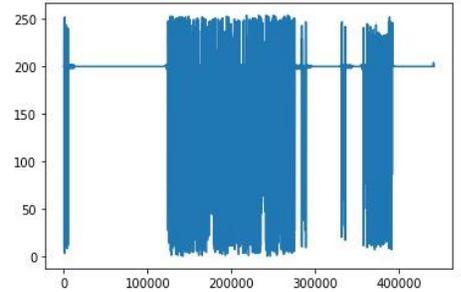
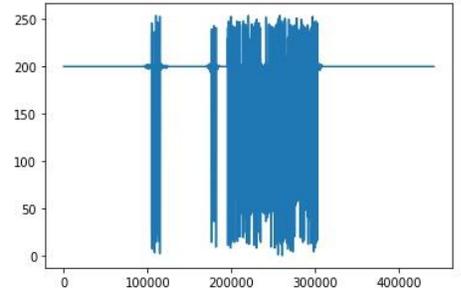
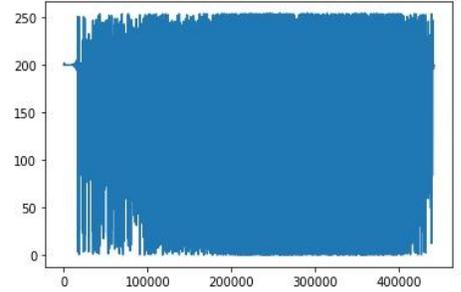
66			196.63	17.84
67			192.77	26.45
68			188.24	33.73
69			149.33	57.67
70			191.6	27.74

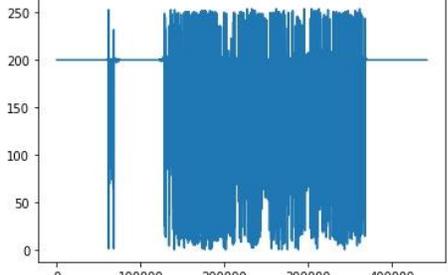
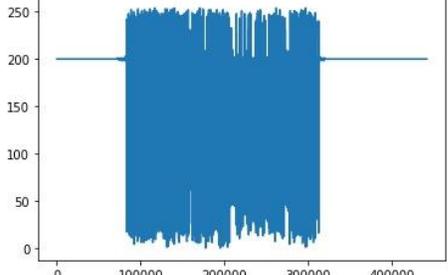
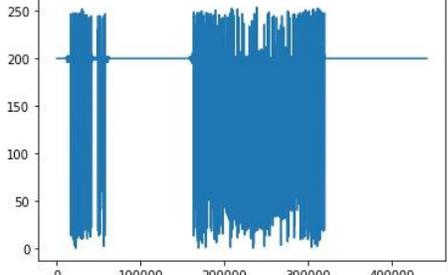
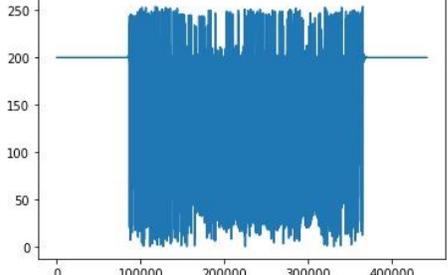
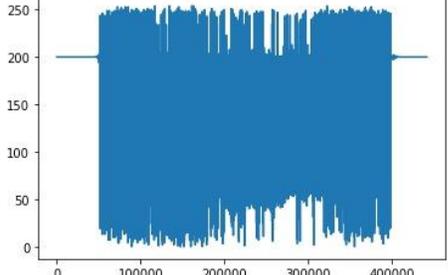
71			176.05	45.7
72			170.87	52.24
73			185.67	41.55
74			188.93	31.69
75			197.15	16.17

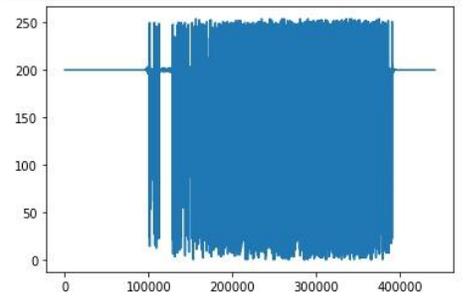
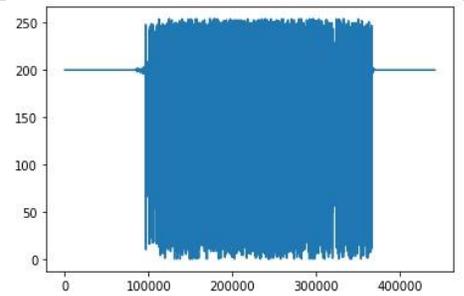
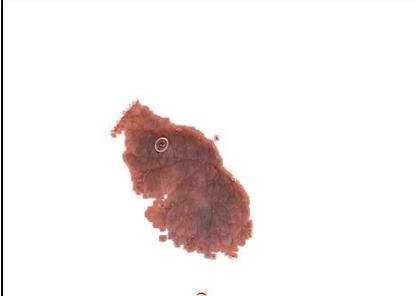
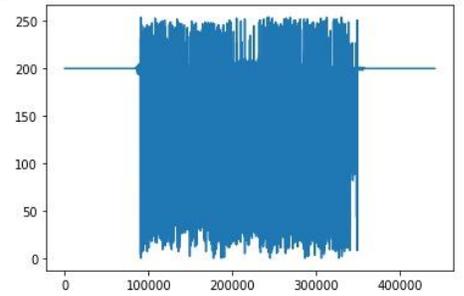
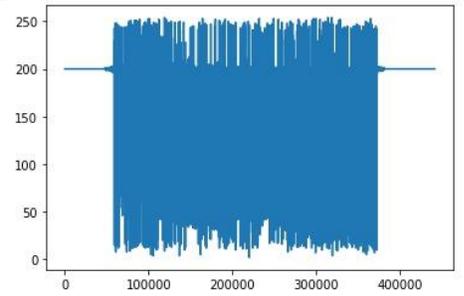
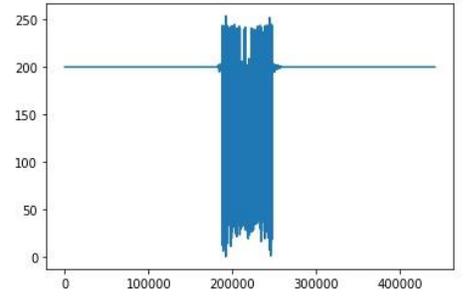
76			193.99	26.74
77			190.7	31.37
78			162.62	50.54
79			151.45	63.4
80			156.69	59.29

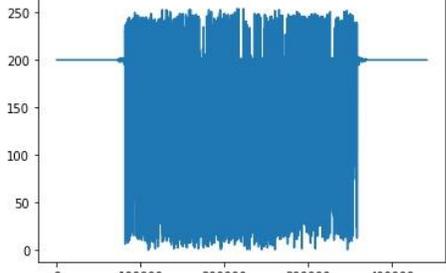
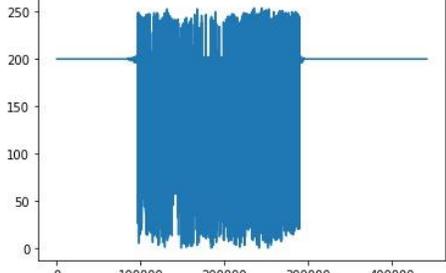
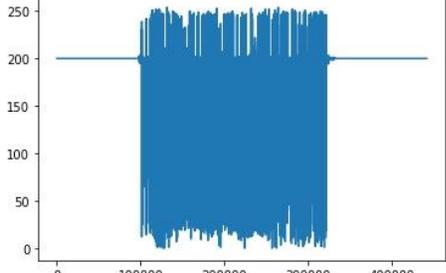
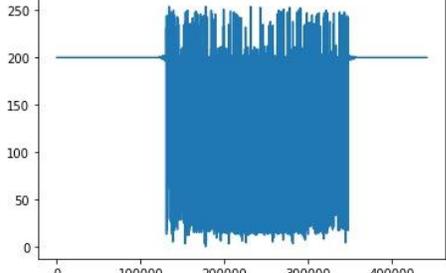
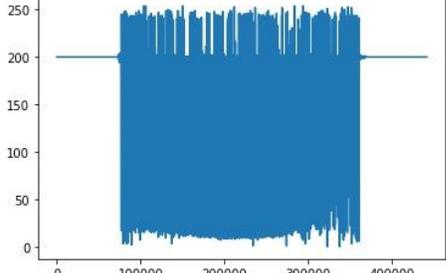
81			138.3	59.81
82			155.63	54.94
83			145.82	59.32
84			147.59	56.4
85			196.31	19.21

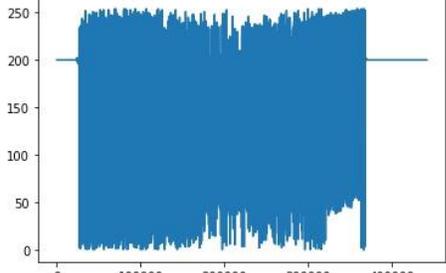
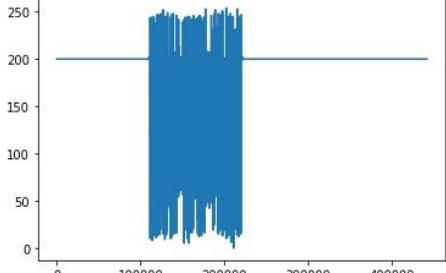
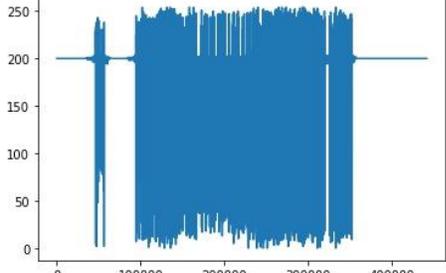
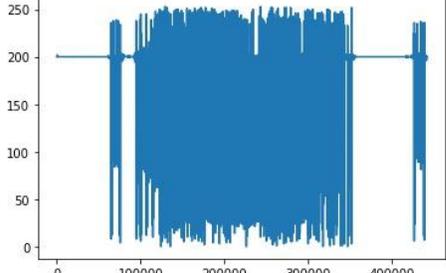
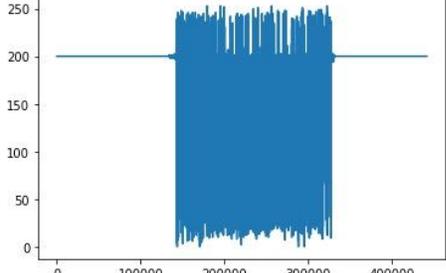
86			166.9	52.2
87			168.3	46.88
88			153.4	53.2
89			157.4	52.9
90			193.36	24.78

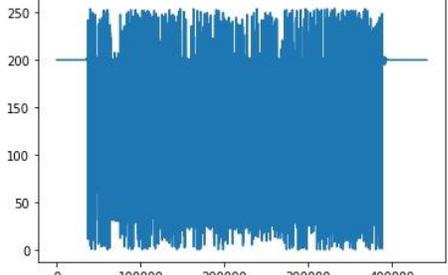
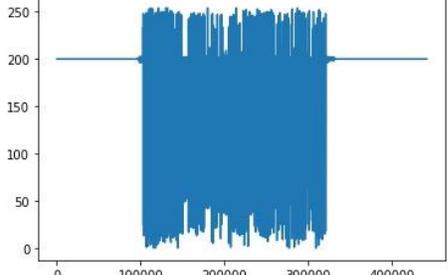
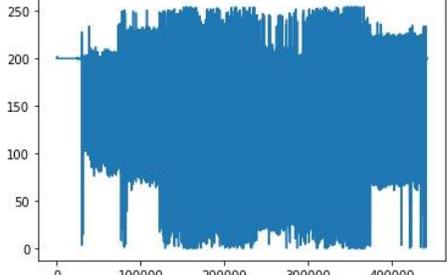
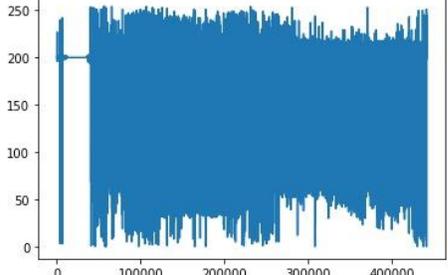
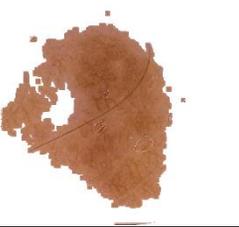
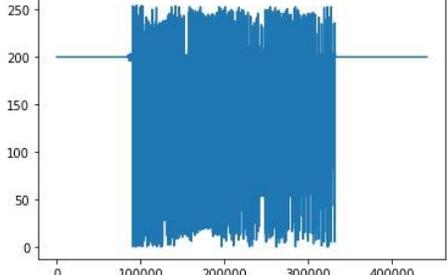
91			173.96	47.38
92			166.06	49.73
93			154.86	67.51
94			167.01	59.11
95			154.48	55.84

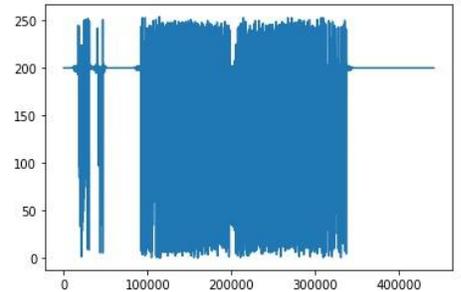
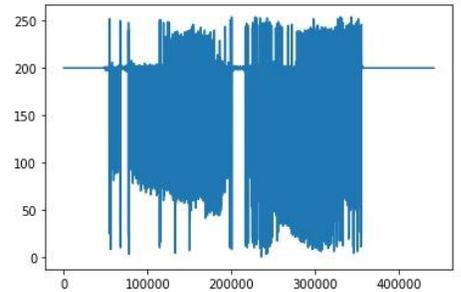
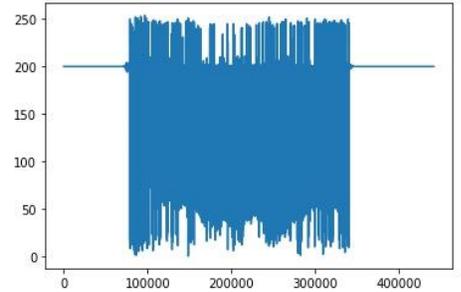
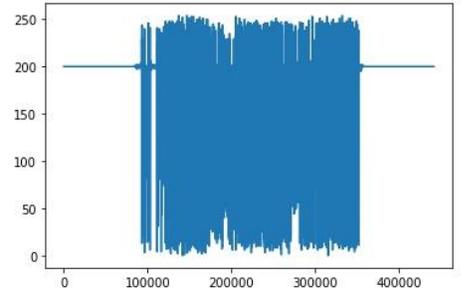
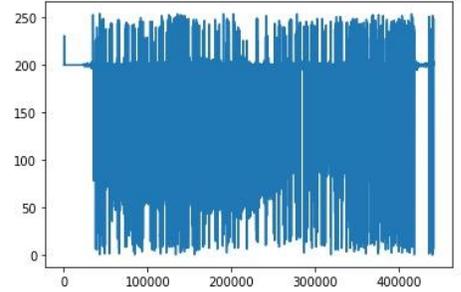
96			163.36	57.97
97			173.28	49.05
98			185.98	34.92
99			176.74	43.18
100			187.58	35.28

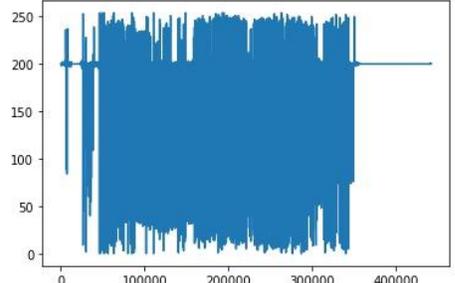
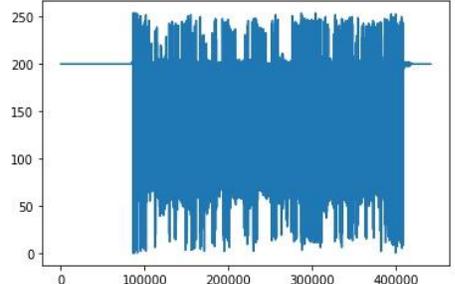
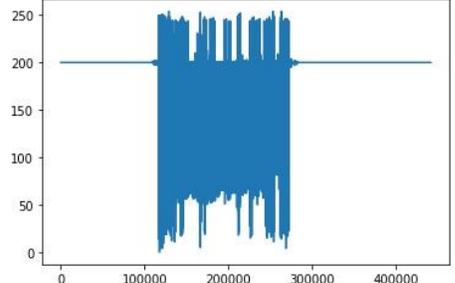
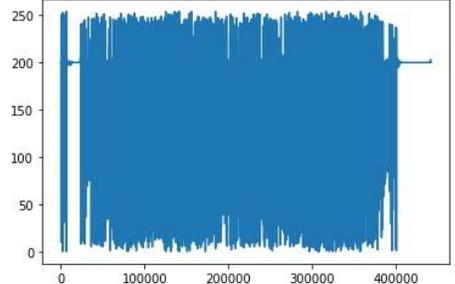
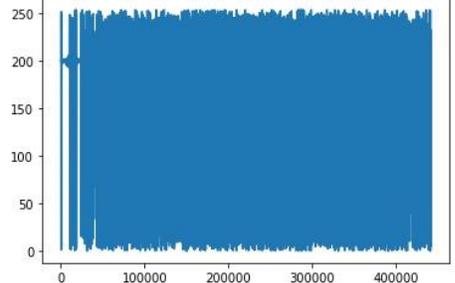
101			192.13	27.02
102			188.88	32.88
103			178.59	45.18
104			197.53	15.23
105			184.54	34.84

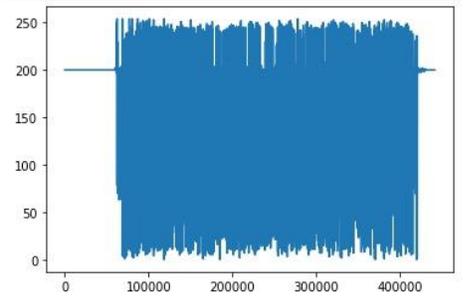
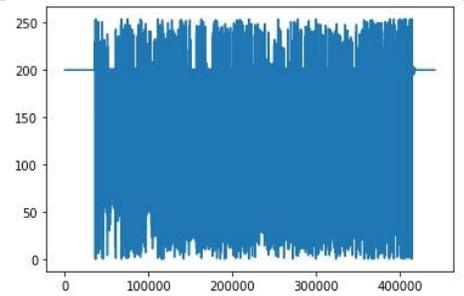
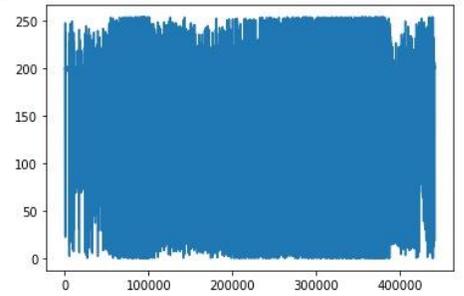
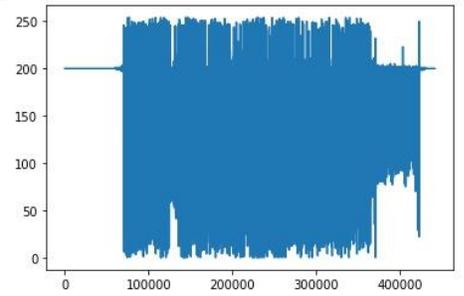
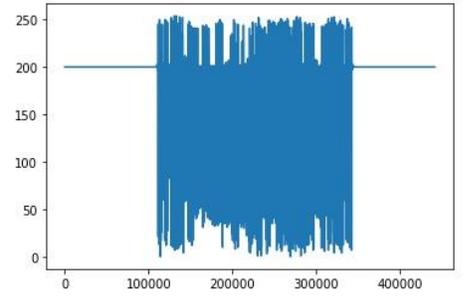
106			182.27	38.47
107			191.45	28.67
108			196.49	18.93
109			187.74	37.41
110			186.3	34.86

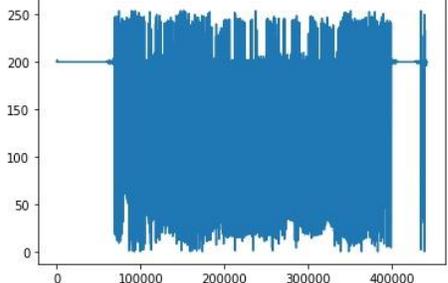
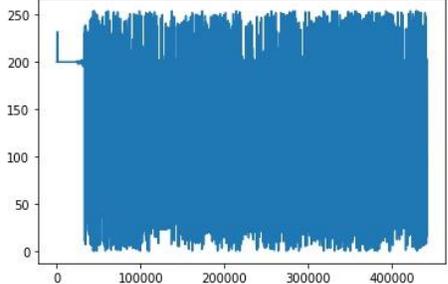
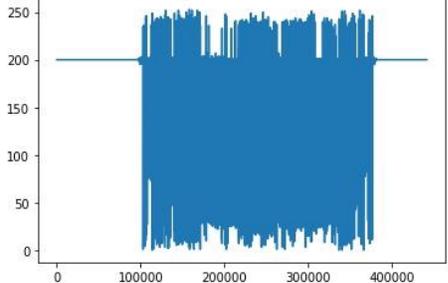
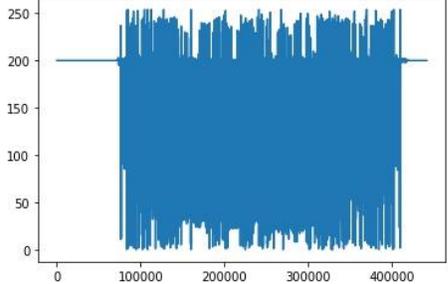
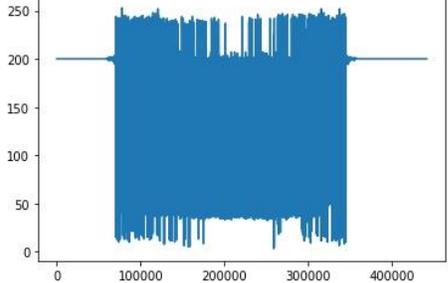
111			178.56	42.92
112			163.82	59.13
113			153.96	63.59
114			128.48	70.51
115			166.84	50.18

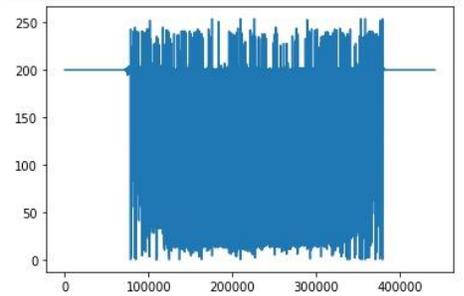
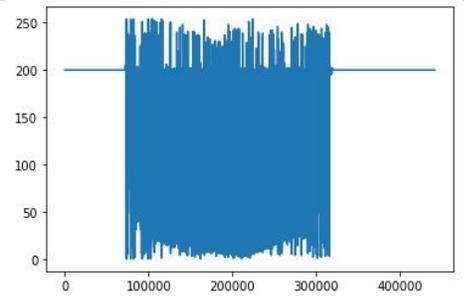
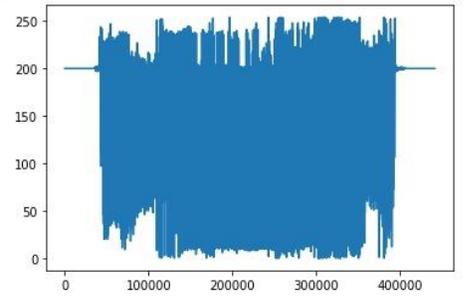
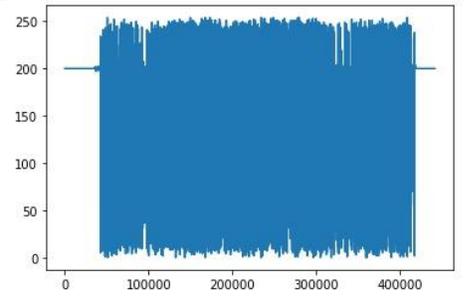
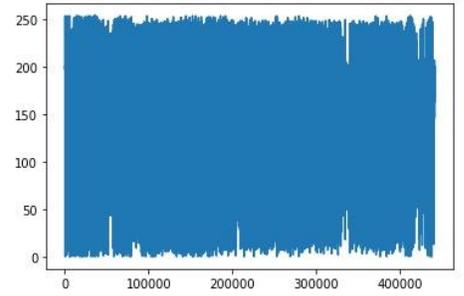
116			125.51	65.92
117			155.55	59.27
118			164.26	57
119			148.35	54.68
120			168.88	48.41

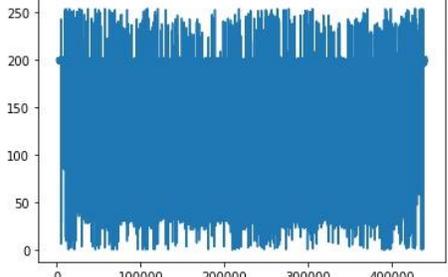
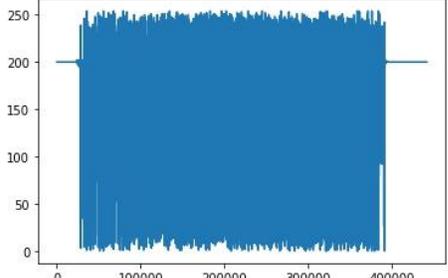
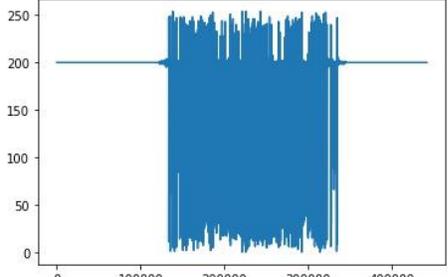
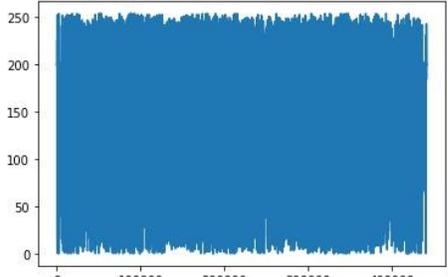
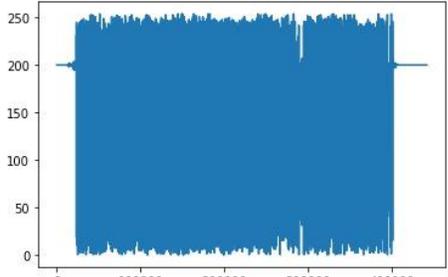
121			165.02	59.21
122			191.1	28.93
123			183.32	38.77
124			195.96	19.52
125			178.93	42.76

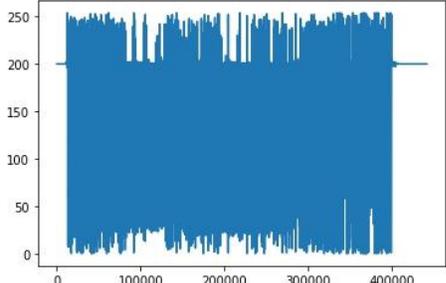
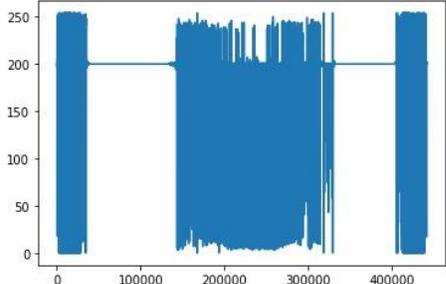
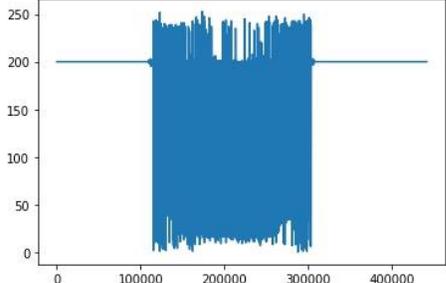
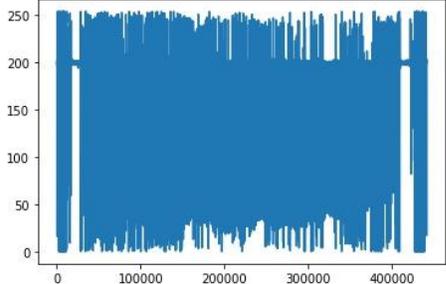
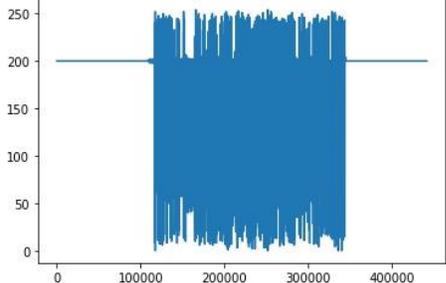
126			162.46	51.53
127			181.55	41.61
128			173.87	47.4
129			179.33	44.57
130			183.87	39.75

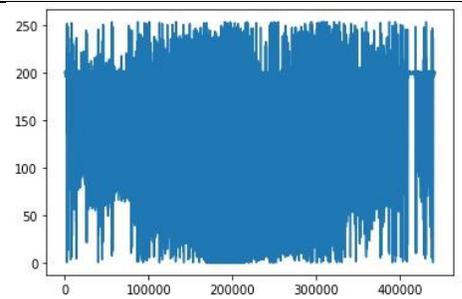
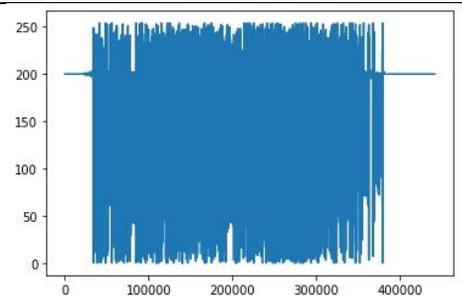
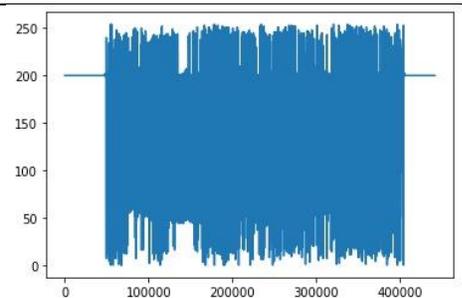
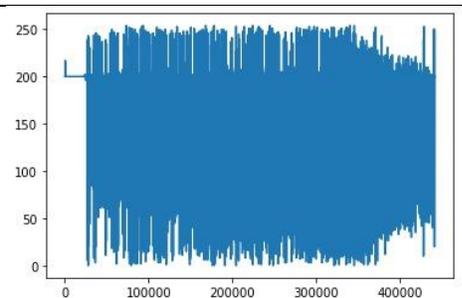
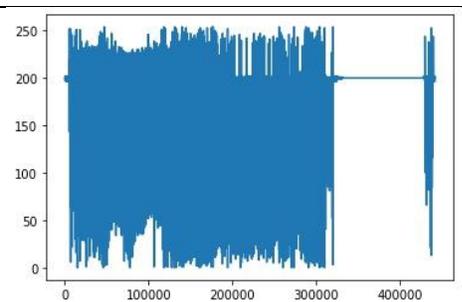
131			191.87	31.26
132			187.77	37.32
133			184.66	38.85
134			143.69	59.35
135			107.73	75.38

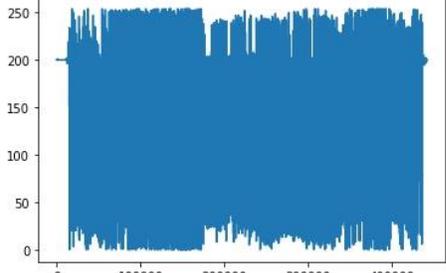
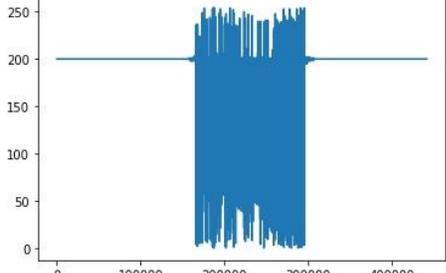
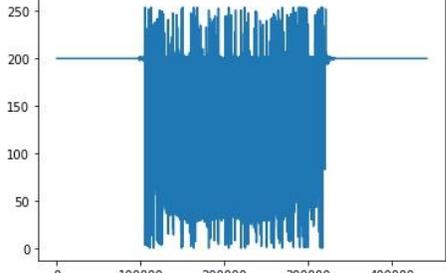
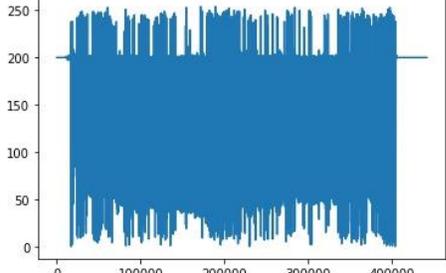
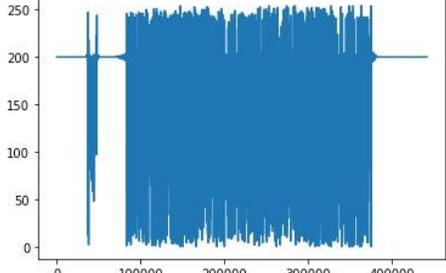
136			134.84	63.99
137			187.46	33.56
138			188.8	34.89
139			191.77	28.66
140			192.76	26.17

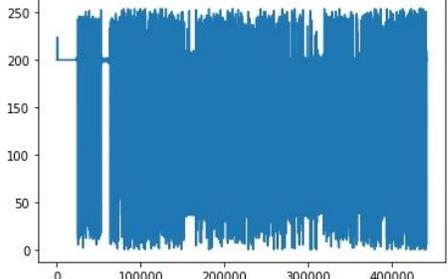
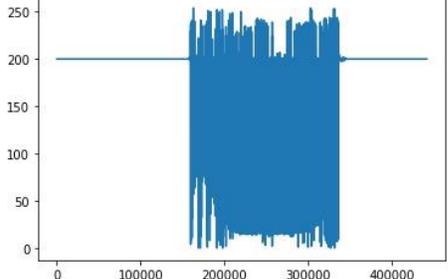
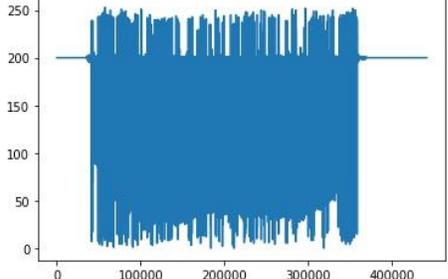
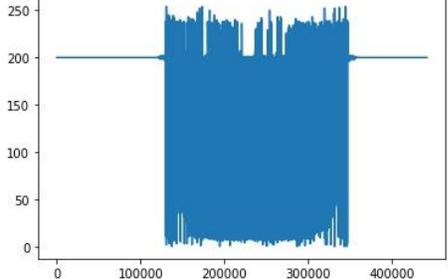
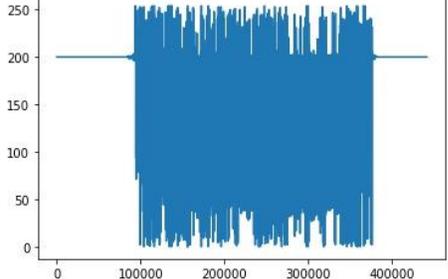
141			195.52	20.91
142			195.81	19.94
143			173.85	52.16
144			178.51	42.52
145			144.06	60.08

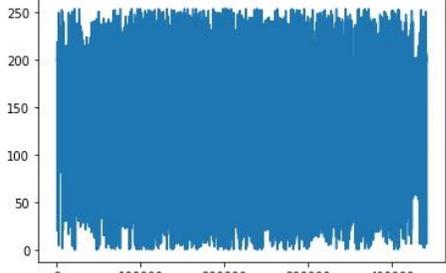
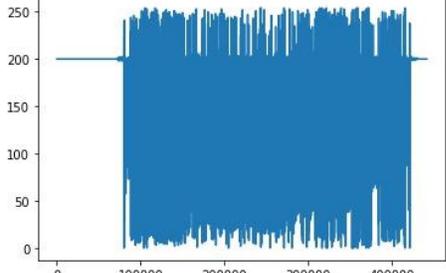
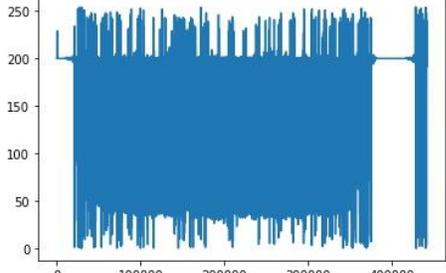
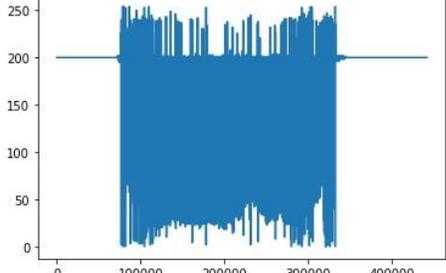
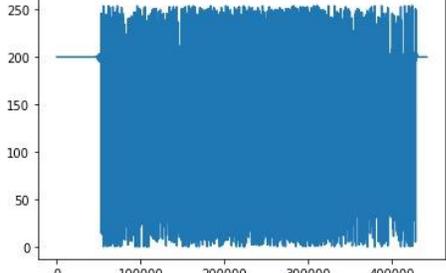
146			180.34	43.52
147			190.71	30.6
148			189.58	31.89
149			182.17	44.05
150			187.8	33.75

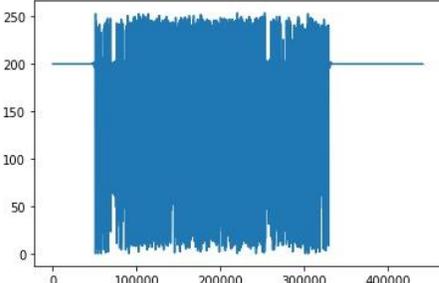
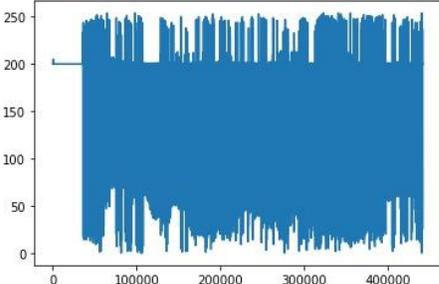
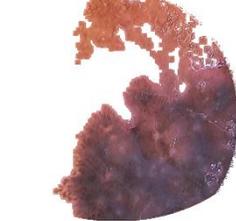
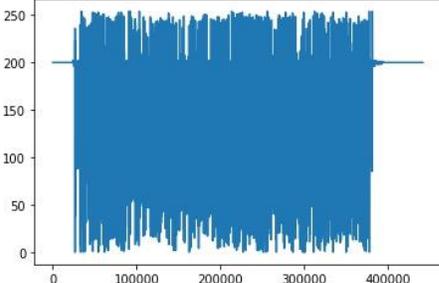
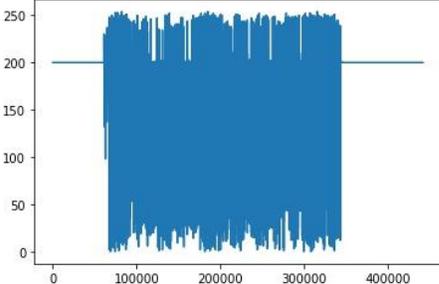
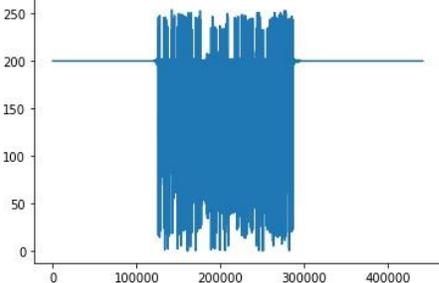
151			177.87	46.12
152			188.39	34
153			189.03	30
154			180.13	43.45
155			194	23.04

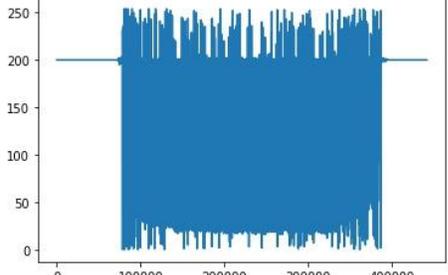
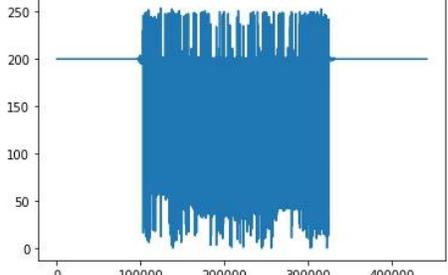
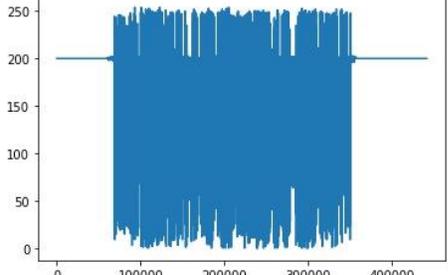
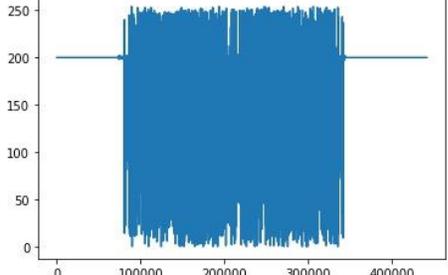
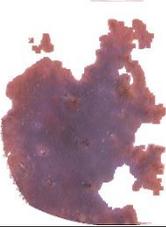
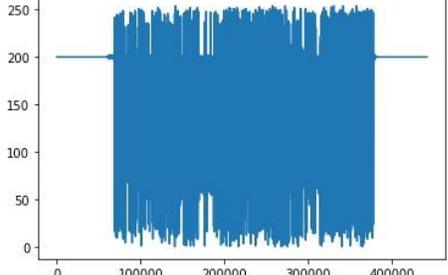
156			186.79	35.98
157			188.26	32.95
158			180.83	48.96
159			163.13	51.34
160			151.22	55.13

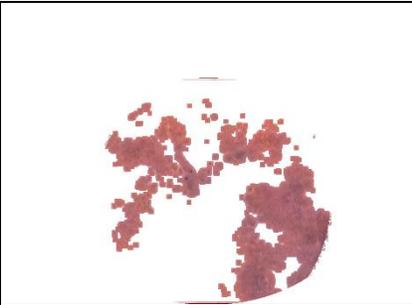
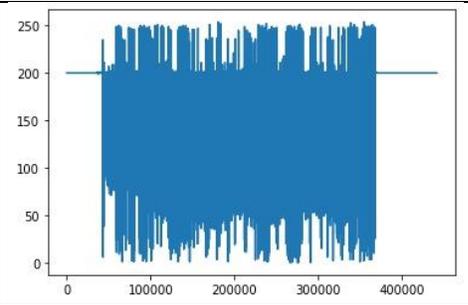
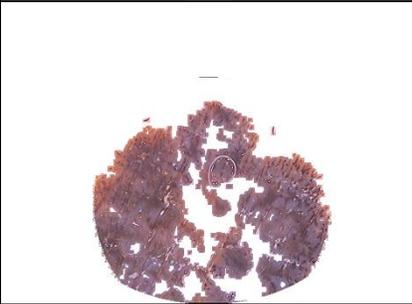
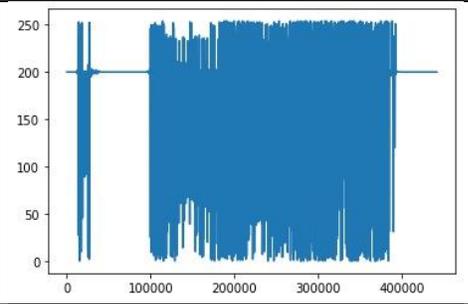
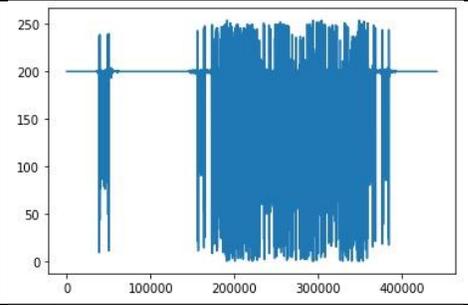
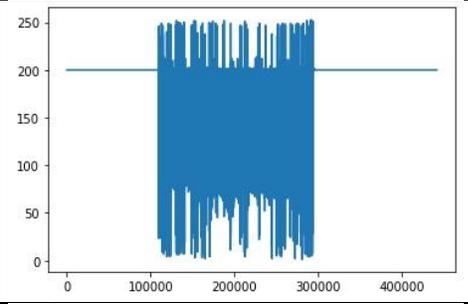
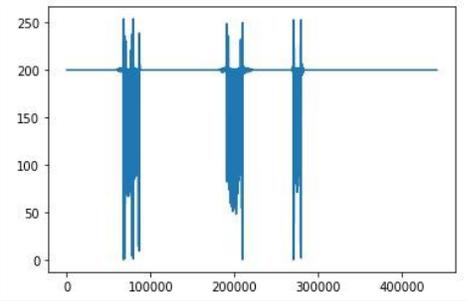
161			193.99	22.68
162			194.7	21.99
163			166.39	48.6
164			182.8	42.31
165			167.19	53.37

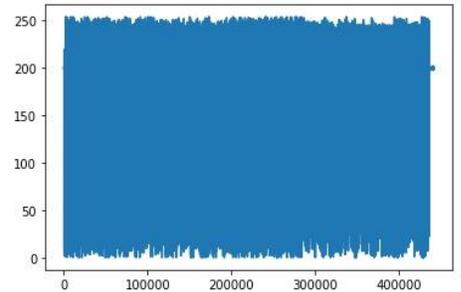
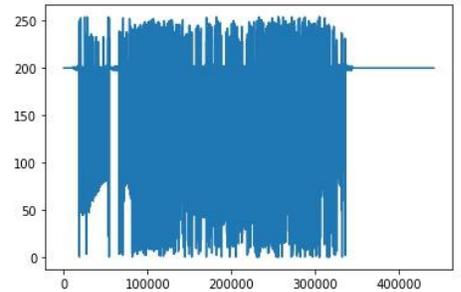
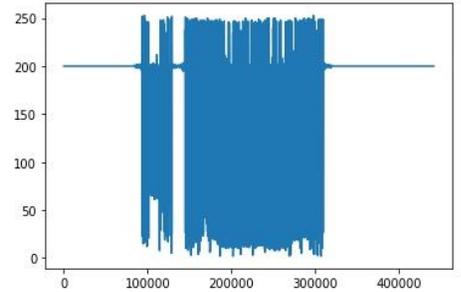
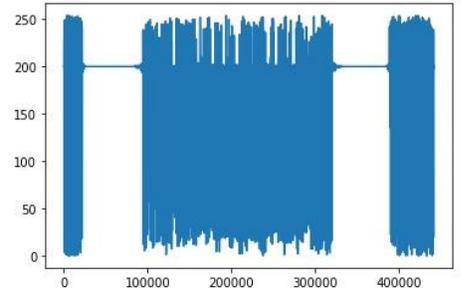
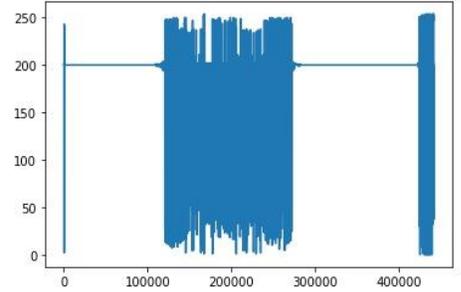
166			171.63	52.18
167			177.99	43.47
168			172.49	47.05
169			158.71	56.8
170			167.84	48.69

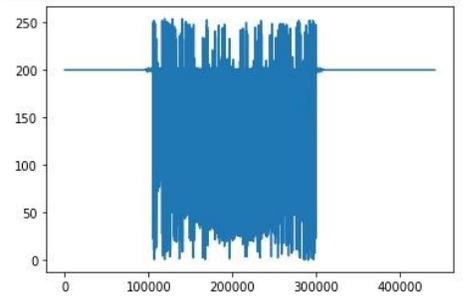
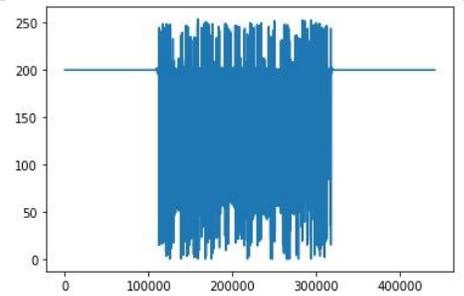
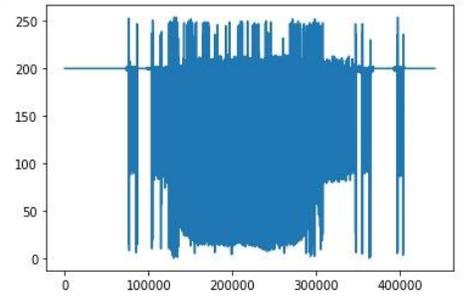
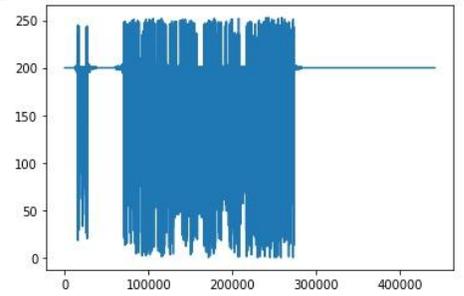
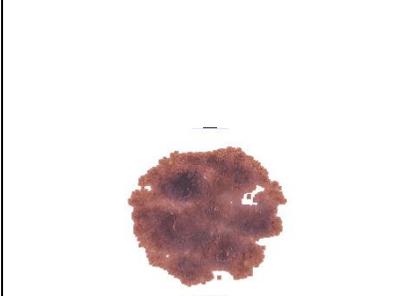
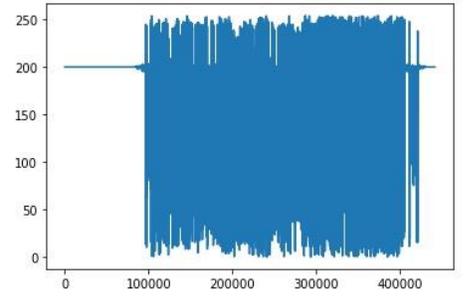
171			173.43	50.02
172			136.36	59.06
173			145.98	58.48
174			180.68	44.82
175			157.36	59.91

176			158.82	63.61
177			157.11	53.61
178			152.82	60.37
179			175.99	44.18
180			102.21	75.53

181			128.54	60.93
182			128.13	56.56
183			138.64	64.6
184			114.71	65.46
185			162.52	54.49

186			173.43	45.51
187			153.78	55.37
188			169.21	52.44
189			168.71	51.63
190			178.53	43.03

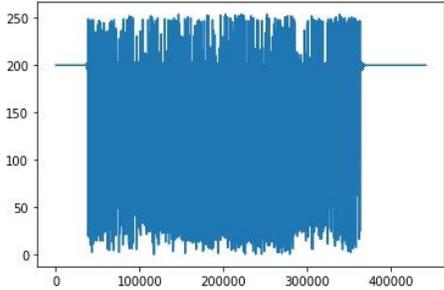
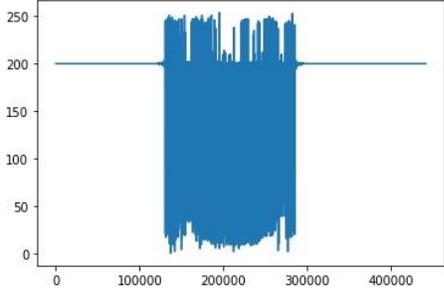
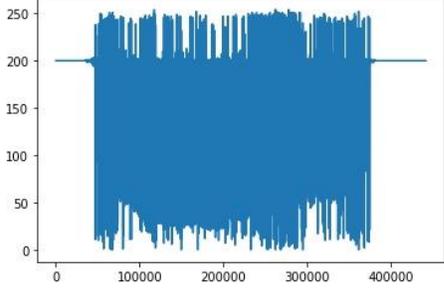
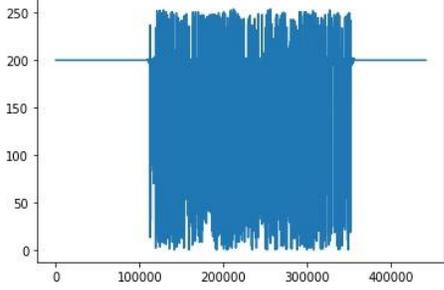
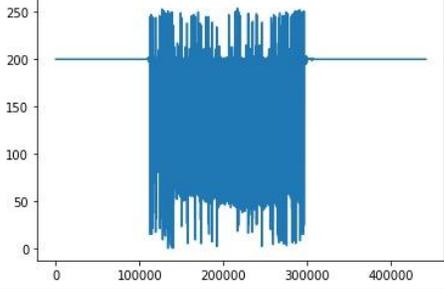
191			167.08	52.96
192			178.43	47.33
193			177.11	49.77
194			163	54.92
195			183.46	40.65

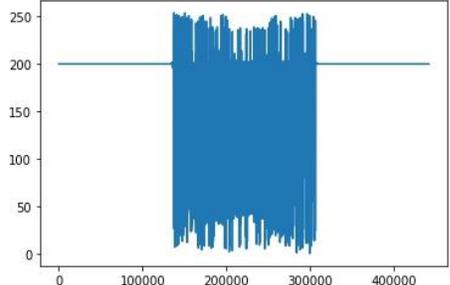
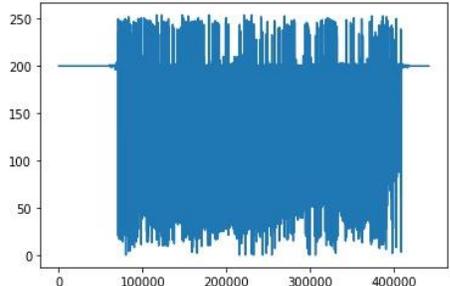
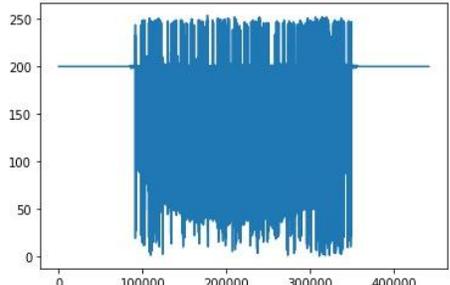
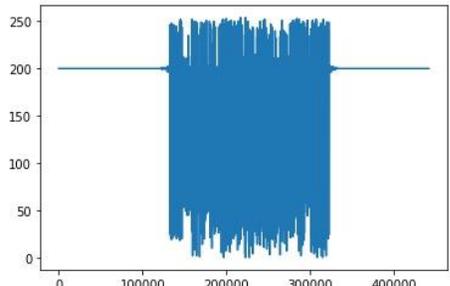
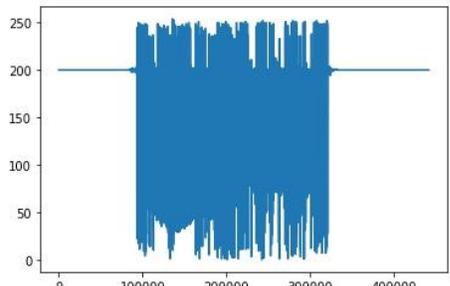
196			192.37	28.06
197			179.73	44.02
198			167.81	49.43
199			165.13	50.38
200			153.17	57.24

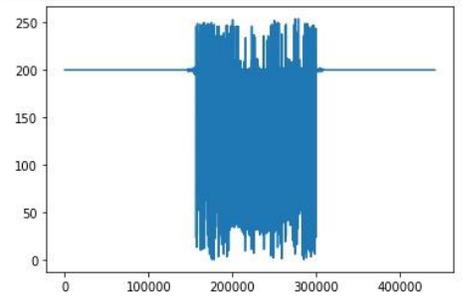
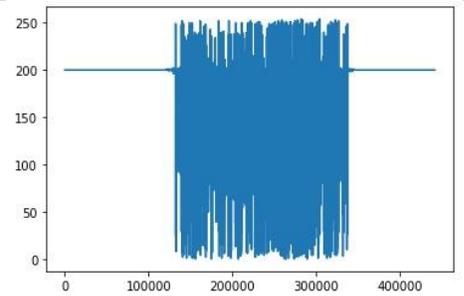
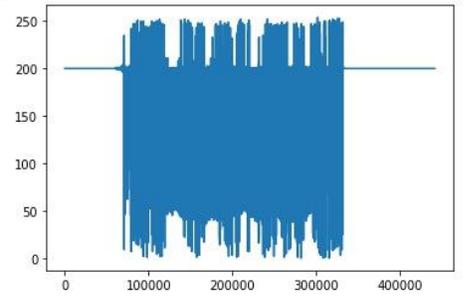
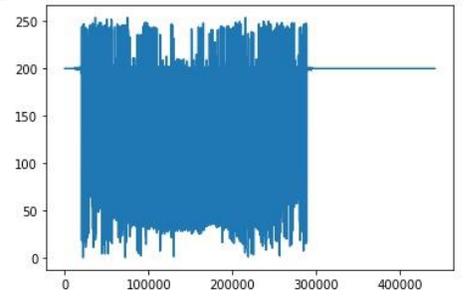
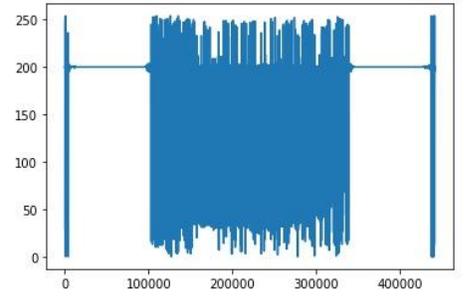
## Apéndice C

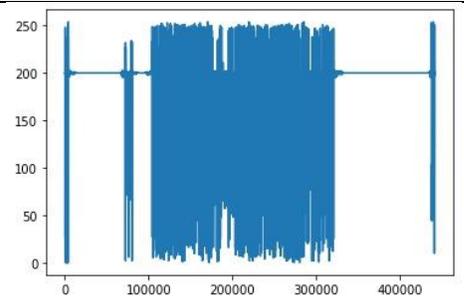
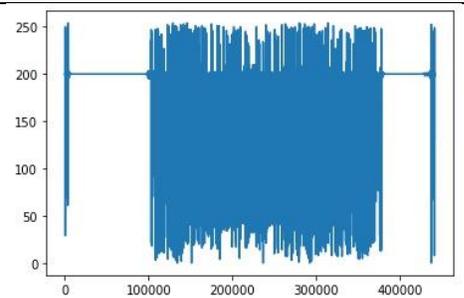
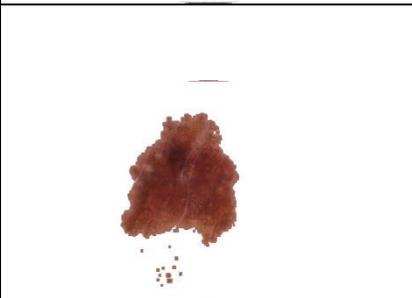
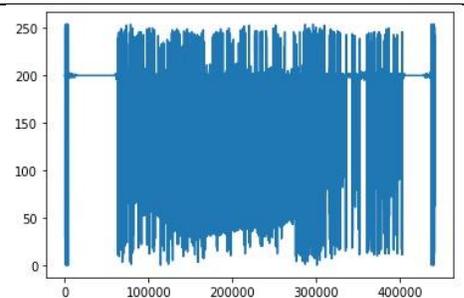
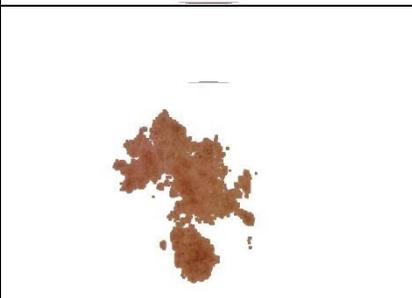
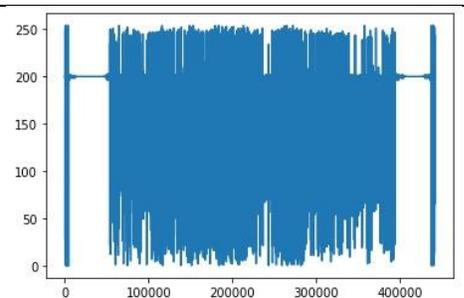
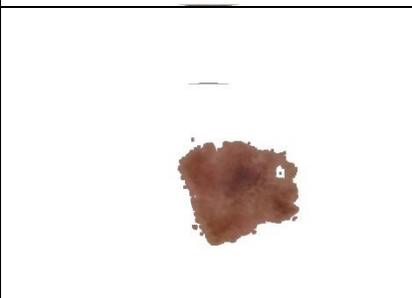
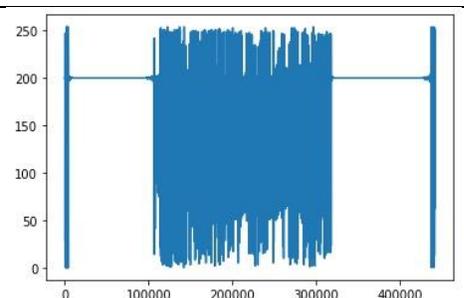
### Gráficas de valores de intensidad clase Benignos

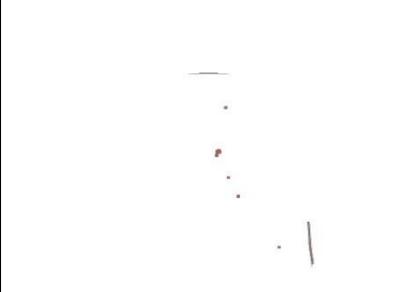
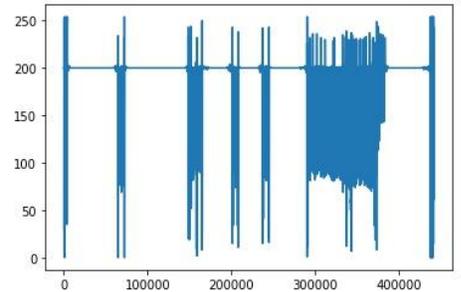
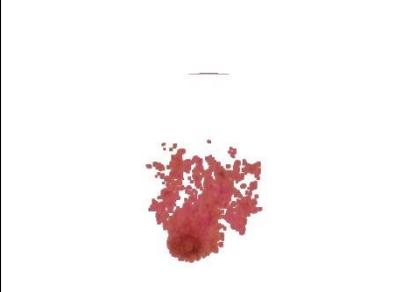
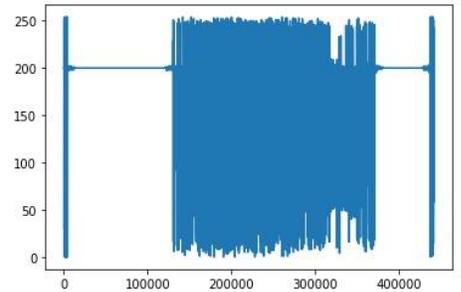
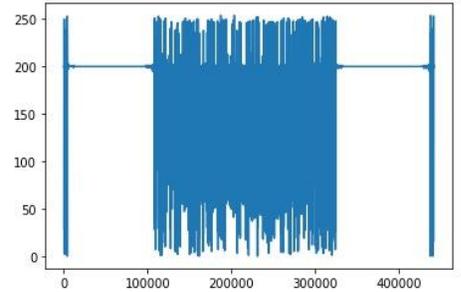
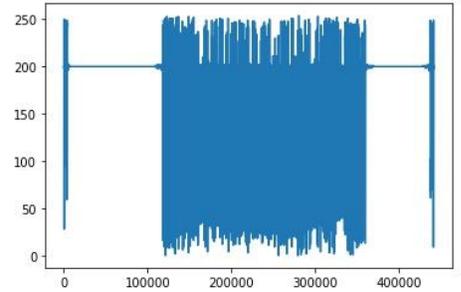
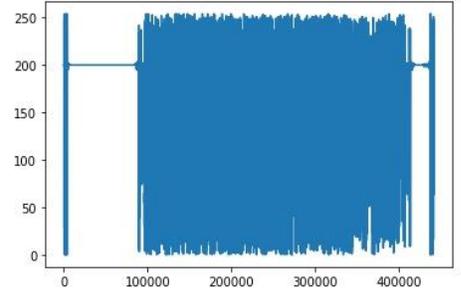
NEOPLASIAS BENIGNAS

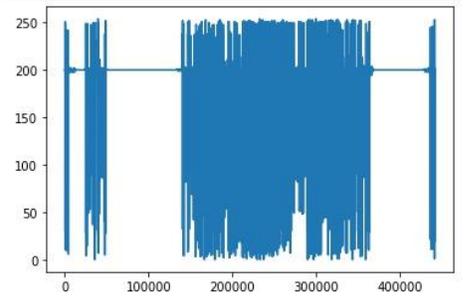
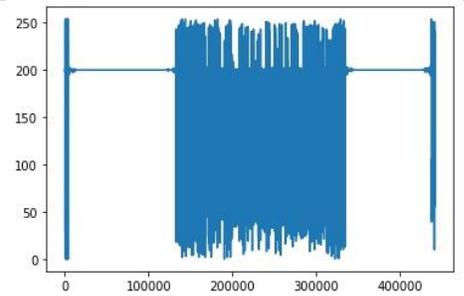
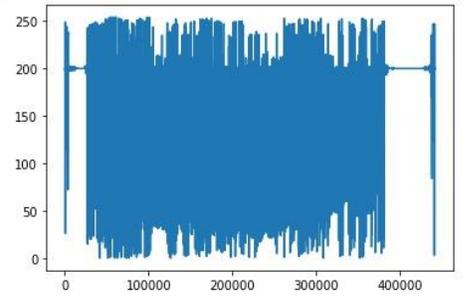
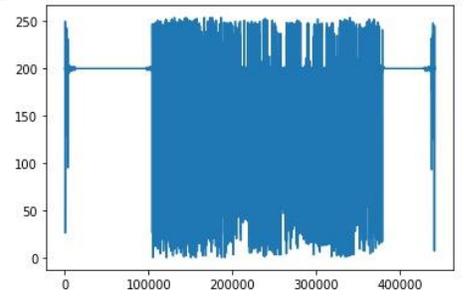
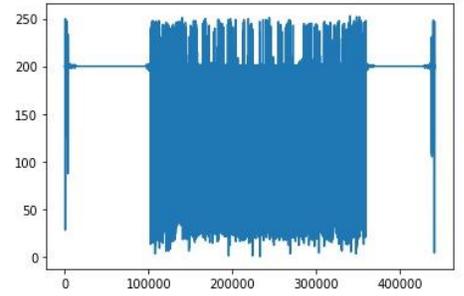
No.	Imagen	Gráfica	Media	Desviación
1			142	72.81
2			188.24	40.65
3			166.18	57.84
4			183.02	43.04
5			191.21	32.33

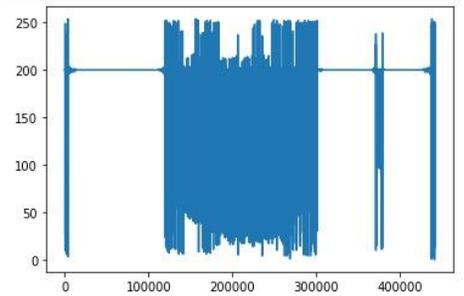
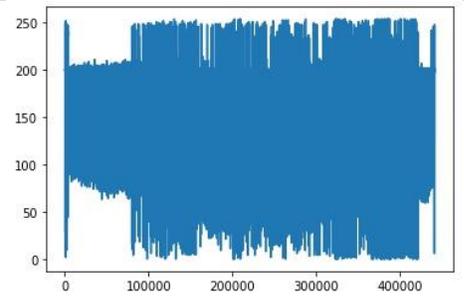
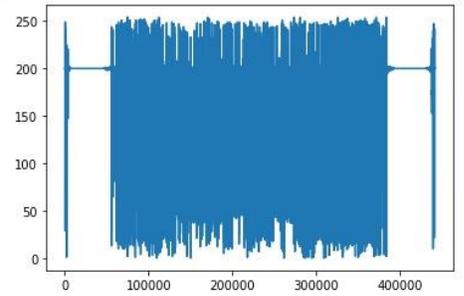
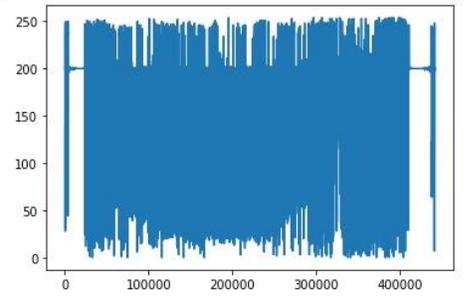
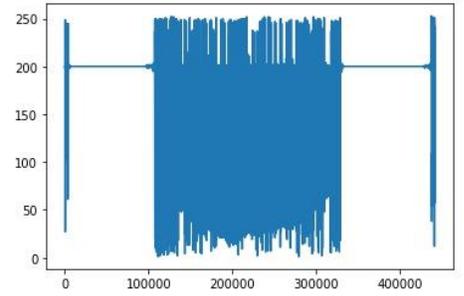
6			189.04	36.84
7			168.52	55.7
8			176.79	49.21
9			190.3	34.36
10			183.37	42.36

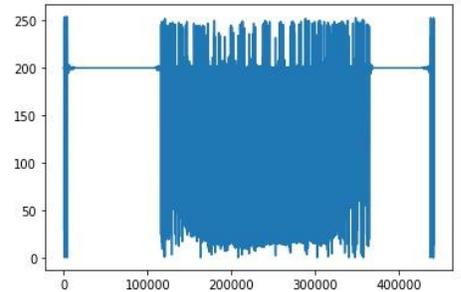
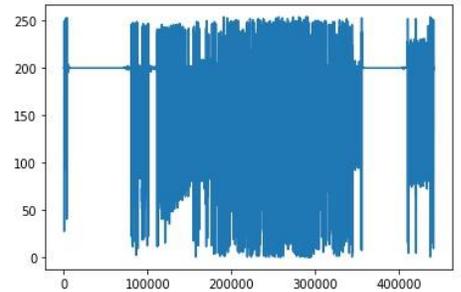
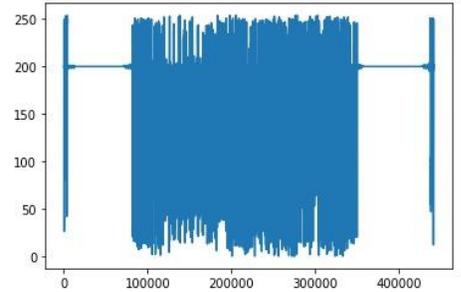
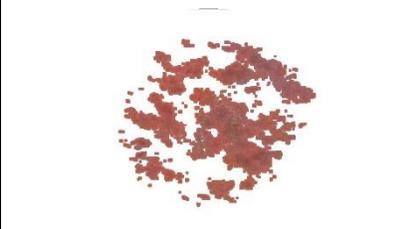
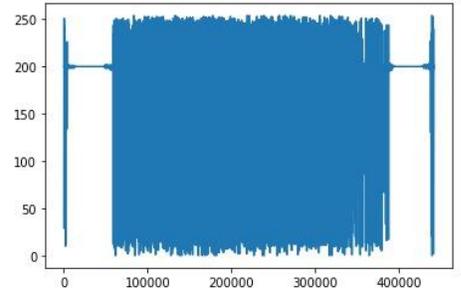
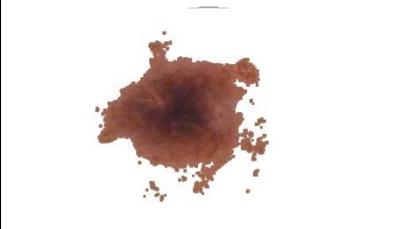
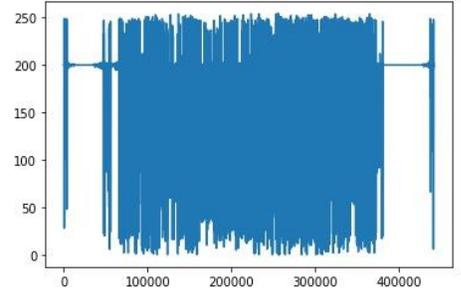
11			192.24	31.39
12			189.42	33.25
13			181.13	46.07
14			173.52	56.24
15			181.3	47.33

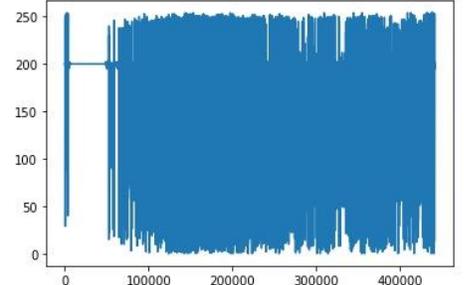
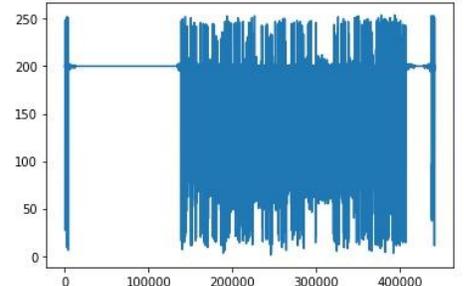
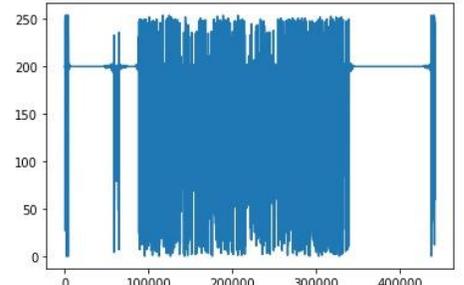
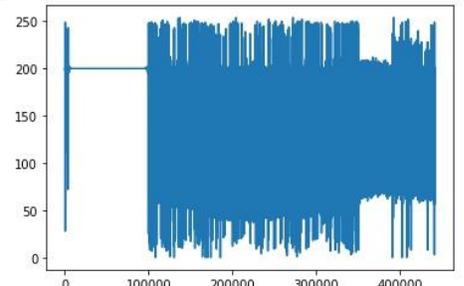
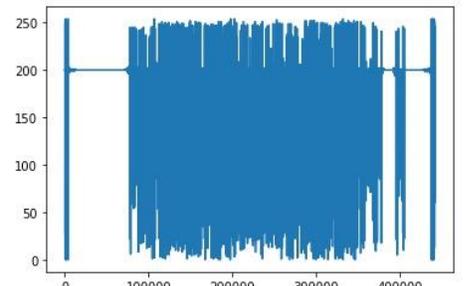
16			184.17	41.06
17			167.02	58.18
18			177.79	50.69
19			182.11	41.86
20			183.13	42.75

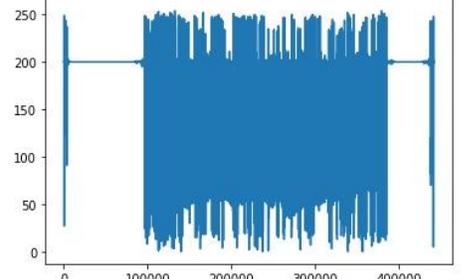
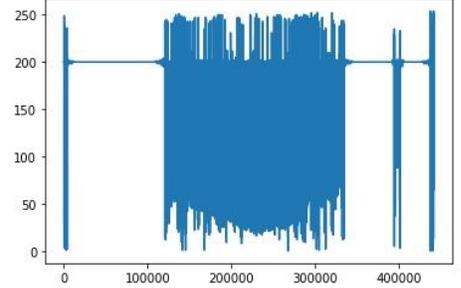
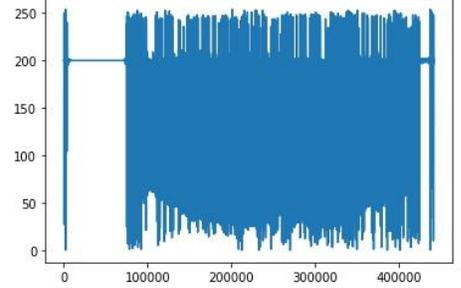
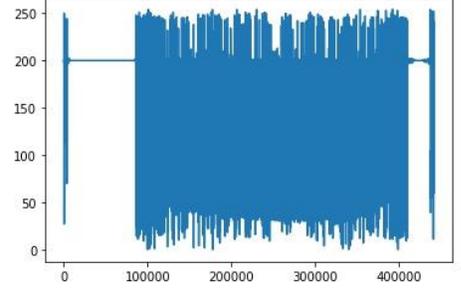
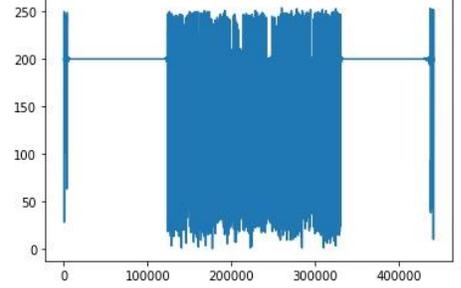
21			199.47	8.15
22			186.58	37.44
23			186.55	38.42
24			170.36	58.93
25			177.18	46.77

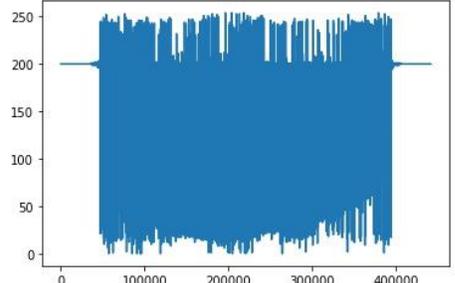
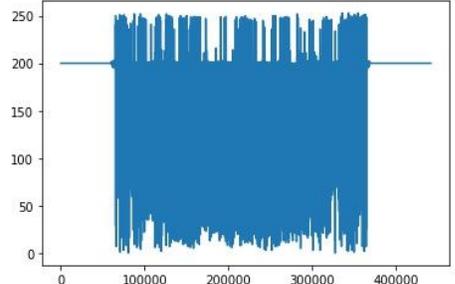
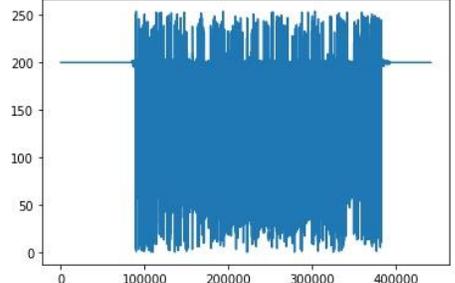
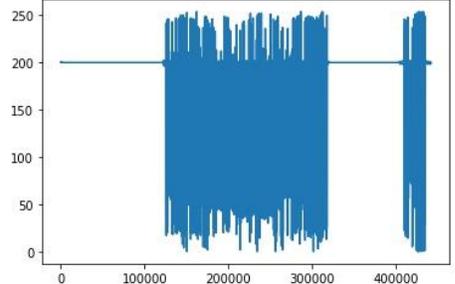
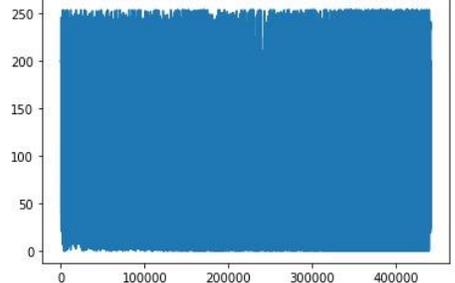
26			195.4	23.34
27			173.73	54.34
28			163.56	61.43
29			162.76	59.28
30			170.07	60.39

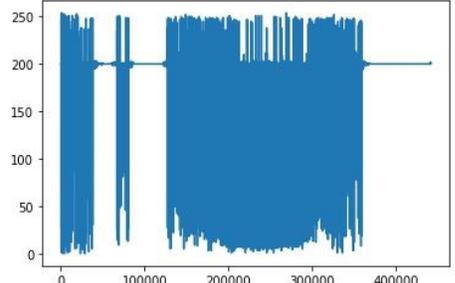
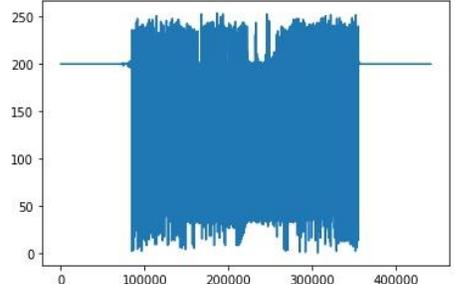
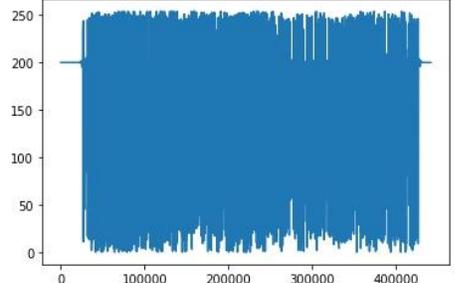
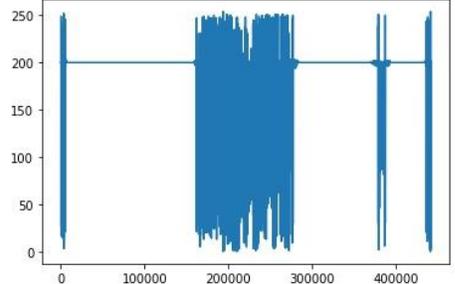
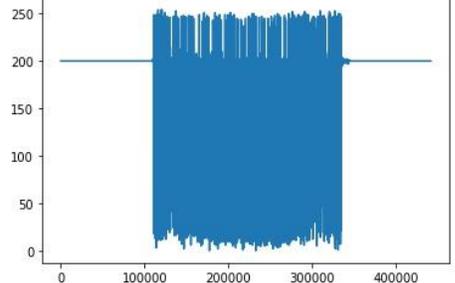
31			189.22	38.46
32			171.68	54.84
33			153.44	62.16
34			168.35	58.99
35			178.44	51.36

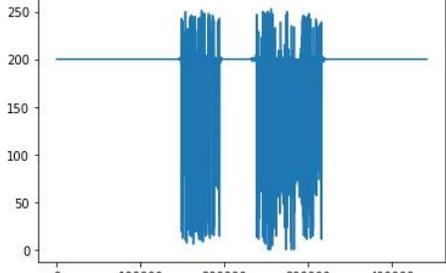
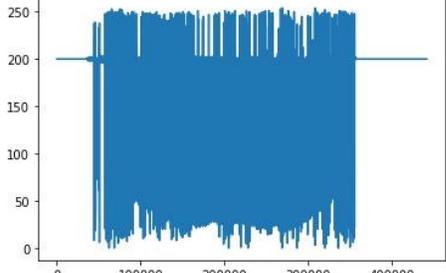
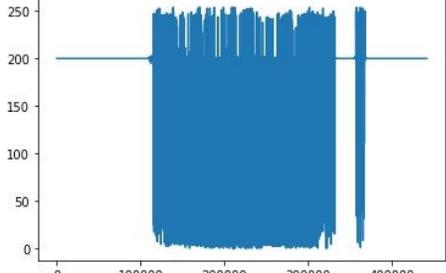
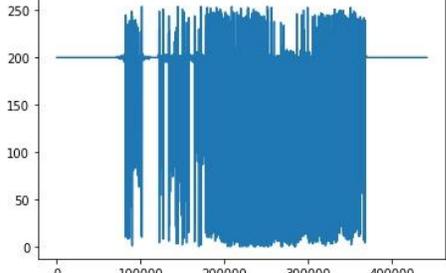
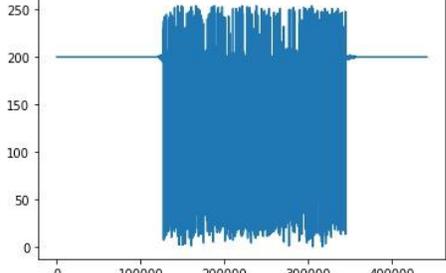
36			181.94	50.18
37			192.78	28.52
38			176.46	49.68
39			176.67	46.72
40			173.22	52.8

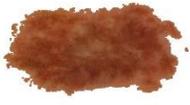
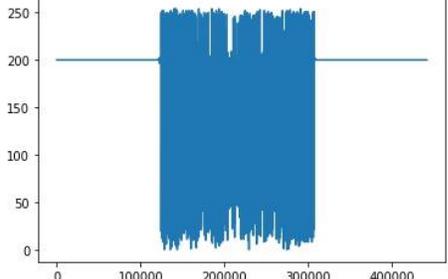
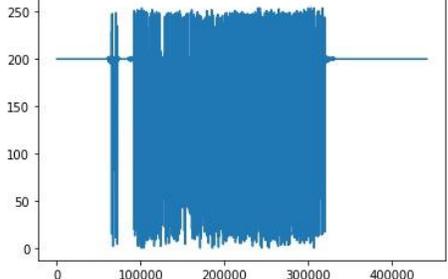
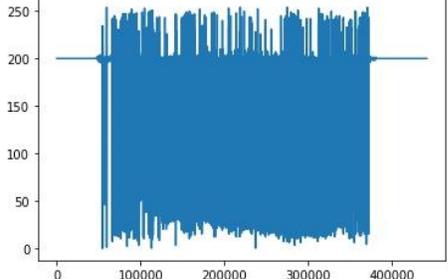
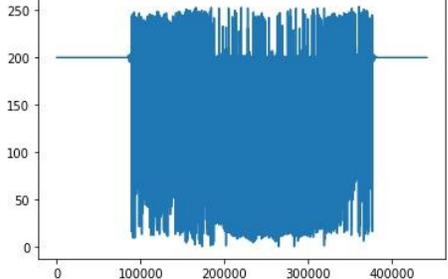
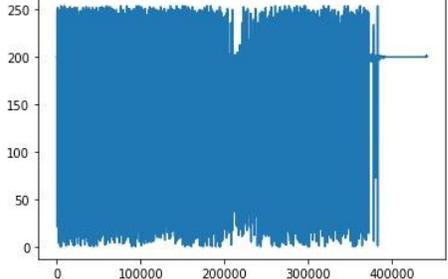
41			174.64	49.2
42			181.38	43.5
43			181.74	42.85
44			173.78	53.31
45			172.52	55.41

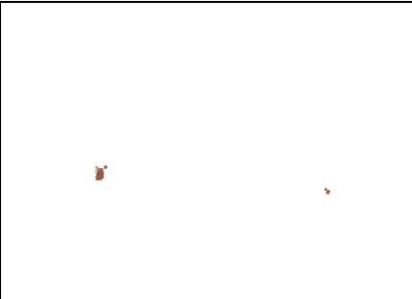
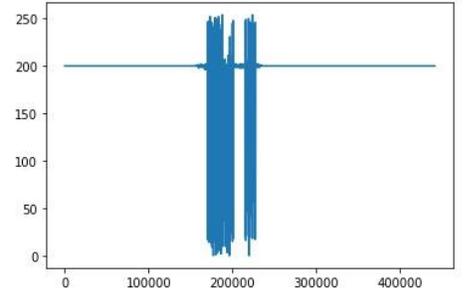
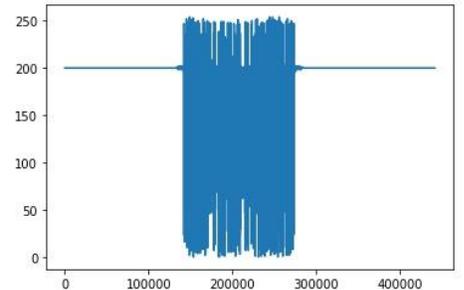
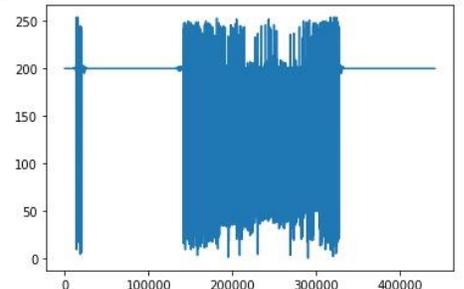
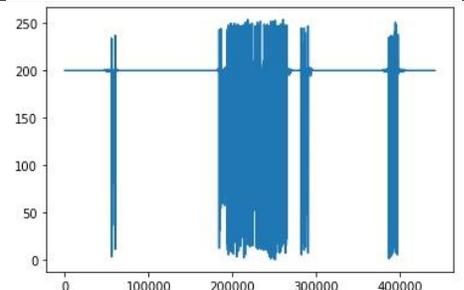
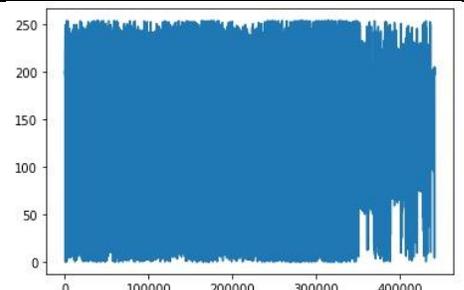
46			169.79	54.23
47			188.63	38.58
48			161.96	63.23
49			157.35	65.37
50			164.92	61.55

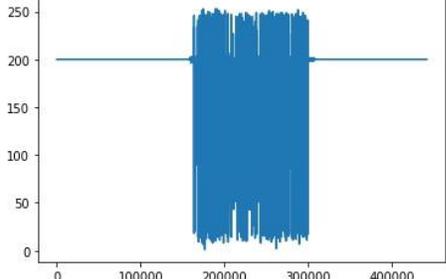
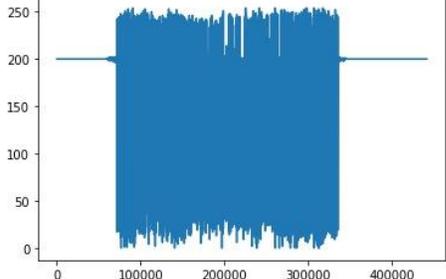
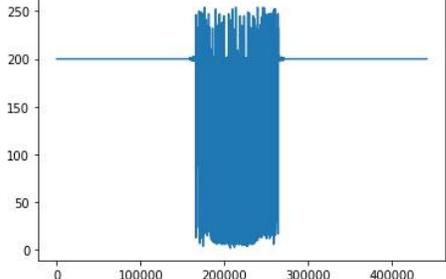
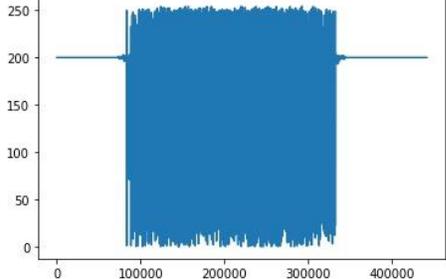
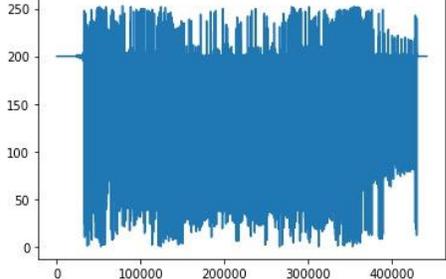
51			148.11	70.92
52			168.78	58.29
53			158.72	61.61
54			183.13	43.32
55			168.56	60.91

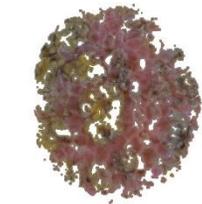
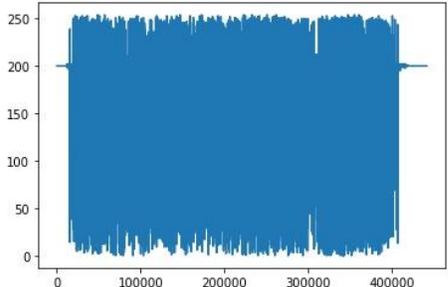
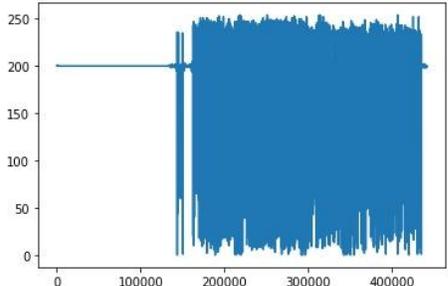
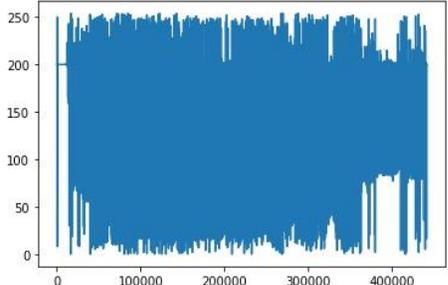
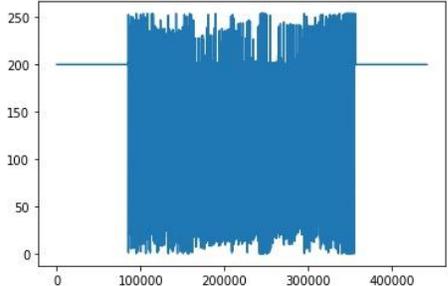
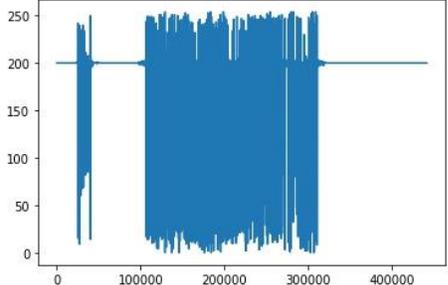
56			181.31	50.59
57			154.9	67.96
58			143.42	58.78
59			194.71	25.35
60			174.54	57.29

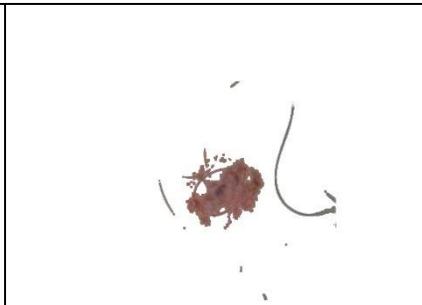
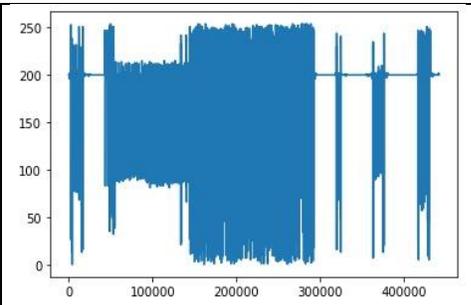
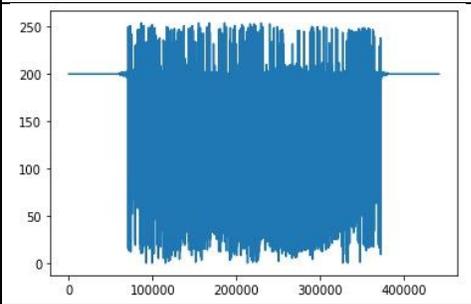
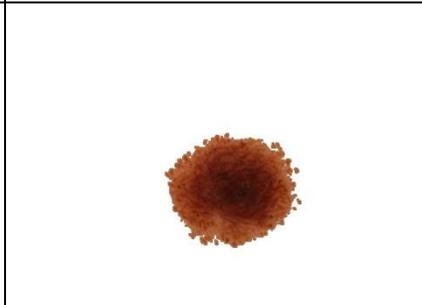
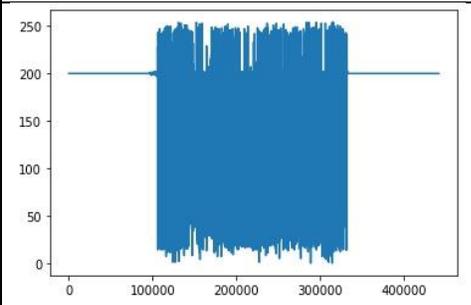
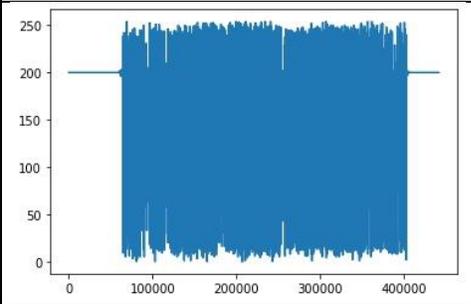
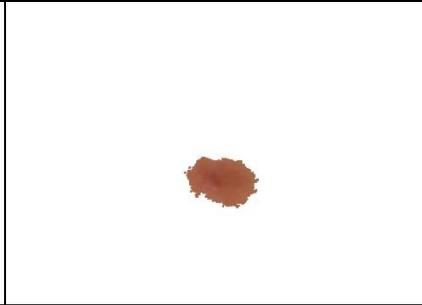
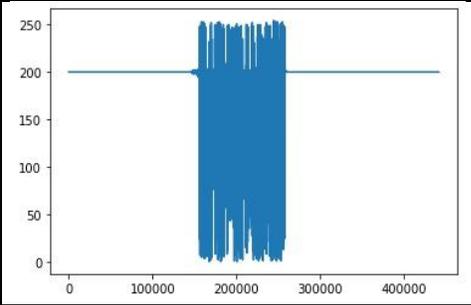
61			198.92	11.42
62			171.1	58.18
63			175.69	61.32
64			182.54	50.83
65			184.53	44

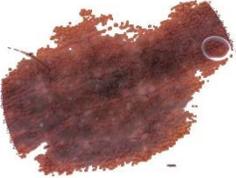
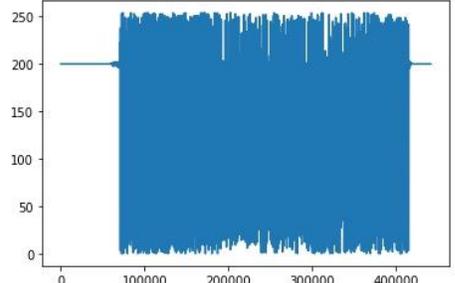
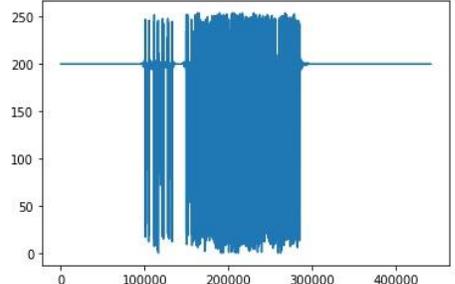
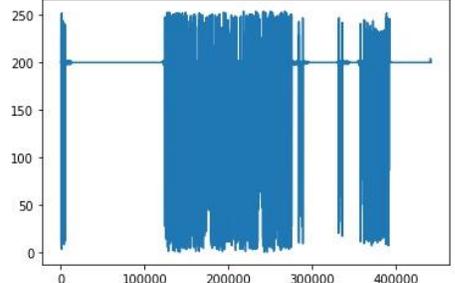
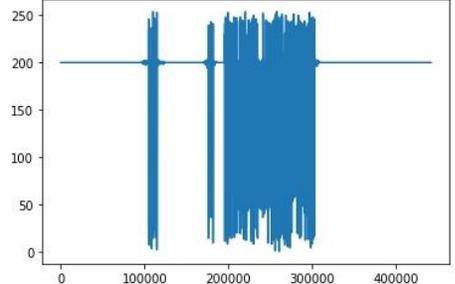
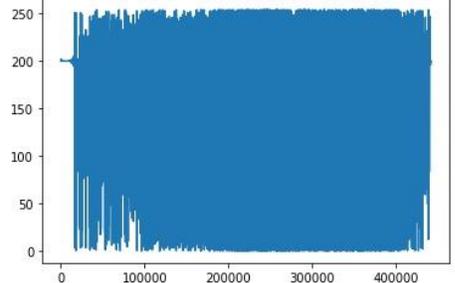
66			176.45	50.36
67			173.68	51.59
68			172.9	58.04
69			184.23	46.09
70			178.06	47.8

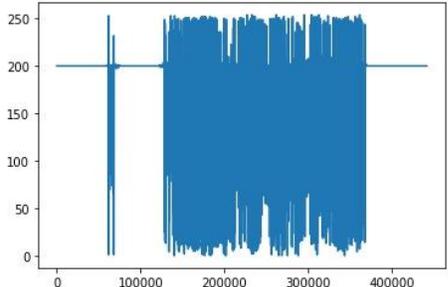
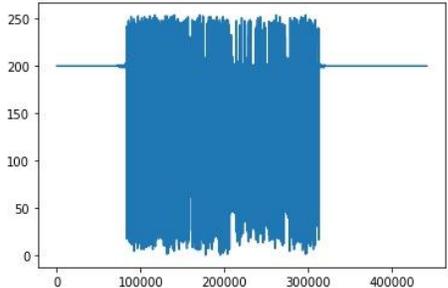
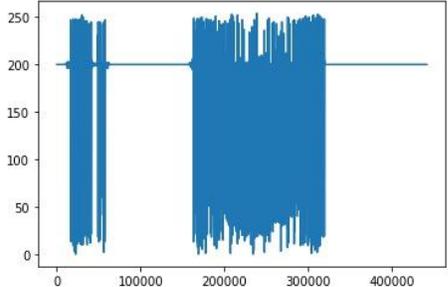
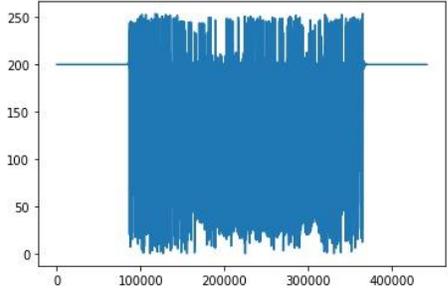
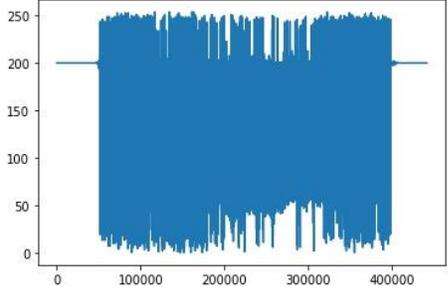
71			199.76	5.48
72			192.51	28.74
73			185.49	42.55
74			197.57	16.87
75			169.81	54.78

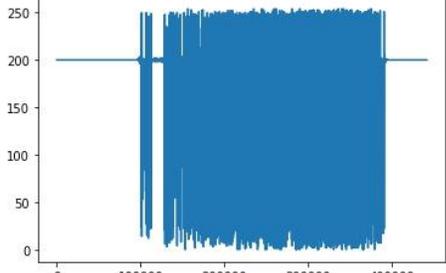
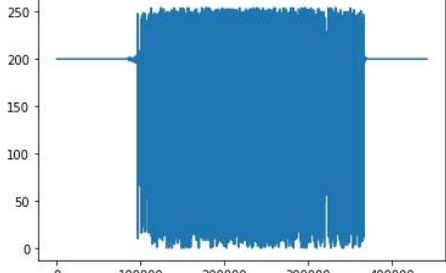
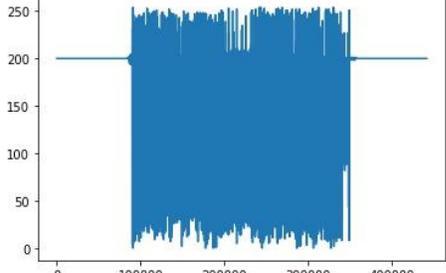
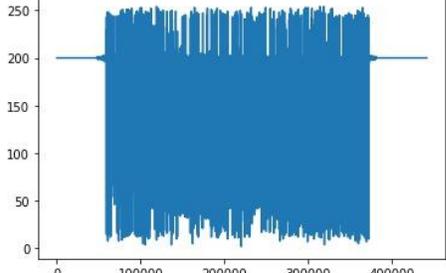
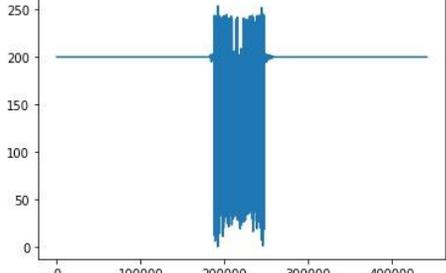
76			195.2	23.97
77			168.54	58.91
78			194.94	28.45
79			176.22	47.97
80			164.39	59.19

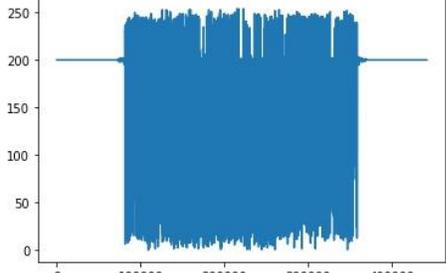
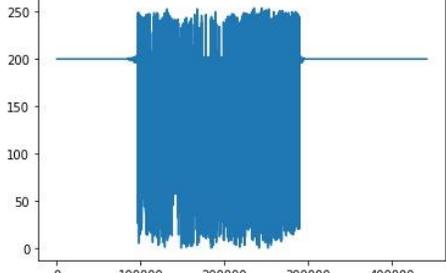
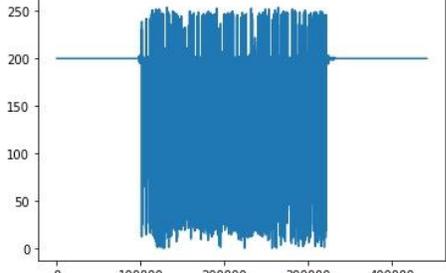
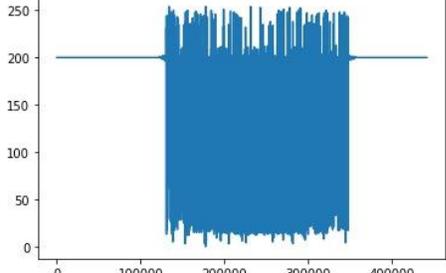
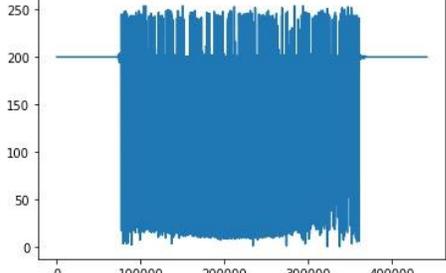
81			160.01	57.51
82			192.64	29.69
83			156.03	67
84			164.07	64.05
85			190.76	34.99

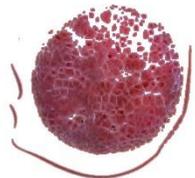
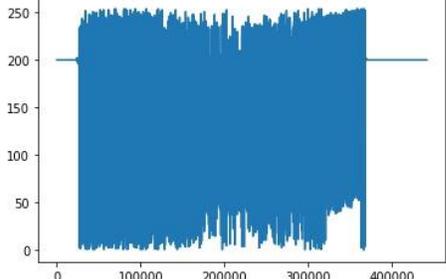
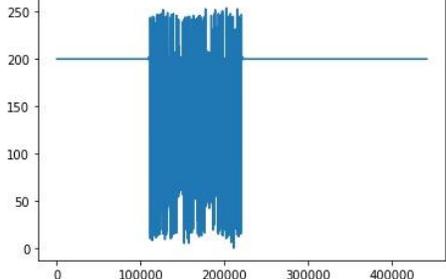
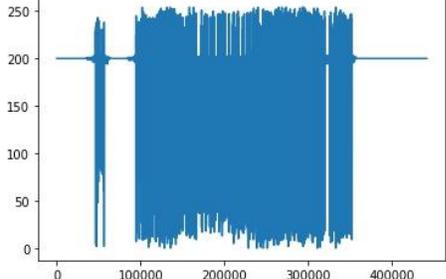
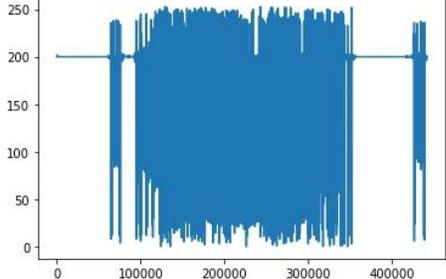
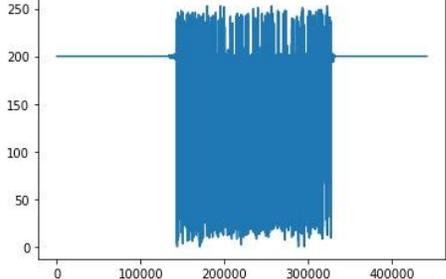
86			192.86	28.21
87			179.09	50.77
88			177.09	53.07
89			183	41.85
90			195.56	22.39

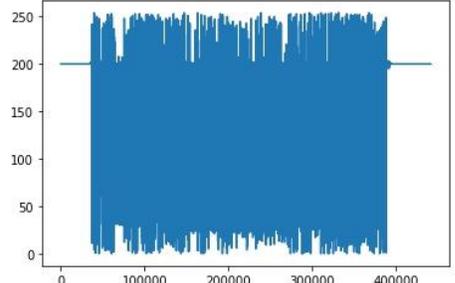
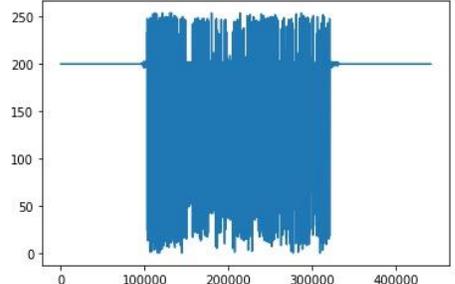
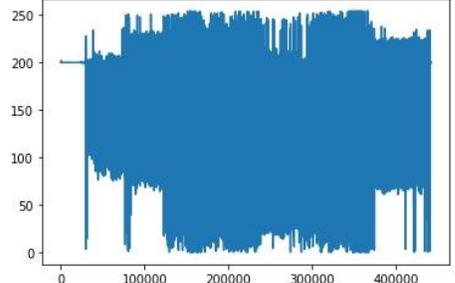
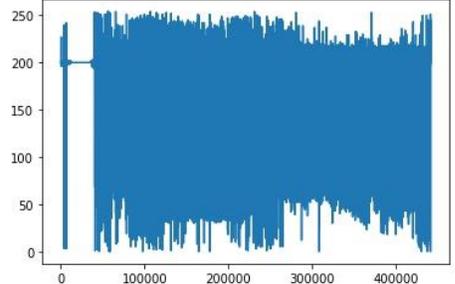
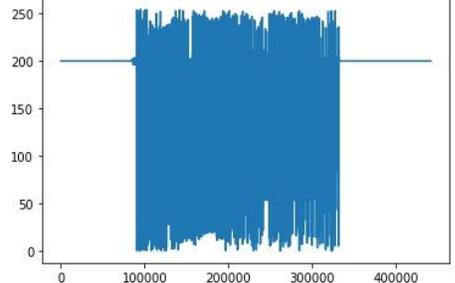
91			148.42	65.26
92			193.29	27.56
93			191.28	31.24
94			196.19	21.47
95			172.88	53.07

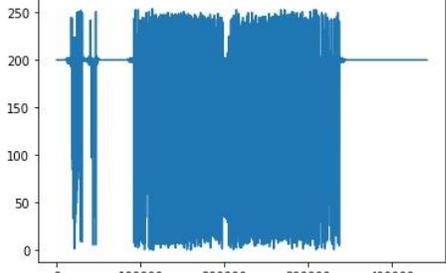
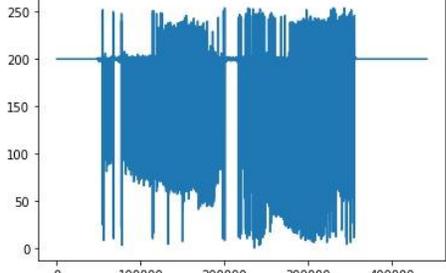
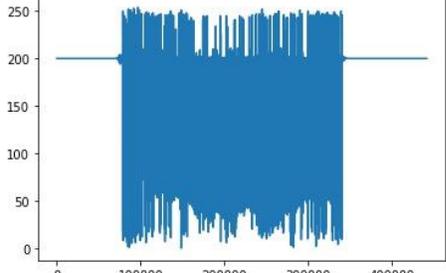
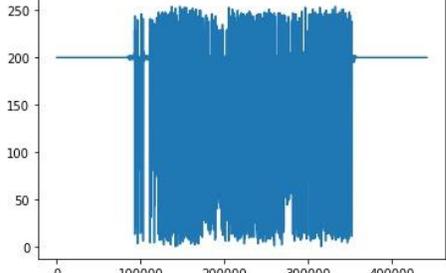
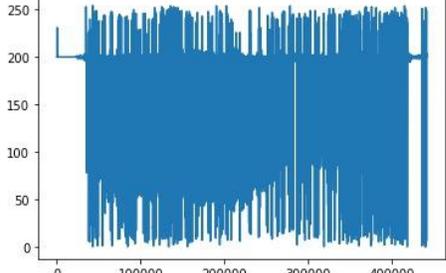
96			182.85	41.54
97			181.18	46.19
98			191.14	32.9
99			181.62	47.59
100			168.36	54.29

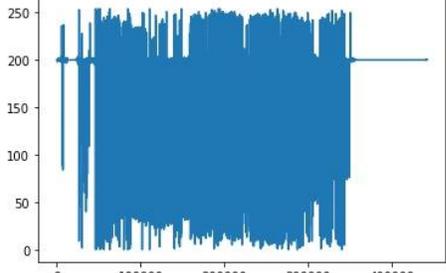
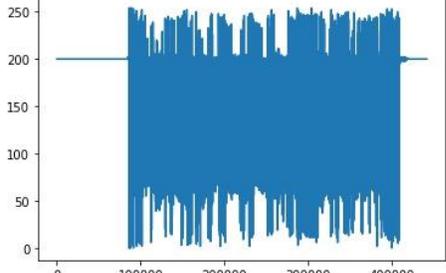
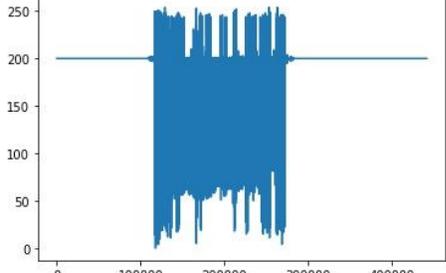
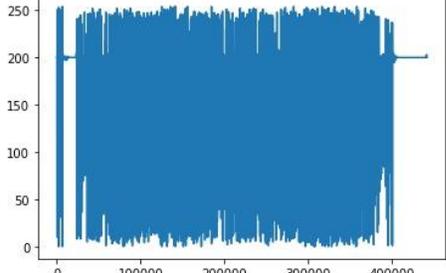
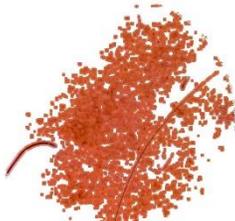
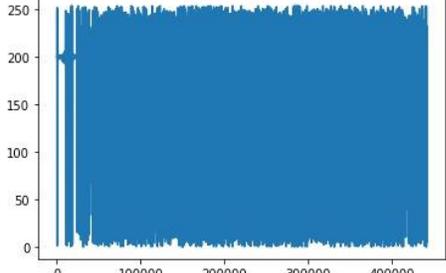
101			193.56	27.48
102			181.45	43.12
103			172.47	55.6
104			173.46	53.61
105			198.33	15.28

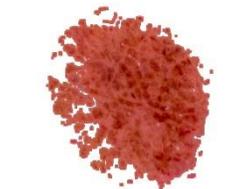
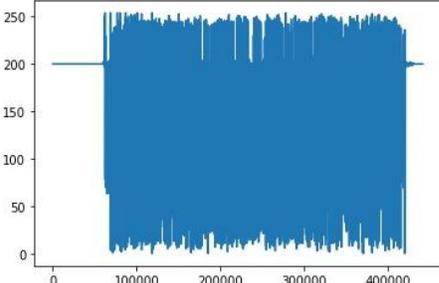
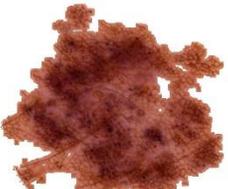
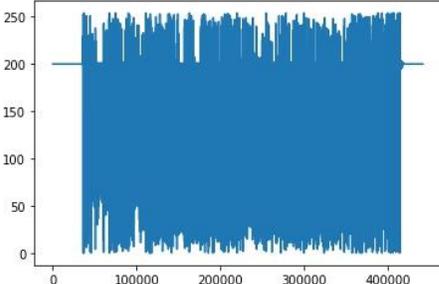
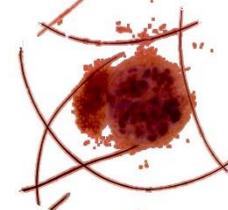
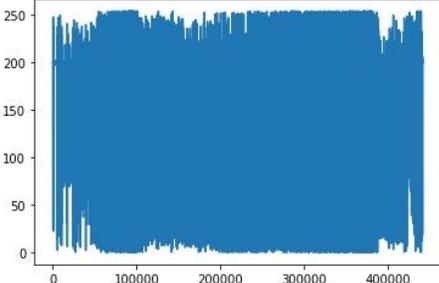
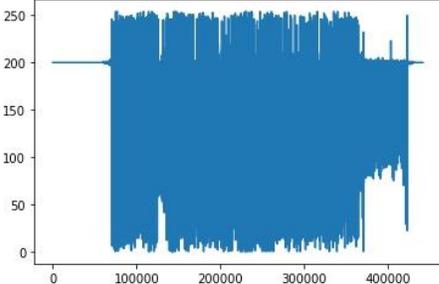
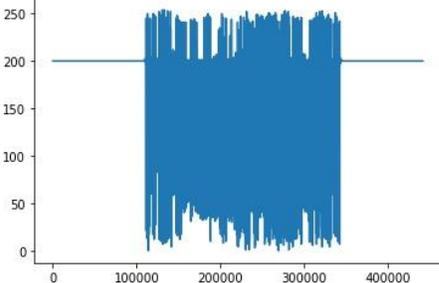
106			170.02	57.68
107			193.76	28.55
108			180.78	48.3
109			185.2	46.51
110			166.53	66.16

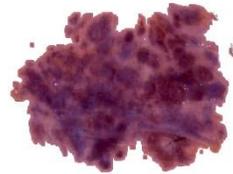
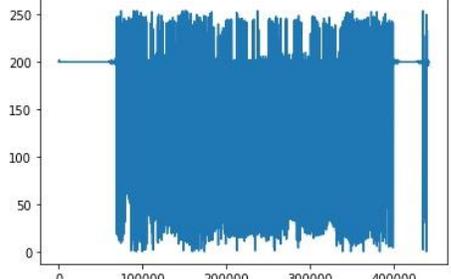
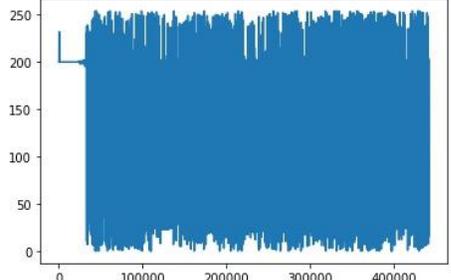
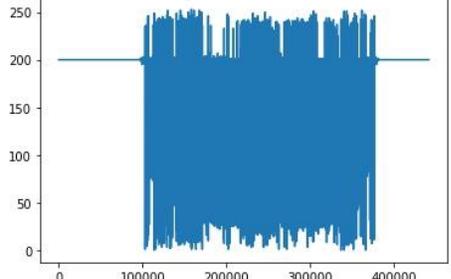
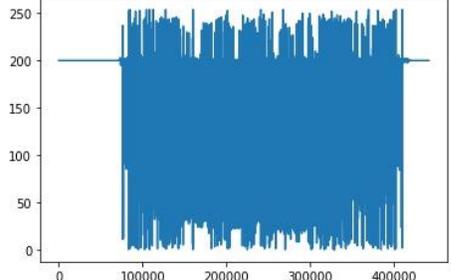
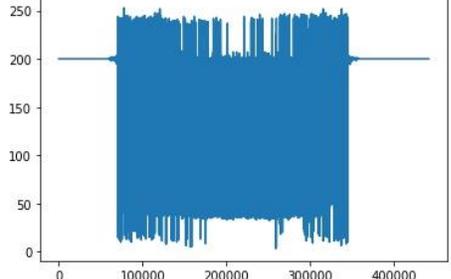
111			168.83	54.29
112			196.61	20.01
113			183.79	42.68
114			188.76	38.26
115			188.22	39.68

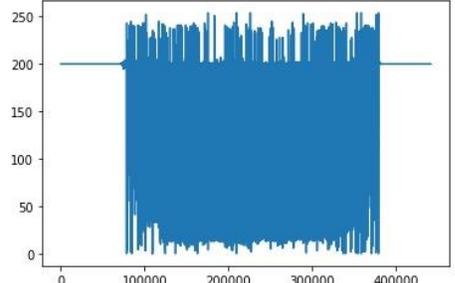
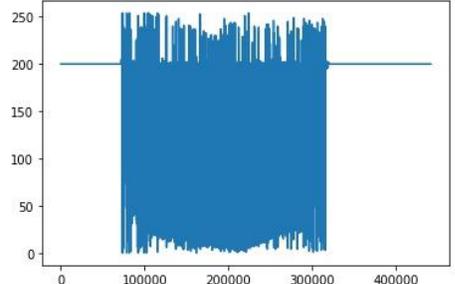
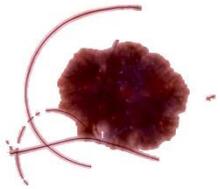
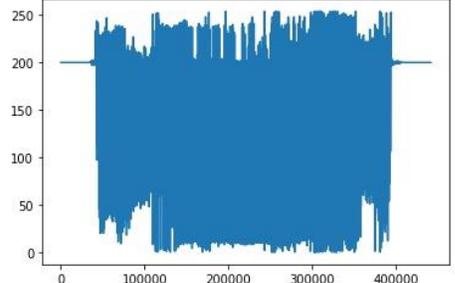
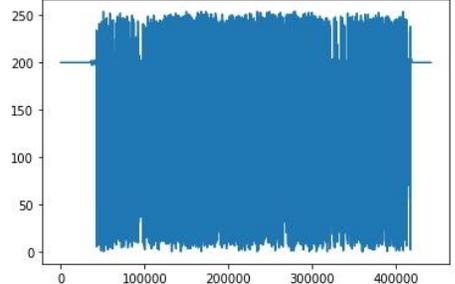
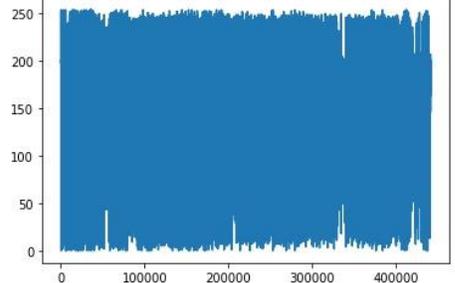
116			148.08	70.38
117			173.65	52.3
118			171.46	57.41
119			182.06	46.26
120			176.68	48.66

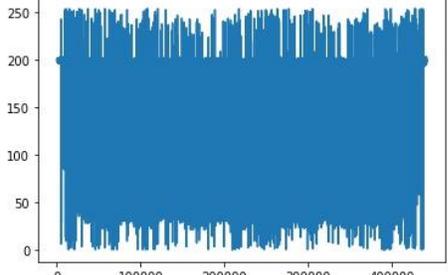
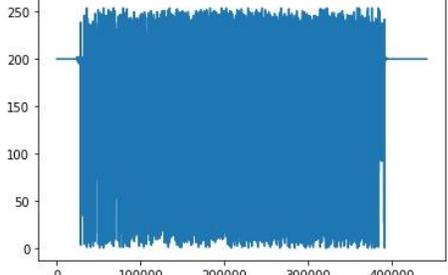
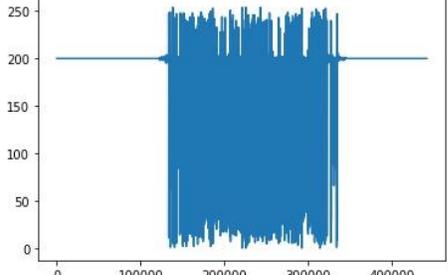
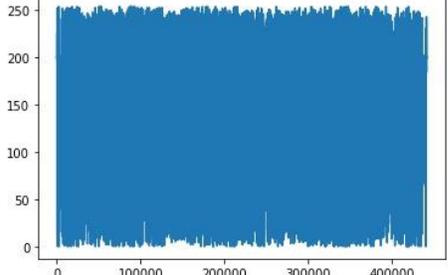
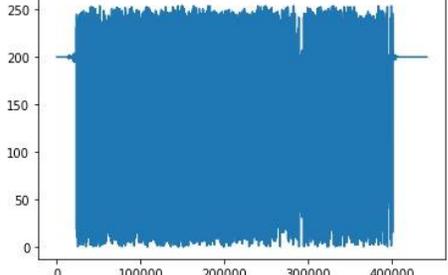
121			181.53	45.36
122			190.69	34.29
123			182.28	45.4
124			182.5	42.05
125			187.27	36.73

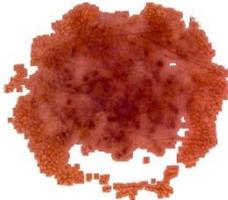
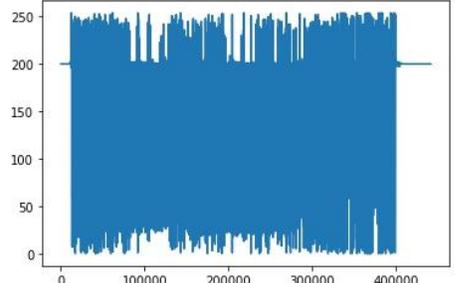
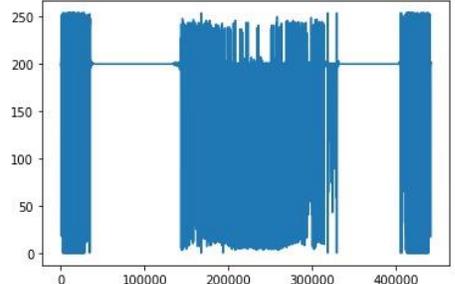
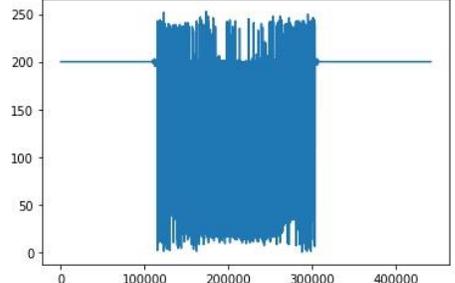
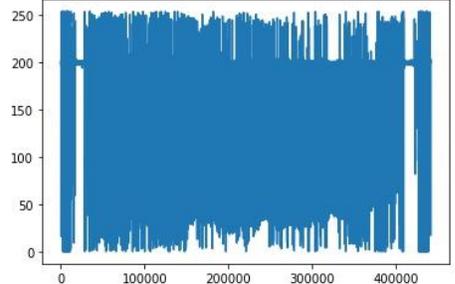
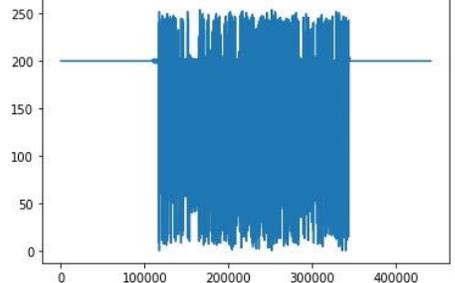
126			172.33	53.49
127			171.38	50.5
128			192.34	29.36
129			162.37	54.82
130			162.66	53.78

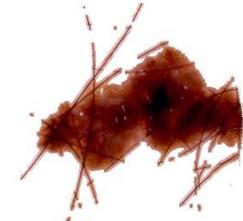
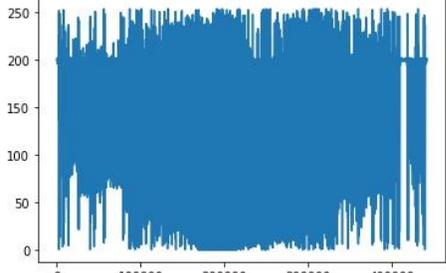
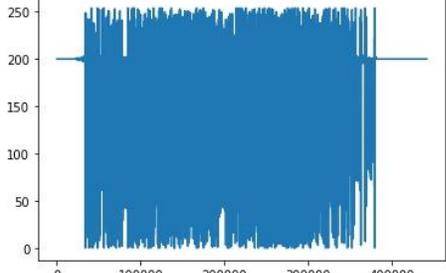
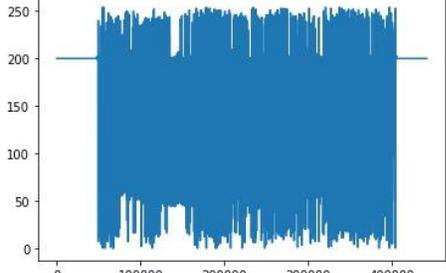
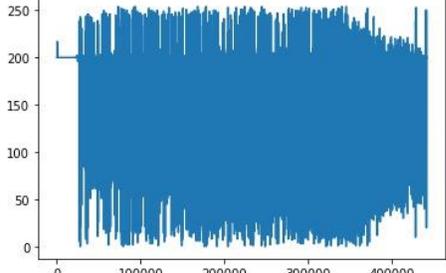
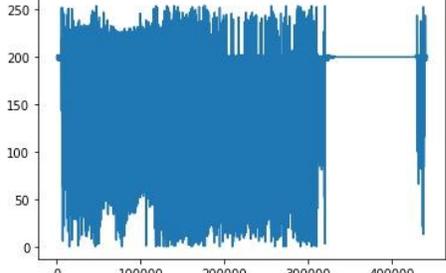
131			165.46	55.75
132			145.36	66.39
133			171.81	58.28
134			159.64	60.96
135			175.89	49.92

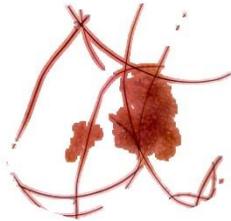
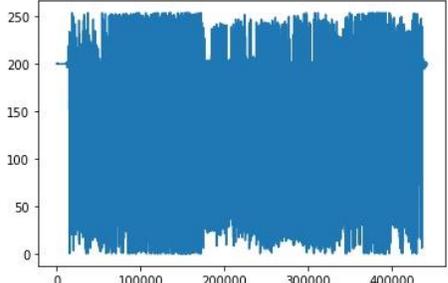
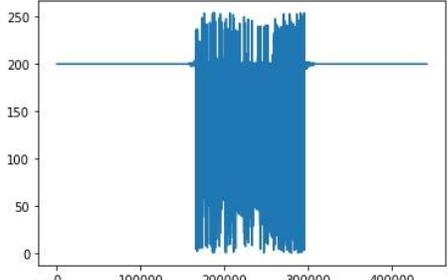
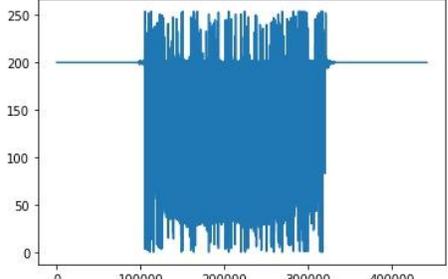
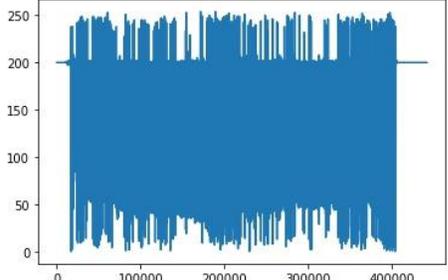
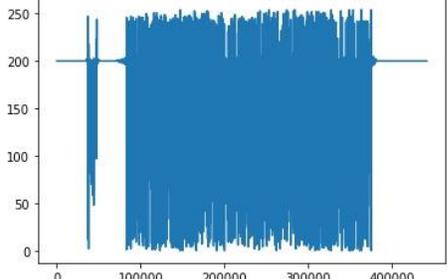
136			146	68.62
137			124.47	75.23
138			178.55	49.62
139			176.59	50.88
140			168.13	60.94

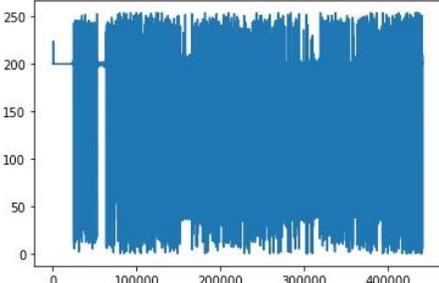
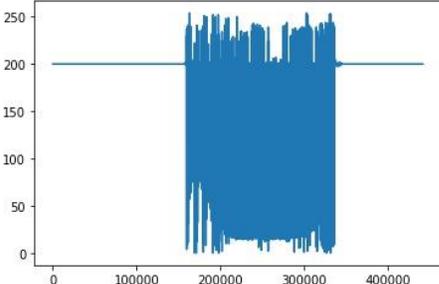
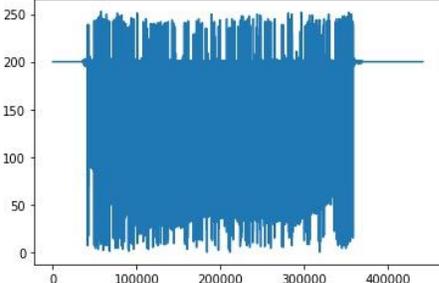
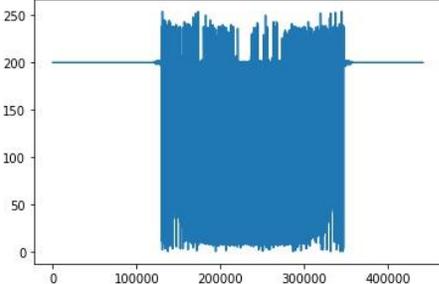
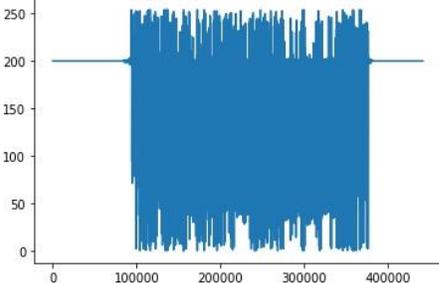
141			154.29	72.25
142			171.53	61.84
143			171.76	61.48
144			166.72	53.92
145			142.4	61.58

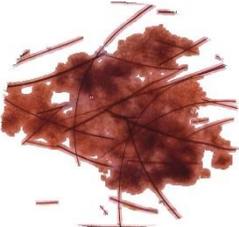
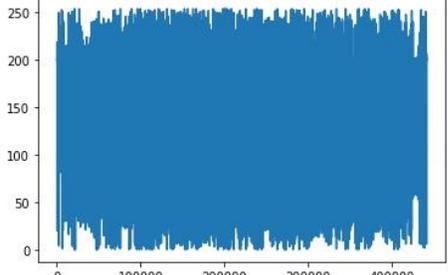
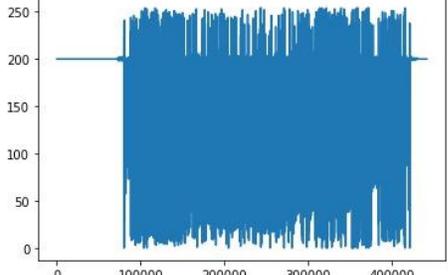
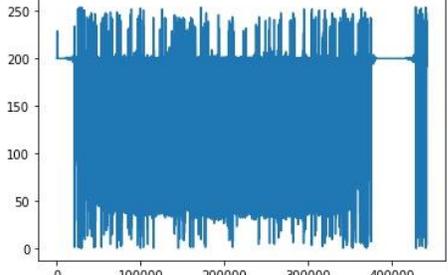
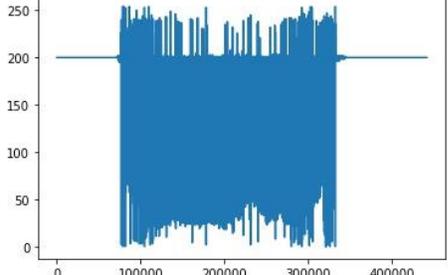
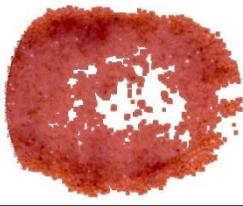
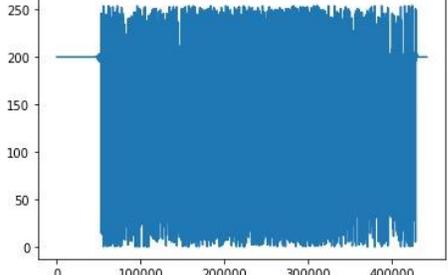
146			138.27	71.75
147			168.45	52.49
148			186.69	41.72
149			154.28	57.18
150			168.86	52.61

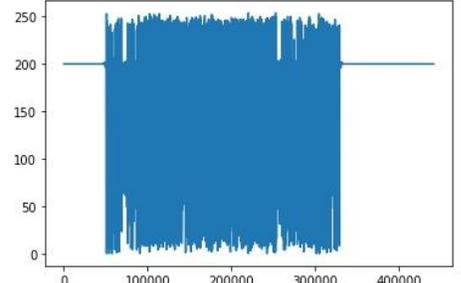
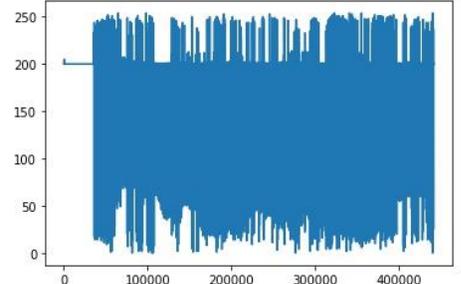
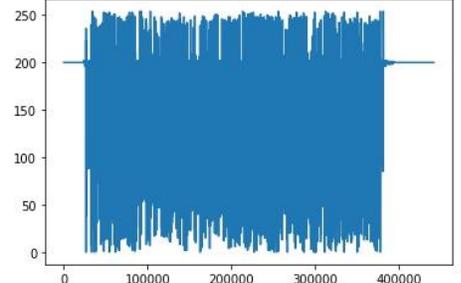
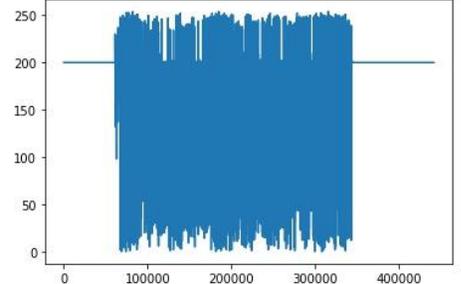
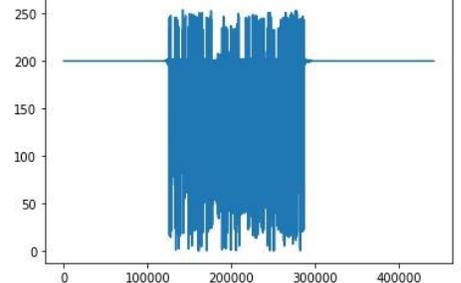
151			144.72	65
152			176.75	59.42
153			184.6	46.16
154			132.45	68.58
155			181.46	43.84

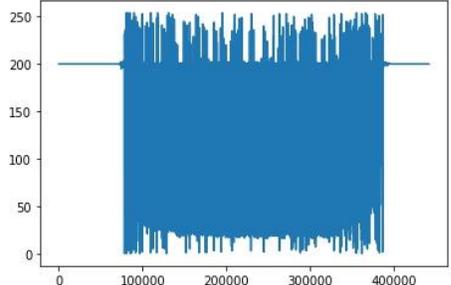
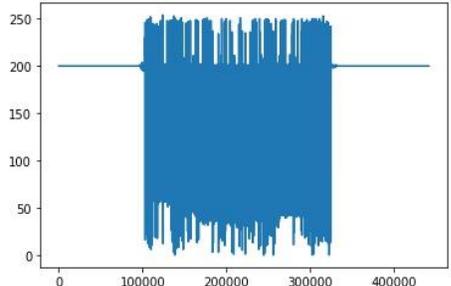
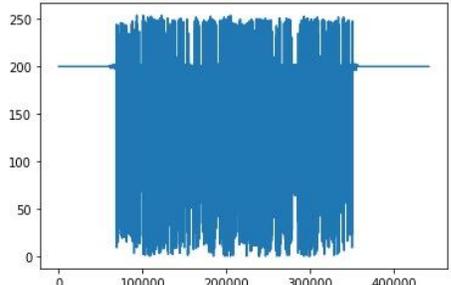
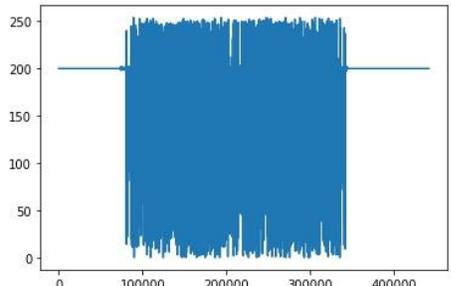
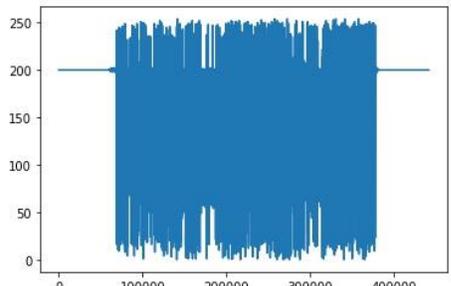
156			162.14	67.13
157			175.62	48.03
158			158.53	59.67
159			166.79	52.93
160			179.88	48.56

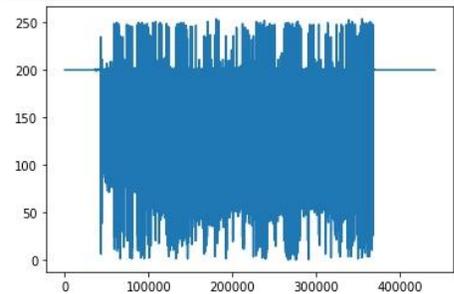
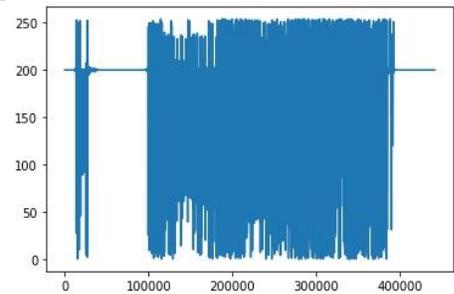
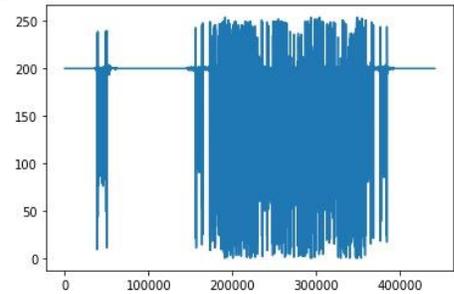
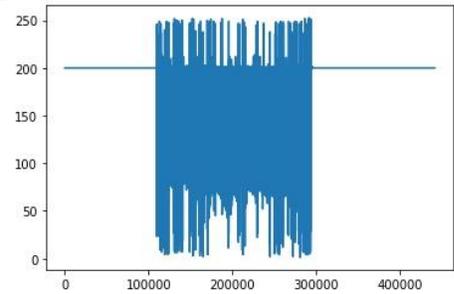
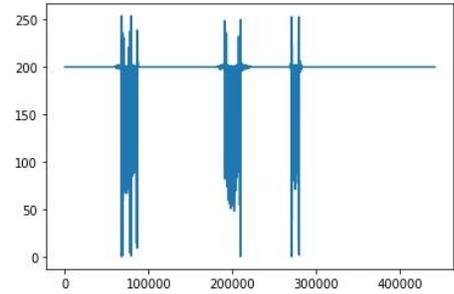
161			182.57	42.46
162			193.26	27.75
163			182.91	46.18
164			149.09	63.38
165			161.73	56.1

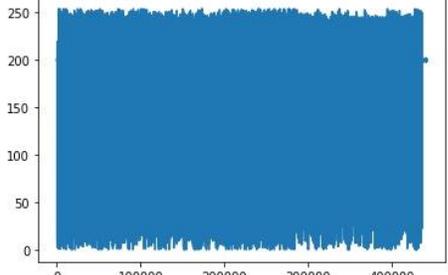
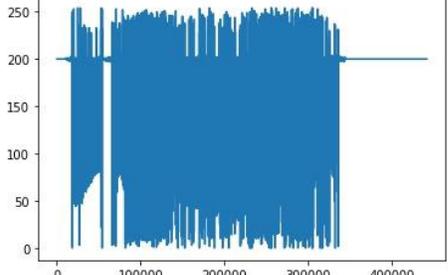
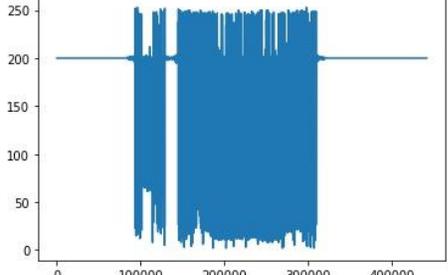
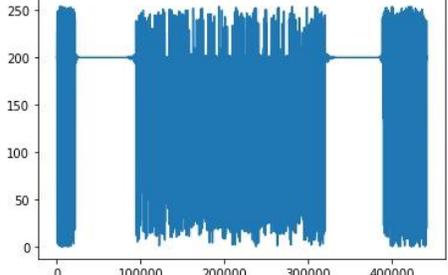
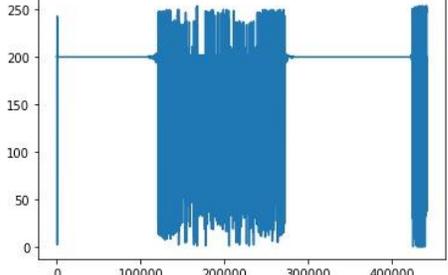
166			150.23	65.28
167			186.86	43.42
168			156.7	65.95
169			174.85	60.05
170			165.89	56.27

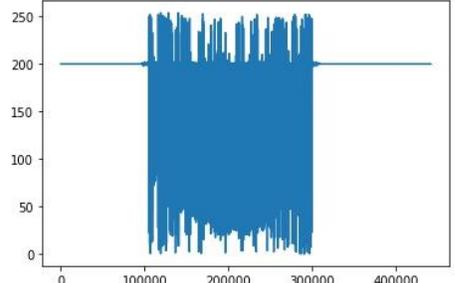
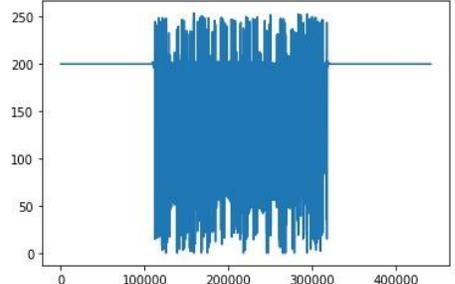
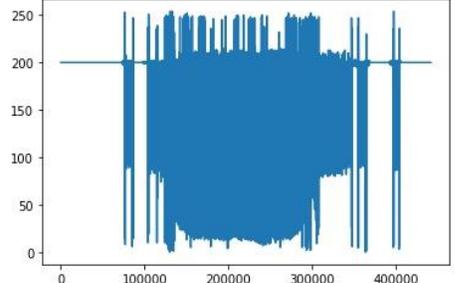
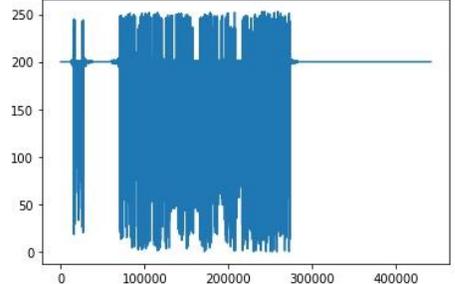
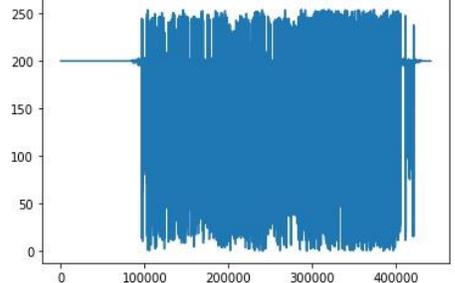
171			150.84	65.54
172			156.36	62.37
173			157.94	65.61
174			178.34	50.41
175			143.9	60.44

176			181.81	41.73
177			152.12	62.32
178			159.07	59.51
179			169.32	56.08
180			185.98	40.12

181			146.6	75.01
182			183.12	44.85
183			175.44	48.19
184			172.56	51.01
185			166.18	53.56

186			175.58	49.44
187			180.32	44.67
188			189.15	34
189			191.08	31.04
190			199.8	4.81

191			179.77	47.83
192			169.68	54.34
193			186.04	44.19
194			169.21	57.84
195			186.86	39.96

196			182.22	45.69
197			180.48	44.77
198			186.68	43.24
199			184.76	40.26
200			161.87	58.02