

HJAIEJ MOHAMED



# ECG SEMAINE 2

NEXT GENERATION IT

# EDA

L'ANALYSE EXPLORATOIRE DE DONNÉES (EDA) EST UNE ÉTAPE CRUCIALE DANS LE PROCESSUS D'ANALYSE DE DONNÉES QUI VOUS AIDE À COMPRENDRE LA STRUCTURE, LES RELATIONS ET LES MODÈLES DANS VOS DONNÉES



# EDA STEPS

**VISUALISATION DES DONNÉES:** TRACER DES ÉCHANTILLONS DES IMAGES DANS VOTRE JEU DE DONNÉES POUR AVOIR UNE COMPRÉHENSION VISUELLE DES DONNÉES. VOUS POUVEZ UTILISER DES OUTILS TELS QUE MATPLOTLIB EN PYTHON POUR CRÉER DES GRAPHIQUES ET DES HISTOGRAMMES.

**DISTRIBUTION DES DONNÉES:** TRACER LA DISTRIBUTION DES TAILLES, DES RÉSOLUTIONS OU DE TOUTE AUTRE CARACTÉRISTIQUE PERTINENTE DES IMAGES. CELA PEUT VOUS AIDER À IDENTIFIER LES TAILLES ET RÉSOLUTIONS LES PLUS COURANTES, AINSI QUE LES ANOMALIES OU LES VALEURS EXTRÊMES.

**STATISTIQUES D'IMAGE:** CALCULER DES STATISTIQUES DE BASE POUR LES IMAGES, TELLES QUE LA MOYENNE, L'ÉCART-TYPE, LA VALEUR MINIMALE ET MAXIMALE POUR CHAQUE CANAL DE L'IMAGE. CES STATISTIQUES PEUVENT FOURNIR UN RÉSUMÉ DE LA DISTRIBUTION GLOBALE DE LA LUMINOSITÉ ET DE LA COULEUR DES IMAGES DANS VOTRE JEU DE DONNÉES.

**CORRÉLATION DES DONNÉES:** VÉRIFIER LES CORRÉLATIONS ENTRE DIFFÉRENTES CARACTÉRISTIQUES OU CANAUX DES IMAGES. PAR EXEMPLE, VOUS POUVEZ VÉRIFIER LA CORRÉLATION ENTRE LA LUMINOSITÉ DE L'IMAGE ET LA TAILLE DE L'IMAGE, OU ENTRE LA RÉSOLUTION DE L'IMAGE ET LA QUALITÉ DE L'IMAGE.

**SÉGMENTATION DES DONNÉES:** ISÉGMENTER LES IMAGES DE VOTRE JEU DE DONNÉES EN DIFFÉRENTES CLASSES OU CATÉGORIES, ET EFFECTUER UNE EDA SÉPARÉE POUR CHAQUE SEGMENT. CELA PEUT VOUS AIDER À IDENTIFIER LES DIFFÉRENCES OU LES MOTIFS AU SEIN DE CHAQUE SEGMENT QUI PEUVENT ÊTRE UTILES POUR LA FORMATION DE MODÈLES D'APPRENTISSAGE AUTOMATIQUE.

**AUGMENTATION DES DONNÉES:** EFFECTUER DES TECHNIQUES D'AUGMENTATION DES DONNÉES SUR VOTRE JEU DE DONNÉES, TELLES QUE LES ROTATIONS, LES RETOURNEMENTS ET LES ZOOMS, POUR VOIR SI LA DISTRIBUTION DES DONNÉES CHANGE. CELA PEUT VOUS AIDER À DÉTERMINER LES TECHNIQUES D'AUGMENTATION DES DONNÉES LES PLUS EFFICACES POUR VOTRE JEU DE DONNÉES.

# **NETTOYAGE DE DONNEES**



# ETAPES NETTOYAGE

**SUPPRESSION DES DUPLICATAS :** SUPPRIMEZ TOUTES LES IMAGES EN DOUBLE DANS LE JEU DE DONNÉES POUR ÉVITER UNE SUR-REPRÉSENTATION DE CERTAINES IMAGES.

**ÉVALUATION DE LA QUALITÉ DE L'IMAGE :** VÉRIFIEZ LES IMAGES QUI SONT FLOUES, DE BASSE RÉSOLUTION OU DE MAUVAISE QUALITÉ DE QUELQUE AUTRE MANIÈRE, ET SUPPRIMEZ-LES DU JEU DE DONNÉES.

**ÉVALUATION DE LA PERTINENCE :** SUPPRIMEZ LES IMAGES QUI NE SONT PAS PERTINENTES POUR L'ANALYSE, TELLES QUE LES IMAGES D'OBJETS OU DE SCÈNES QUI NE SONT PAS D'INTÉRÊT.

**CORRECTION DE FORMAT D'IMAGE :** ASSUREZ-VOUS QUE TOUTES LES IMAGES DANS LE JEU DE DONNÉES SONT AU MÊME FORMAT ET DE LA MÊME TAILLE, ET CONVERTISSEZ LES IMAGES QUI SONT DANS UN FORMAT DIFFÉRENT AU FORMAT STANDARD.

**ÉTIQUETAGE DE L'IMAGE :** ASSUREZ-VOUS QUE TOUTES LES IMAGES SONT ÉTIQUETÉES CORRECTEMENT ET DE MANIÈRE COHÉRENTE, ET CORRIGEZ LES ÉTIQUETTES QUI SONT INCORRECTES OU INCOHÉRENTES.

**AUGMENTATION DES DONNÉES :** AUGMENTEZ LES DONNÉES EN APPLIQUANT DES TECHNIQUES TELLES QUE LA ROTATION, LE REDIMENSIONNEMENT ET L'INVERSION POUR AUGMENTER LA TAILLE DU JEU DE DONNÉES ET AMÉLIORER LA ROBUSTESSE DES MODÈLES QUI SERONT FORMÉS À PARTIR DE CELUI-CI.

**PARTITIONNEMENT DES DONNÉES :** PARTITIONNEZ LES DONNÉES EN ENSEMBLES D'ENTRAÎNEMENT, DE VALIDATION ET DE TEST, EN VEILLANT À CE QUE CHAQUE ENSEMBLE REPRÉSENTE L'ENSEMBLE DU JEU DE DONNÉES ET CONTIENNE UNE DISTRIBUTION SIMILAIRE DE CLASSES.

# QUALITE DE L'IMAGE

**VISUAL INSPECTION:** TAKE A LOOK AT THE IMAGE TO SEE IF IT APPEARS VISUALLY CLEAR AND IN FOCUS. THIS CAN GIVE YOU A QUICK IDEA OF WHETHER THE IMAGE MEETS YOUR QUALITY STANDARDS.

**IMAGE RESOLUTION:** CHECK THE RESOLUTION OF THE IMAGE, WHICH IS USUALLY EXPRESSED IN PIXELS. A HIGHER RESOLUTION CAN RESULT IN A CLEAKER IMAGE, BUT IT MAY ALSO INCREASE THE SIZE OF THE IMAGE AND THE AMOUNT OF MEMORY REQUIRED TO STORE IT.

**IMAGE BRIGHTNESS:** ENSURE THAT THE IMAGE IS NOT TOO DARK OR TOO BRIGHT. IDEALLY, THE IMAGE SHOULD HAVE A WELL-BALANCED BRIGHTNESS THAT ALLOWS ALL DETAILS TO BE CLEARLY VISIBLE.

**IMAGE SHARPNESS:** LOOK FOR SIGNS OF IMAGE BLURRING, SUCH AS A LACK OF DEFINITION OR SOFT EDGES. THIS CAN BE CAUSED BY FACTORS SUCH AS CAMERA SHAKE, MOTION BLUR, OR A LOW-QUALITY LENS.

**IMAGE NOISE:** CHECK FOR THE PRESENCE OF IMAGE NOISE, WHICH CAN APPEAR AS RANDOM SPECKLES OR GRAINS IN THE IMAGE. THIS CAN BE A RESULT OF A HIGH ISO SETTING, LONG EXPOSURE TIMES, OR LOW LIGHT CONDITIONS.

**IMAGE ARTIFACTS:** CHECK FOR THE PRESENCE OF IMAGE ARTIFACTS, SUCH AS COMPRESSION ARTIFACTS OR ALIASING. THESE CAN BE CAUSED BY IMAGE PROCESSING, RESIZING, OR OTHER OPERATIONS THAT CAN AFFECT IMAGE QUALITY.

**IMAGE CONTRAST:** ENSURE THAT THE IMAGE HAS GOOD CONTRAST, MEANING THAT THERE IS A CLEAR DIFFERENCE BETWEEN THE LIGHTEST AND DARKEST AREAS OF THE IMAGE. POOR CONTRAST CAN MAKE IT DIFFICULT TO SEE DETAILS IN THE IMAGE.



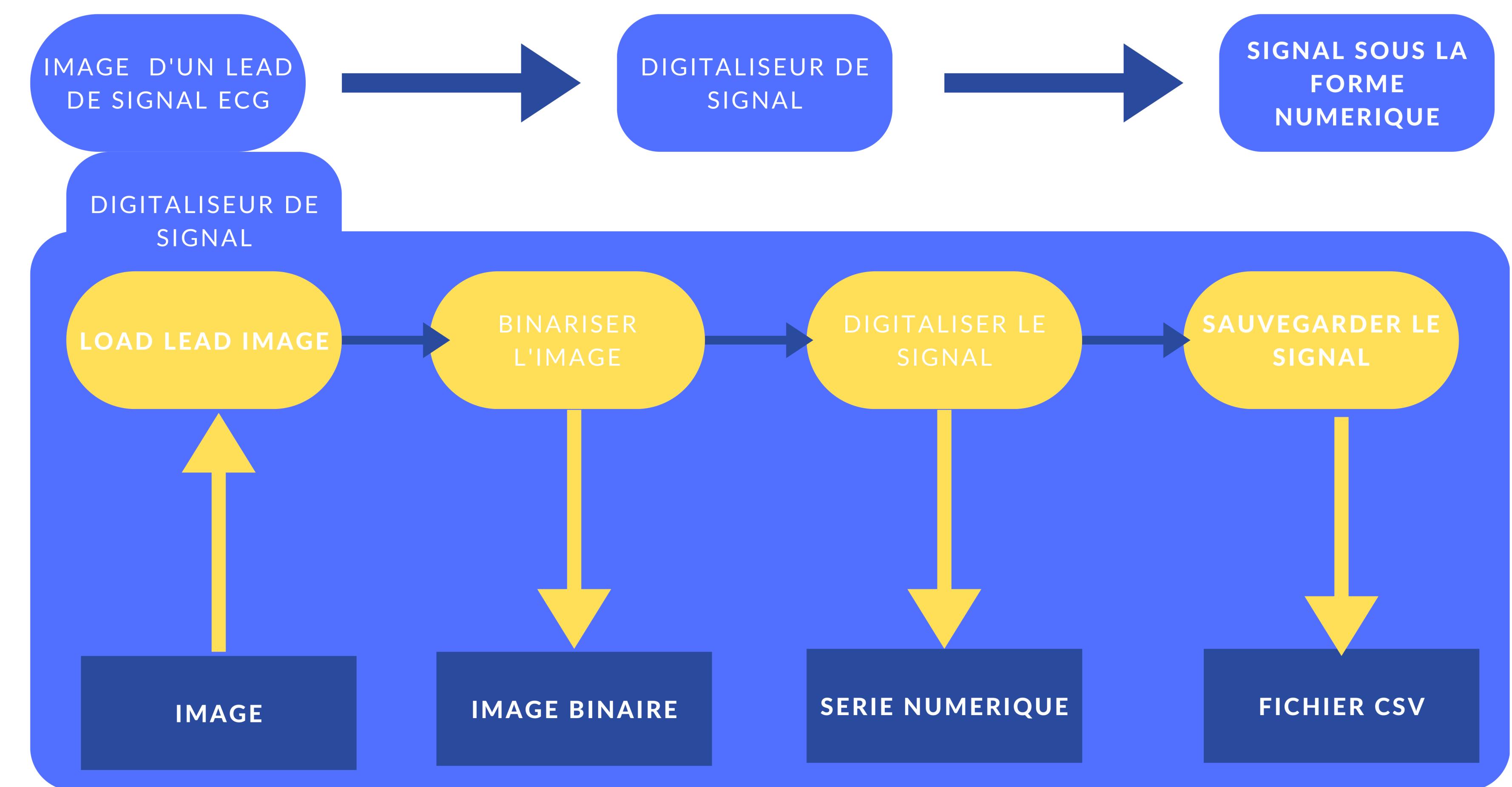
# **DIGITALISATION**

## **D'UN SIGNAL ECG**

### **A PARTIR D'UN**

# **LEAD**

# PRINCIPE



# METHODES

**MÉTHODE DE SEUILLAGE :** CETTE MÉTHODE CONSISTE À CONVERTIR L'IMAGE EN NOIR ET BLANC EN UTILISANT UN SEUIL. TOUT PIXEL AYANT UNE VALEUR SUPÉRIEURE AU SEUIL EST CONVERTI EN BLANC, TANDIS QUE TOUT PIXEL AYANT UNE VALEUR INFÉRIEURE AU SEUIL EST CONVERTI EN NOIR. CETTE MÉTHODE EST SOUVENT UTILISÉE POUR LA RECONNAISSANCE DE CARACTÈRES ET LA DÉTECTION DE CONTOURS.

**MÉTHODE DE DÉTECTION DE CONTOURS :** CETTE MÉTHODE CONSISTE À DÉTECTOR LES CONTOURS DES OBJETS DANS L'IMAGE EN UTILISANT DES FILTRES. LES CONTOURS SONT ENSUITE UTILISÉS POUR EXTRAIRE DES CARACTÉRISTIQUES DE L'IMAGE. CETTE MÉTHODE EST SOUVENT UTILISÉE POUR LA RECONNAISSANCE DE FORMES ET L'ANALYSE D'IMAGES.

**MÉTHODE DE TRANSFORMATION DE FOURIER :** CETTE MÉTHODE CONSISTE À APPLIQUER UNE TRANSFORMATION DE FOURIER À L'IMAGE POUR CONVERTIR L'INFORMATION SPATIALE EN INFORMATION FRÉQUENTIELLE. LES COEFFICIENTS DE FOURIER SONT ENSUITE UTILISÉS POUR EXTRAIRE DES CARACTÉRISTIQUES DE L'IMAGE. CETTE MÉTHODE EST SOUVENT UTILISÉE POUR LA RECONNAISSANCE DE FORMES ET LA VISION PAR ORDINATEUR.

# METHODE DE SEUILAGE (THRESHOLD)



- OPEN CV

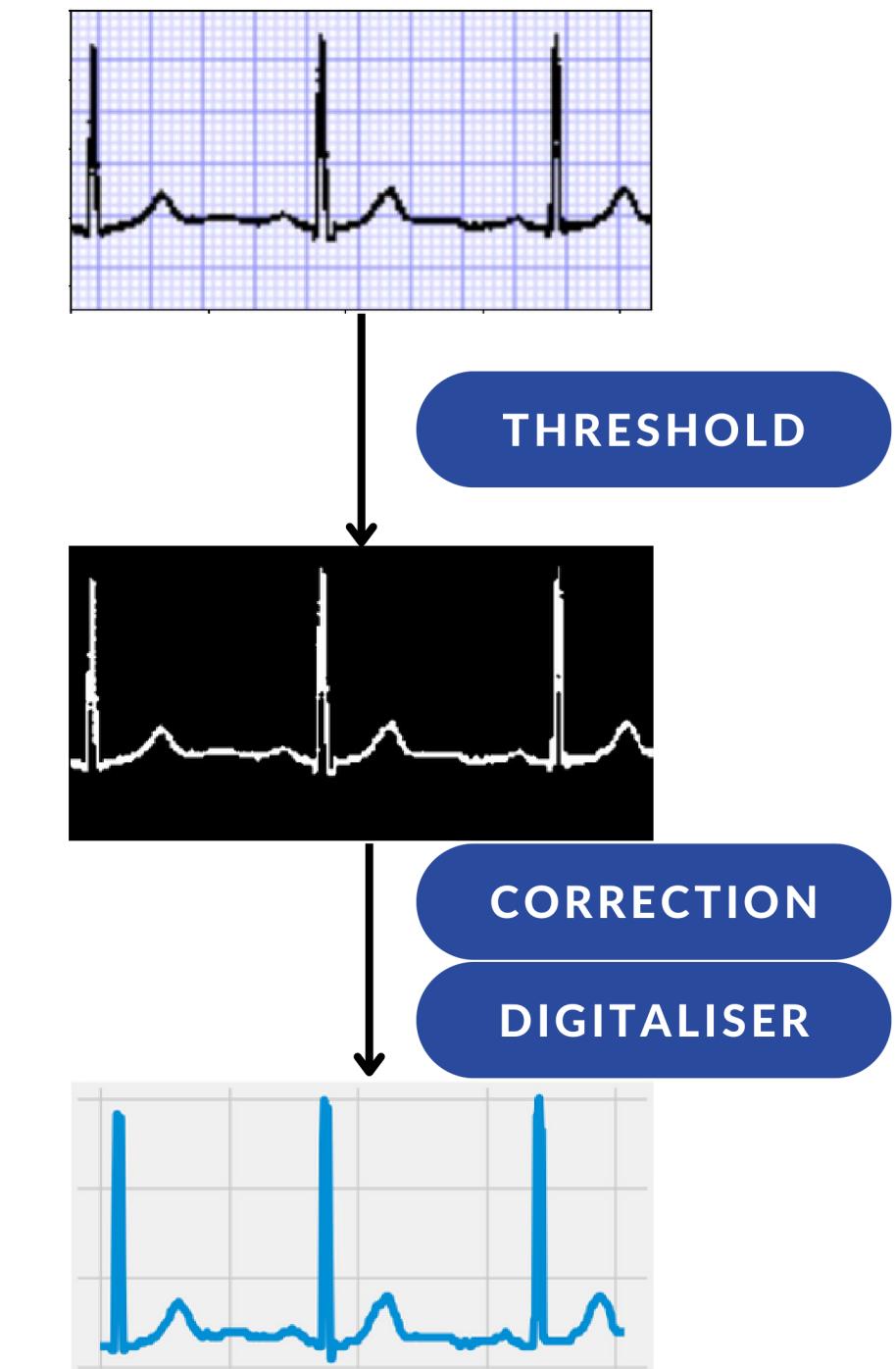
- MATPLOTLIB

- PANDAS

# SEUILLAGE +F.C

LA METHODE S'AGIT D'APPLIQUER LE SEUILLAGE(THRESHOLD) POUR BINARISER L'IMAGE(SIGNAL BLANC (255) ET LE RESTE EN NOIR (0))

LORS DE LA DIGITALISATION ON VA CONSIDRER QUE LES POINTS QUI SONT TRES LOIN DE L'AUTRES SONT DES BRUITS ET DONC EN VA UTILISER UN FACTEUR DE CORRECTION (QUI EST LA DISTANCE ENTRE DEUX POINTS CONSECUTIFS) POUR ELIMINER LES POINTS QUI INFLUENCENT SUR LA CONTINUITE DE SIGNAL (AUTREMENT DIT LOIN DU SIGNAL)



COMMAND : PYTHON LEAD\_DIGITILIZER.PY IMAGE.JPG THRESHOLD CORRECTION\_FACTOR

# TEST & LIMITES 1

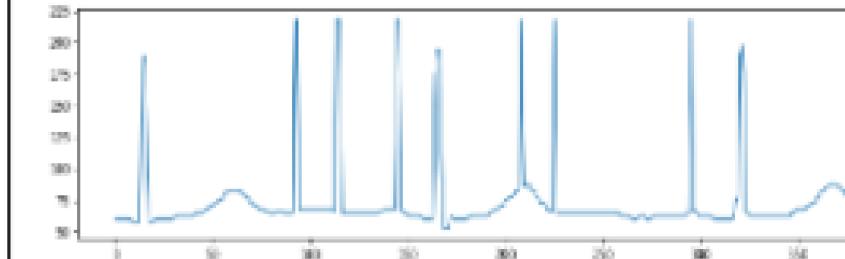
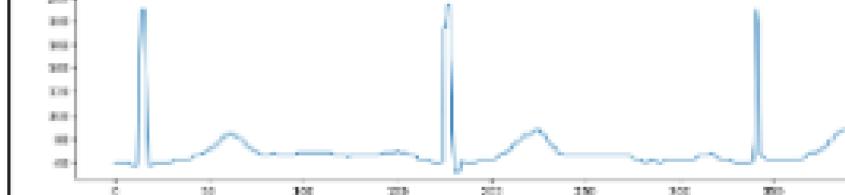
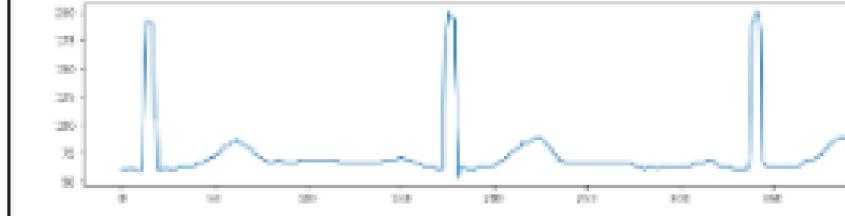
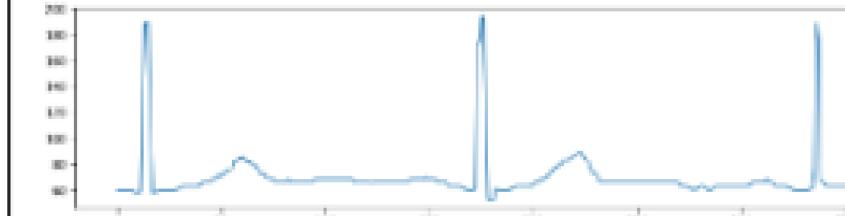
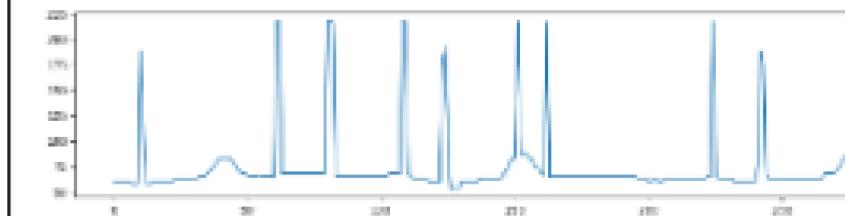
Thresh  
old

facteur de  
correction

resultats

commentaires

image réelle

10	10		pics supplémentaires.
20	10		identique avec image entrée
120	10		la variation de paramètre thresh a influence sur signal obtenue mais reste identique à l'image entrée.
20	5		la diminution de facteur de correction n'a aucune influence sur l'image => on a une forte continuité.
10	2		Le facteur de correction n'a aucune influence sur le pics.



LE THRESHOLD EST  
SUFFISANT POUR  
LA DIGITALISATION

# TEST & LIMITES 2

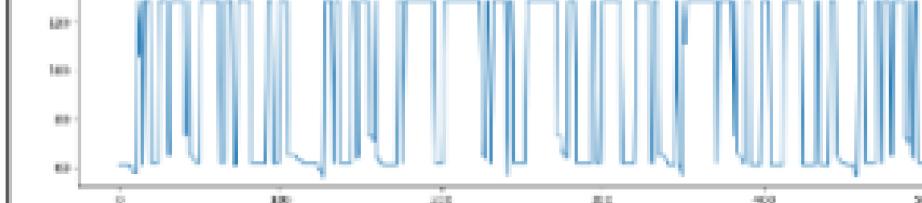
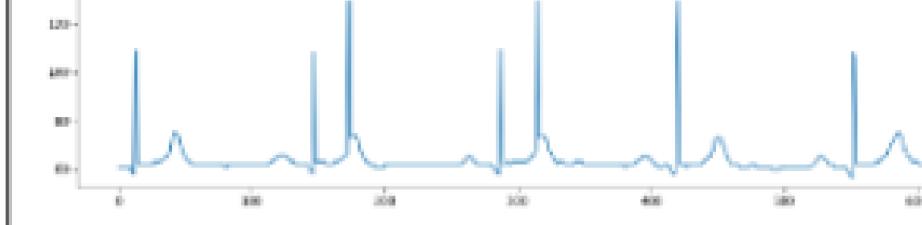
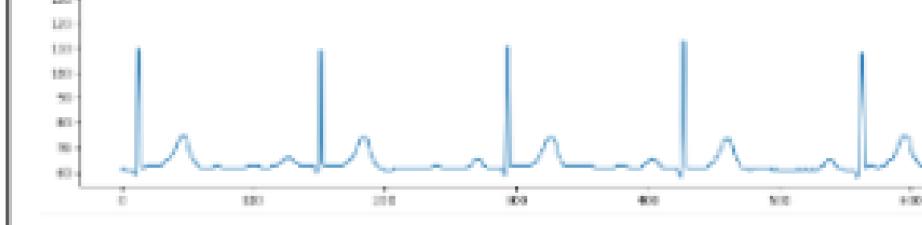
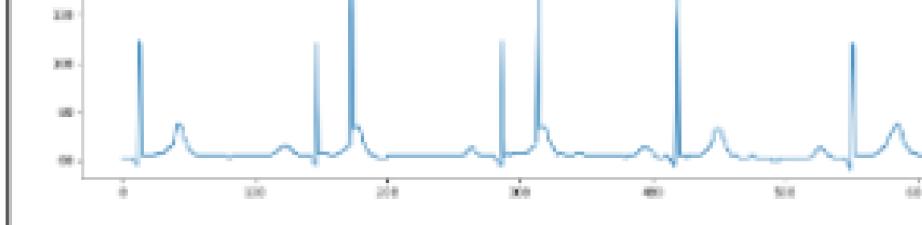
Threshold

facteur de correction

resultats

commentaires

image réelle

Threshold	facteur de correction	resultats	commentaires	image réelle
20	10		rien ,pas de variation	
80	10		variation non identique	
120	10		variation presque identique mais on des pics supplémentaires	
150	10		variation identique au signal d'entrée	
120	5		facteur de correction n'a pas une grande influence sur l'élimination des pics supplémentaires	

LE THRESHOLD EST  
SUFFISANT POUR  
LA DIGITALISATION

# TEST & LIMITES 3

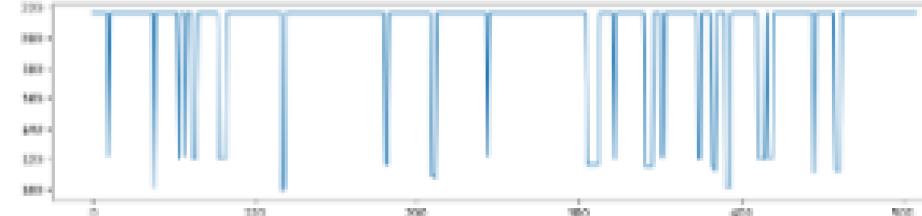
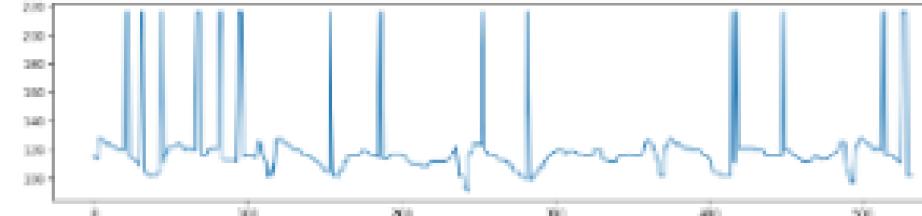
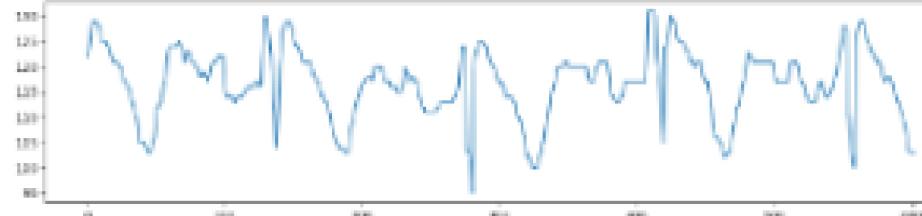
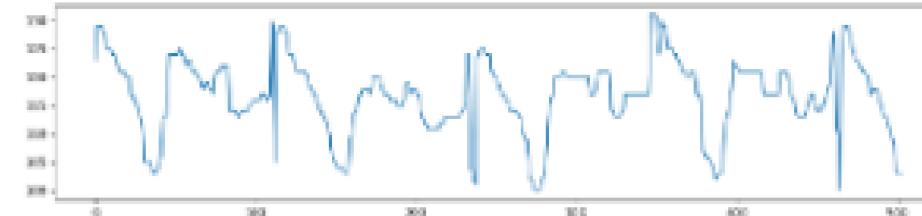
Threshold

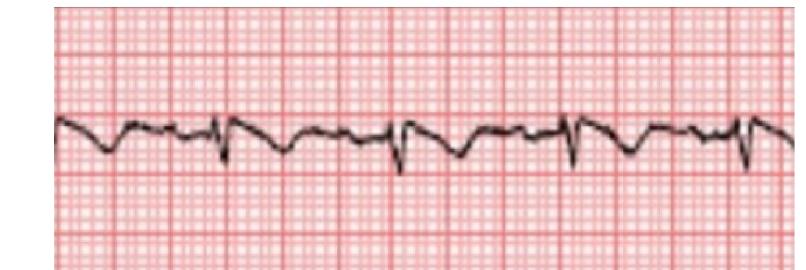
facteur de correction

resultats

commentaires

image réelle

20	10		non identique
80	10		on a un signal presque identique, mais on a des pics supplémentaires (bruits)
120	10		on a une variation de signal identique
120	5		on a perdu quelques points mais on a toujours la même variation.



LE THRESHOLD EST SUFFISANT POUR LA DIGITALISATION

L'UTILISATION DE FACTEUR DE CORRECTION (CONSERVATEUR DE CONTINUITE) ELIMINE LES BRUITS MAIS DIMINUE LE NOMBRE DE POINTS (DONC D'INFORMATION)

# TEST & LIMITES 4

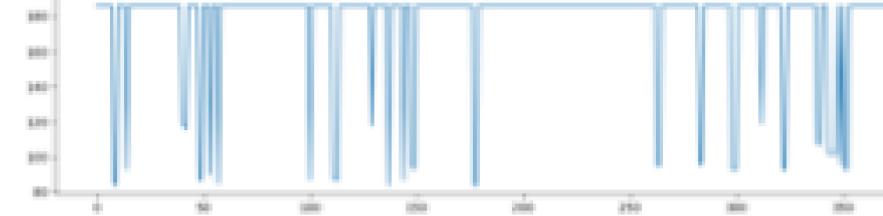
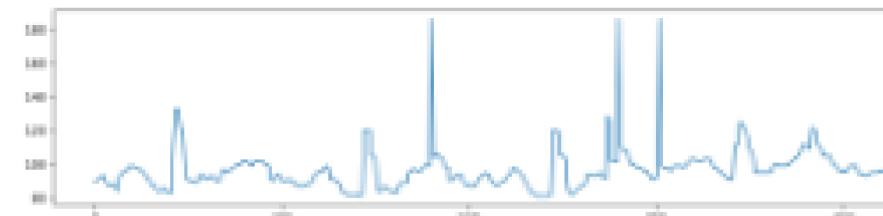
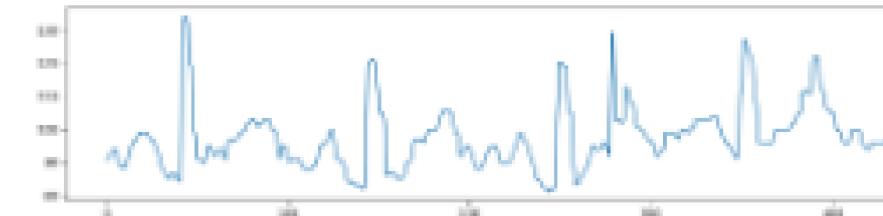
Threshold

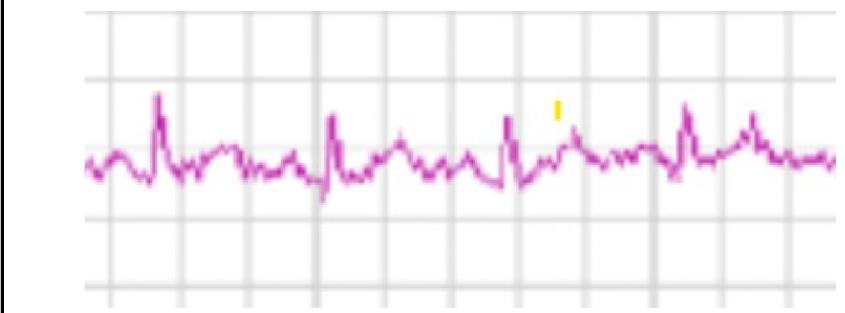
facteur de correction

resultats

commentaires

image réelle

20	10		rien ,pas de variation
80	10		rien ,pas de variation
120	10		variation non identique
180	10		variation presque identique au signal d'entrée mais avec des pics supplémentaires.
200	10		variation identique au signal d'entrée



LE THRESHOLD EST  
SUFFISANT POUR  
LA DIGITALISATION

# TEST & LIMITES 5

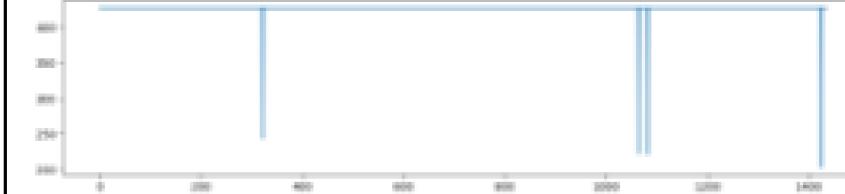
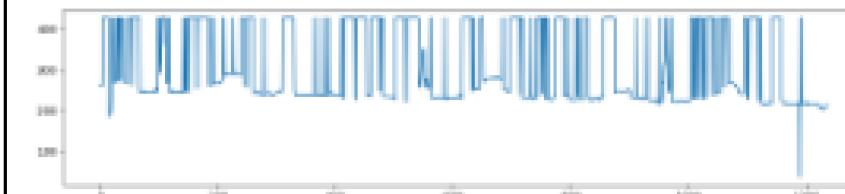
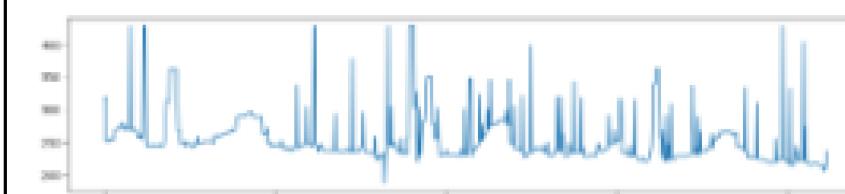
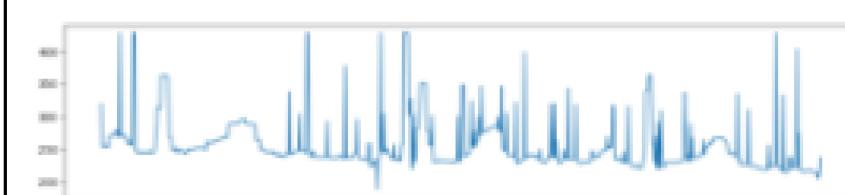
Threshold

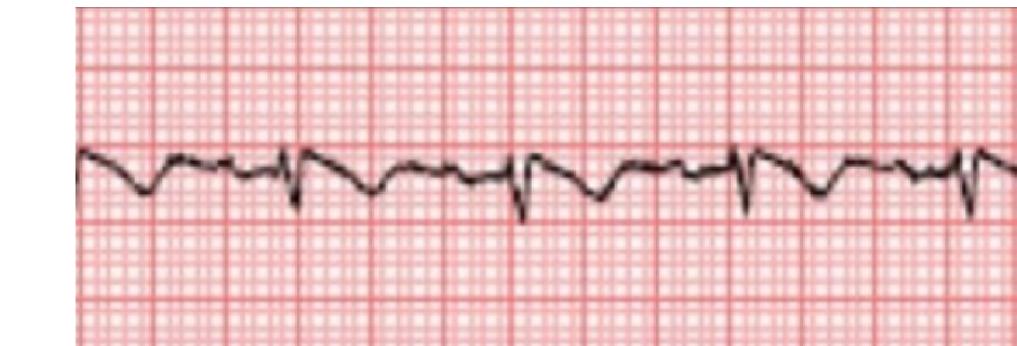
facteur de correction

resultats

commentaires

image réelle

20	10		rien ,pas de variation
80	10		rien ,pas de variation
120	10		petite variation
150	10		variation non identique
170	10		on a plusieurs pics supplémentaires
170	5		le facteur de correction n'a presque aucun influence sur l'élimination des pics supplémentaires

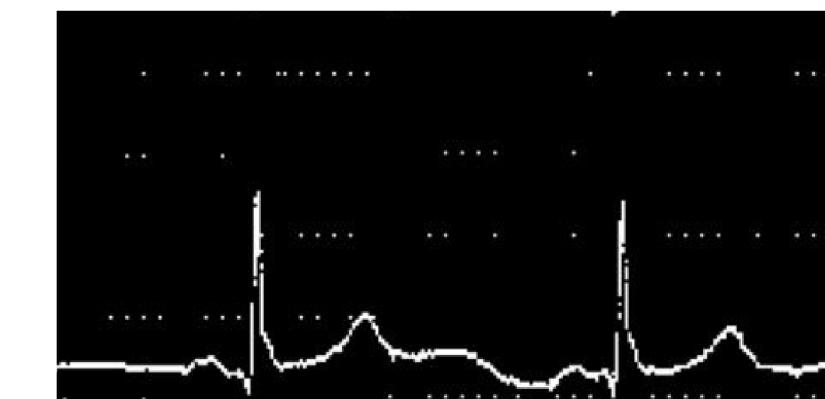
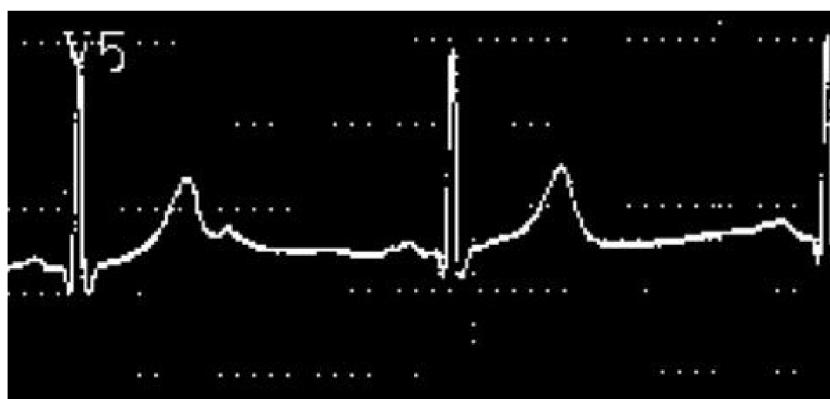


LE THRESHOLD EST  
INSUFFISANT POUR  
LA DIGITALISATION

# CONCLUSION

1. L'UTILISATION DE SEUILLAGE N'EST PAS AUTOMIQUE ON DOIT TOUJOURS TESTER DES DIFFERENTES VALEURS DE SEUILLAGE POUR TROUVER LE MEILLEUR RESULTAT.

2. PARFOIS ON NE PEUT PAS OBTENIR UN SEUILLAGE PARFAIT , ET ON OBSERVE DES BRUITS AVEC LE SIGNAL .



3. CE BRUIT ON NE PEUT PAS ELIMINER NI AVEC LE FACTEUR DE CORRECTION NI AVEC LE SUILLAGE , ET ON A CONSTATÉ QUE CE BRUIT DUE À LA SIMILARITÉ ENTRE LA COULEUR DE SIGNAL ET LA COULEUR DE GRID.

4. DANS LA SUITE ,ON VA TESTER QUELQUE FILTRES POUR ELIMINER LES BRUITS .