

Utilisation de l'entropie d'énergie de cisaillement et du classificateur de machine à vecteur de support pour traiter les signaux sismiques microsismiques et désertiques faibles

Comptes Rendus. Géoscience, Tome 352 (2020) no. 1, pp. 103-113.

Yue Li ; Shiyu Fan ; Chao Zhang ; Baojun Yang

(traduction post-éditée par N. Bacaër, suggestions d'amélioration : nicolas.bacaer@ird.fr)



Résumé

La détection de signaux de faible amplitude est une procédure clé dans l'exploration sismique microsismique et désertique de forage. Habituellement, les signaux sont difficiles à détecter en raison de leur faible amplitude et de leur contamination par le bruit. Pour résoudre ce problème, nous proposons une méthode combinant l'entropie d'énergie de cisaillement avec une machine à vecteurs de support (SVM) pour détecter des signaux de faible amplitude. Dans la méthode proposée, la caractéristique du signal est extraite à l'aide de l'entropie d'énergie de cisaillement. Le signal est plus faiblement représenté dans le domaine du cisaillement en raison de la caractéristique multi-échelles et multidirectionnelle de la transformée de cisaillement, qui favorise l'extraction des caractéristiques du signal. En outre, dans le calcul de l'entropie d'énergie de cisaillement, nous utilisons la corrélation des coefficients de cisaillement pour améliorer la différence entre le signal et le bruit dans le domaine du cisaillement. L'entropie d'énergie de cisaillement permet au SVM d'obtenir un résultat de classification plus précis par rapport à d'autres caractéristiques traditionnelles telles que l'amplitude et l'énergie. Les résultats des données synthétiques et de terrain montrent que notre méthode est plus efficace que le STA / LTA et le réseau de neurones convolutifs pour la détection de signaux microsismiques de faible amplitude et de signaux sismiques du désert.

1. Introduction

La technologie de surveillance de la fracturation hydraulique avec des signaux microsismiques est utilisée pour juger de la forme et de la tendance d'extension de la fracturation grâce à une analyse des données microsismiques enregistrées [Maxwell et Urbancic 2001]. Cependant, les signaux microsismiques de faible amplitude sont généralement submergés par un fort bruit [Zhang et Van der Baan 2018 a]. De même, de plus en plus de difficultés surviennent dans la détection des signaux sismiques du désert en raison de la similitude entre le signal et le bruit [Li et Li 2016]. Ainsi, nous devons développer une méthode de détection de signal sismique de faible amplitude précise et rapide. Différentes approches ont été proposées: la méthode STA / LTA détecte les signaux en fonction de la variation du rapport de l'amplitude ou de l'énergie dans le domaine temporel [Allen 1982]. Bien que le coût de calcul du STA / LTA soit faible, sa qualité diminue lorsque le rapport signal sur bruit (SNR) de l'enregistrement est faible. Le critère d'information Akaike (AIC) proposé ici est basé sur un modèle autorégressif utilisé pour déterminer les points limites de deux séquences stationnaires avec des propriétés statistiques différentes pour la détection du signal [Leonard 2000]. L'AIC ne fonctionne pas bien non plus pour la détection de signaux sismiques de faible amplitude. Les méthodes basées sur la détection de signaux de réseau neuronal convolutif (CNN) ont attiré beaucoup d'attention [Yuan et al. 2018 ; Samaneh et coll. 2018 ; Xiong et coll. 2018], mais leurs performances dépendent de l'existence de grands ensembles de données d'entraînement.

La transformée en ondelettes est largement utilisée pour la détection de signaux sismiques [Wang 2009]. Il s'agit d'une méthode basée sur l'analyse temps-fréquence, qui présente une grande flexibilité et un calcul rapide [Mallat et Hwang 1992]. Il existe deux types de méthodes de détection de signal utilisant la transformée en ondelettes: l'une utilise la transformée en ondelettes pour le prétraitement afin de supprimer le bruit [Wang 2009]; l'autre utilise certaines caractéristiques du signal dans le domaine des ondelettes telles que l'entropie d'énergie, l'entropie multi-échelles [Jia et al. 2016]. Cependant, la transformée en ondelettes ne peut pas effectuer de décomposition multidirectionnelle. Des méthodes d'analyse multidirectionnelle telles que la transformation en cisaillement sont proposées pour traiter ce problème [Lim 2010]. La transformée en cisaillement a une erreur d'approximation plus faible pour les signaux que les autres méthodes d'analyse multi-échelles et multidirectionnelles [Zhang et Van der Baan 2018 b, 2019].

Dans cet article, nous proposons une méthode combinant l'entropie d'énergie de cisaillement avec SVM pour détecter un signal microsismique et sismique. Le signal peut obtenir une représentation plus éparse dans le domaine de cisaillement en raison de la caractéristique multidirectionnelle de la transformée de cisaillement, qui favorise l'extraction des caractéristiques du signal. En outre, nous effectuons le traitement de corrélation à travers les échelles pour améliorer la différence entre le signal et le bruit dans le domaine cisaillement. Ensuite, nous calculons l'entropie d'énergie de cisaillement en tant que caractéristique du signal. Comparé aux caractéristiques traditionnelles telles que l'amplitude et l'énergie, SVM peut donner un résultat de classification plus précis avec l'entropie d'énergie de cisaillement. Le SVM a une précision plus élevée pour les petits ensembles de données d'entraînement que les méthodes de détection de signaux d'apprentissage en profondeur [Zhang et al. 2004]. Nous utilisons SVM au lieu d'un seuil pour obtenir une détection automatique. Cela évite les erreurs de jugement causées par la sélection des seuils. Les tests montrent que la méthode proposée peut détecter efficacement les signaux microsismiques et sismiques à faible SNR.

2. Entropie d'énergie de cisaillement

2.1. Shearlet

Guo et Labate [2006, 2007] ont combiné la théorie des ondelettes complexes avec l'analyse géométrique multi-échelles pour construire une représentation clairsemée d'une fonction multidimensionnelle: la représentation cisaillement. Les cisaillements $\psi_{j,k,m}$ sont un exemple particulier d'ondelettes composites dans $L^2(\mathbb{R}^2)$, qui peut être construite en appliquant des dilatations, des transformations de cisaillement et des traductions à une fonction mère appropriée. En dimension 2, les cisaillements peuvent s'écrire sous la forme suivante:

$$\{\psi_{j,k,m}(x) = |\det A_0|^{j/2} \psi(S_0^k A_0^j x - m): \\ j, k \in \mathbb{Z}, m \in \mathbb{Z}^2\}, \quad (1)$$

où ψ est une fonction appropriée à bande limitée dans $L^2(\mathbb{R}^2)$, pour plus de détails, voir [Easley et al. 2008]. Les paramètres j, k, m représentent respectivement l'échelle, la direction et la position. A_0 est une matrice de dilatation anisotrope qui est associée à la transformation d'échelle, et S_0 est appelée matrice de cisaillement qui est associée à la transformation de direction, ici les deux matrices sont données comme:

$A_0 = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$, $S_0 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. La transformée en cisaillement d'une fonction pour une certaine échelle, direction et position est définie comme suit:

$$S_f(j, k, m) = \langle f, \psi_{j,k,m} \rangle, \quad (2)$$

où $S_f(j, k, m)$ représentent les coefficients décomposés après la transformée cisaillement. Le symbole $\langle \cdot, \cdot \rangle$ désigne le produit scalaire interne.

2.2. Entropie d'énergie de cisaillement

La transformée de cisaillement peut décomposer les données en différentes échelles et directions, ce qui donne les coefficients de cisaillement $S_f(j, k, m)$. L'énergie des signaux n'est concentrée que dans quelques

directions en raison de leur corrélation spatiale à ces creux uniquement [Zhao et al. 2016 ; Zhang et Van der Baan 2019]. Nous définissons l'énergie du coefficient de cisaillement comme:

$$E(j, k) = \sum_m S_f^2(j, k, m) \quad (3)$$

où $E(j, k)$ est l'énergie du coefficient de cisaillement dans la j ème échelle et la k ème direction. La différence entre le signal et le bruit devient faible lorsque le SNR est faible. Ainsi, nous décidons de faire la corrélation entre deux échelles adjacentes pour mettre en évidence les caractéristiques des signaux. L'énergie de corrélation entre deux échelles adjacentes est définie comme suit:

$$ES_{j,j+1}(k) = E(j, k) \cdot E(j + 1, k), \quad (4)$$

où $E(j, k)$ et $E(j + 1, k)$ expriment respectivement l'énergie de deux échelles adjacentes dans la même direction. Grâce à (4), nous pouvons obtenir la direction du signal par la plus grande distribution d'énergie. Ensuite, nous faisons la corrélation avec les coefficients de direction du signal entre deux échelles adjacentes pour améliorer davantage les coefficients associés aux signaux. Enfin, nous divisons les coefficients améliorés en certains segments, l'énergie de ces segments est exprimée par $E_1, E_2, \dots, E_\varepsilon$. L'énergie totale E est égale à la somme de $E_1, E_2, \dots, E_\varepsilon$. Posons $p_\tau = E_\tau/E$, et $\sum_{\tau=1}^\varepsilon p_\tau = 1$. L'entropie d'énergie de cisaillement est définie comme:

$$W = - \sum_\tau p_\tau \log(p_\tau). \quad (5)$$

L'entropie est utilisée pour mesurer le caractère aléatoire des signaux. Dans le processus de détection de signal, le caractère aléatoire fait que le bruit a des valeurs d'entropie plus grandes que les signaux [Rezek et Roberts 1998]. Puisque le signal et le bruit présentent des caractéristiques différentes dans le domaine du cisaillement, l'entropie d'énergie de cisaillement proposée peut encore améliorer la différence entre le signal et le bruit.

3. Détection de signal basée sur l'entropie d'énergie de cisaillement et SVM

SVM est un classificateur et sa théorie de base est la maximisation des marges. Il a une structure simple et un faible coût de calcul [Adankon et Cheriet 2009 ; Hu et coll. 2013]. Pour la détection de signal, nous obtenons un classificateur SVM via l'ensemble d'apprentissage, puis nous pouvons utiliser ce classificateur SVM entraîné pour déterminer si l'entrée représente des signaux. Les théories spécifiques de SVM sont les suivantes: Le principe principal de la classification SVM est de trouver la surface de classification optimale. La surface de classification optimale peut être transformée en résolvant l'équation suivante [Widodo et Yang 2007]:

$$\min_{\omega, b, \xi_i} \frac{1}{2} \|\omega\|^2 + \gamma \sum_{i=1}^n \xi_i, \quad (6)$$

tel que

$$y_i(\omega^T \phi(z_i) + b) = 1 - \xi_i, \quad \xi_i \geq 0, i = 1, \dots, n \quad (7)$$

où ω est le vecteur normal, qui détermine la direction de l'hyperplan; γ est le paramètre de pénalité qui impose un compromis entre l'erreur d'apprentissage et la complexité du modèle; z_i est l'échantillon d'entrée et y_i correspond à son étiquette, la valeur de y_i pourrait être + 1 ou - 1; i représente le numéro de l'échantillon; b est le décalage, qui détermine la distance entre l'hyperplan et l'origine; ξ_i est la variable d'écart, qui est utilisée pour mesurer l'écart des données par rapport aux conditions idéales. ϕ est une fonction non linéaire, elle peut mapper les données d'entrée dans un espace de caractéristiques de grande dimension. Les détails de la solution de ce problème d'optimisation, voir Widodo et Yang [2007]. Après avoir résolu l'équation ci-dessus, la fonction de décision finale est donnée par:

$$D(z) = \text{sign} \left(\sum_{i=1}^n c_i y_i K(z, z_i) + b \right), \quad (8)$$

où c_i est le multiplicateur de Lagrange. Grâce à (8), nous pouvons obtenir un classificateur à deux classes pour déterminer si les entrées sont des signaux ou non. Si $\sum_{i=1}^n c_i y_i K(z, z_i) + b > 0$, $D(z) = +1$ représente l'échantillon positif, inversement, $D(z) = -1$ représente l'échantillon négatif. $K(z, z_i)$ représente le noyau de la fonction de base radiale gaussienne, il est défini comme:

$$K(z, z_i) = \exp \left(\frac{-\|z - z_i\|^2}{2\sigma^2} \right), \quad (9)$$

où σ est le paramètre de largeur d'une fonction, qui contrôle la plage radiale de la fonction. Le paramètre de pénalité γ et le paramètre σ dans les fonctions du noyau sont deux facteurs clés affectant la précision du classifieur SVM. Si le paramètre de pénalité est trop petit ou trop grand, l'algorithme d'apprentissage sera sous-ajusté ou sur-ajusté. Lorsque σ est réglé trop petit, le rayon de la zone d'influence des vecteurs de support inclut uniquement le vecteur de support lui-même et aucune quantité de régularisation avec γ ne pourra empêcher le sur-ajustement. Inversement, si σ est trop grand, le modèle ne peut pas saisir la complexité ou la «forme» des données. Par conséquent, dans cet article, nous utilisons la validation croisée de type «laisser-un-out» pour obtenir des paires de γ et σ qui peuvent atteindre une grande précision.

Pour entraîner un classificateur SVM, nous choisissons d'abord des échantillons de signal et de bruit pour créer le jeu de données d'apprentissage, où le nombre d'échantillons de bruit et d'échantillons de signal est le même. Les échantillons de signal sont générés en utilisant des ondelettes de Ricker avec différentes amplitudes et fréquences dominantes. Les échantillons de bruit sont générés en utilisant une combinaison de bruit gaussien blanc (WGN) et de bruit réel avec différents niveaux. Ensuite, nous calculons les entropies d'énergie de cisaillement de ces échantillons et les mettons dans SVM pour former un classificateur final. Lorsque nous avons obtenu un classificateur SVM, la détection du signal se fait en deux étapes:

Étape 1: Nous décomposons les données d'entrée dans un domaine de cisaillement pour trouver les coefficients de cisaillement dans la direction du signal. Ensuite, nous les corrélons avec une échelle adjacente pour améliorer les coefficients du signal.

Étape 2: Nous calculons les entropies d'énergie des coefficients de cisaillement améliorés obtenus à partir de l'étape 1. Ensuite, les entropies d'énergie sont utilisées comme entrée dans le classificateur SVM formé pour l'identification du signal.

4. Expériences

4.1. Données microsismiques synthétiques

Pour vérifier la fiabilité de cette méthode, nous avons simulé le signal microsismique contenant l'onde P et l'onde S. L'amplitude de l'onde P est plus petite que celle de l'onde S, comme le montre la figure 1. La fréquence du signal microsismique est élevée et la fréquence dominante de l'ondelette réellement reçue est d'environ 200 Hz. [Gao et coll. 2018 ; Maxwell et Urbancic 2001]. Nous avons fixé les fréquences dominantes de deux signaux microsismiques à 200 Hz [Zhu et al. 2016] qui est proche de la situation réelle. La fréquence d'échantillonnage est de 1000 Hz. Les amplitudes de l'onde P et Svague sont de 0,5 et 0,2. La figure 2 est l'enregistrement bruyant où le bruit gaussien blanc (WGN) a été ajouté avec le SNR de -8 dB.

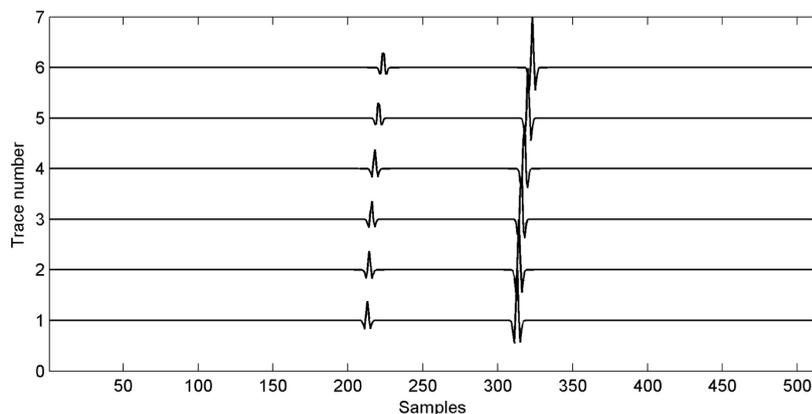


Figure 1. Enregistrement microsismique synthétique.

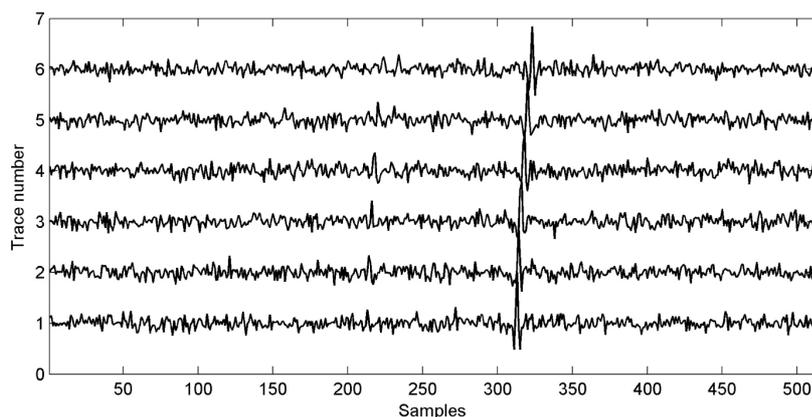


Figure 2. Enregistrement bruyant contaminé par WGN avec un SNR de -8 dB.

L'enregistrement est décomposé en 4 échelles par la transformée en cisaillement. Étant donné que la plupart des signaux microsismiques se concentrent sur les échelles de haute fréquence et que les échelles plus grandes correspondent à des fréquences plus élevées, nous choisissons les échelles les plus grandes pour la détection. Dans cet article, nous divisons les enregistrements en 4 échelles et nous traitons les deux plus grandes échelles: 3e et 4e échelles. La figure 3 montre les coefficients de cisaillement dans différentes directions à la 3e échelle. Nous pouvons voir que la plupart des signaux sont concentrés dans la 4ème direction, tandis que le bruit est distribué dans toutes les directions. L'énergie de corrélation de la première trace entre les échelles adjacentes est représentée sur la figure 4. On voit que l'énergie de corrélation dans la 4ème direction est évidemment plus grande que les autres. Cela indique que nous pouvons obtenir des directions de signal plus précises grâce au traitement de corrélation. La figure 5 (a) montre la forme d'onde de la première trace. On voit que le signal microsismique avant de faible amplitude est submergé par un bruit difficile à identifier. Les coefficients de cisaillement après le traitement de corrélation sont représentés sur la figure 5 (b), les coefficients de cisaillement associés aux signaux microsismiques sont tous distingués du bruit, ce qui facilite l'extraction des caractéristiques d'entropie d'énergie et la détection SVM.

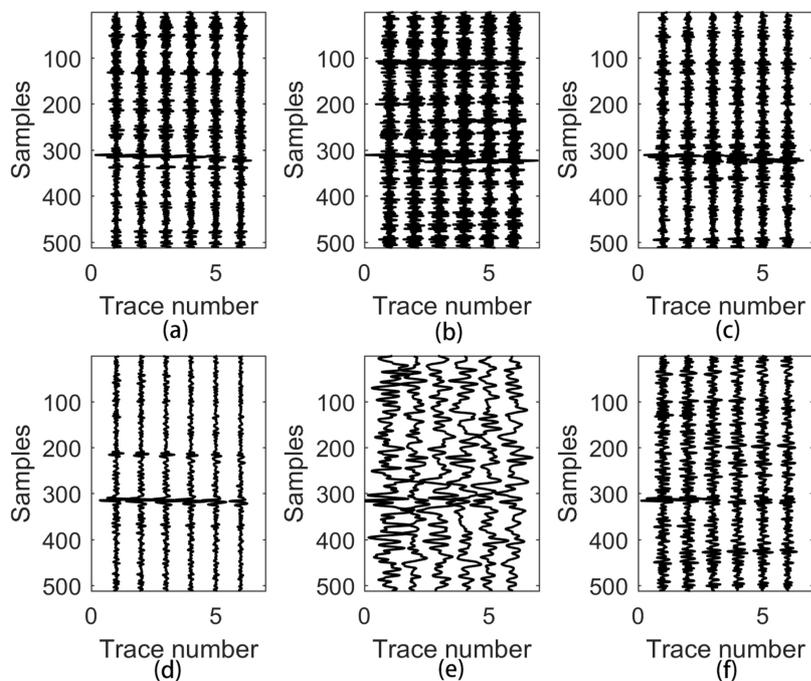


Figure 3. (a) - (f) Correspondent aux coefficients de cisailment de la première à la sixième direction à la troisième échelle.

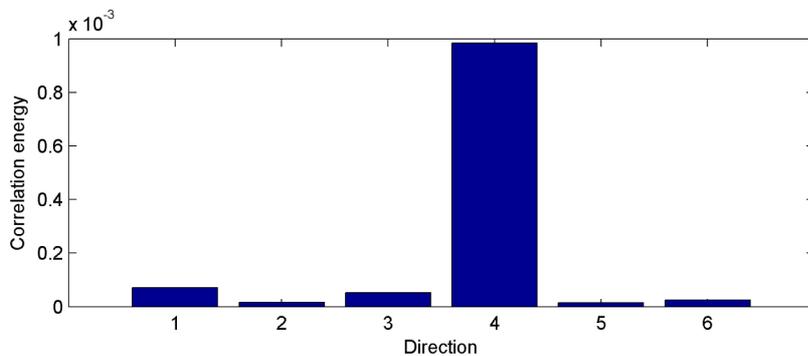


Figure 4. L'énergie de corrélation de la première trace entre les échelles adjacentes ($j = 3$ et $j = 4$).

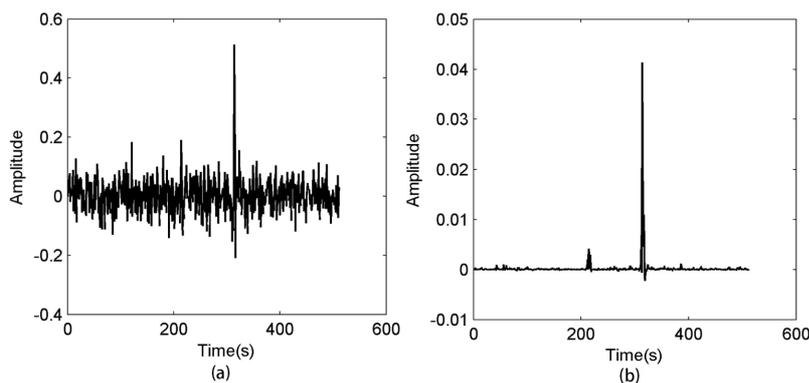


Figure 5. (a) La forme d'onde de la première trace de la figure 2 . (b) Coefficients de Shearlet de la première trace après traitement de corrélation.

Vient ensuite le processus de formation du SVM. Pour l'ensemble d'apprentissage du classificateur SVM, nous avons choisi ici 1000 groupes d'échantillons de signal et 1000 groupes d'échantillons de bruit pour former l'ensemble d'apprentissage. Les échantillons de signal ont été générés de manière aléatoire en utilisant des ondelettes de Ricker avec des amplitudes allant de 0,1 à 1 et des fréquences allant de 100 à 500 Hz. Les échantillons de bruit ont été générés aléatoirement avec WGN et du bruit réel avec différents niveaux. Grâce à la formation, nous pouvons obtenir le classificateur SVM formé.

Nous avons comparé la méthode proposée avec le STA / LTA et CNN, et les résultats de détection sont présentés dans la figure 6 . La méthode proposée peut détecter avec précision deux signaux microsismiques et il n'y a pas d'erreur de jugement. Dans le cas d'un SNR faible, nous pouvons toujours détecter avec

précision les signaux microsismiques, alors qu'il existe de nombreuses erreurs de détection avec les méthodes STA / LTA et CNN. Il est difficile de choisir un seuil approprié pour la méthode STA / LTA surtout lorsque le SNR est faible. CNN réalise une détection précise à un SNR élevé, mais cela ne fonctionne pas non plus pour un SNR bas. En revanche, la méthode proposée évite le réglage du seuil et le résultat de détection de la méthode proposée est précis à faible SNR.

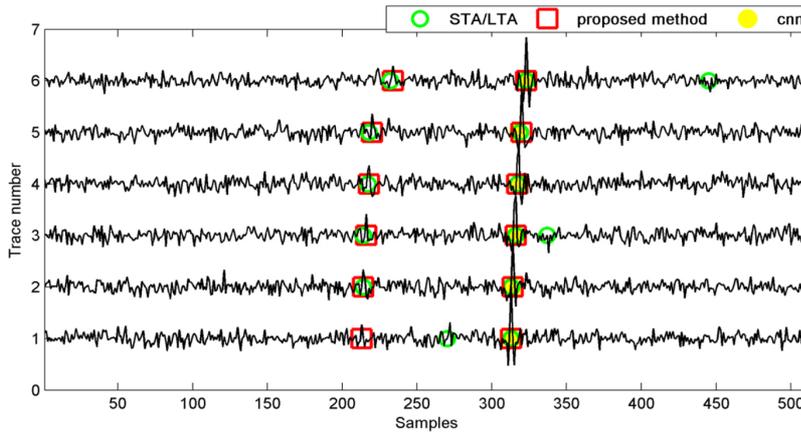


Figure 6. Résultats de détection de la figure 2 par la méthode proposée, STA / LTA et CNN.

4.2. Expériences statistiques

Afin de vérifier la validité de la méthode proposée, nous prélevons 500 groupes d'échantillons pour réaliser l'expérience. Une ondelette de Ricker est utilisée pour simuler le signal. La fréquence dominante des ondelettes de Ricker est de 200 Hz. La fréquence d'échantillonnage est de 1000 Hz. Nous ajoutons différents WGN aux ondelettes de Ricker pures et modifions les amplitudes des ondelettes de Ricker. Les résultats de détection de la méthode proposée, STA / LTA et CNN sont listés dans les tableaux 1, 2 et 3, respectivement.

Tableau 1. Précision de la détection du signal synthétique par la méthode proposée

Amplitude	SNR			
	-5 dB	-6 dB	-7 dB	-8 dB
1	100%	99.2%	92.4%	87%
0.5	100%	98.6%	92.2%	85%
0.2	99.8%	98.2%	91.4%	80%

Tableau 2. Précision de la détection du signal synthétique par STA / LTA

Amplitude	SNR			
	-5 dB	-6 dB	-7 dB	-8 dB
1	97%	90.2%	86.4%	70%
0.5	90%	89.5%	85.6%	69%
0.2	87%	87%	83.4%	63.4%

Tableau 3. Précision de la détection du signal synthétique par CNN

Amplitude	SNR			
	-5 dB	-6 dB	-7 dB	-8 dB
1	97.8%	94%	89%	76%
0.5	92.5%	89.6%	85%	73.1%
0.2	91%	88%	84.6%	68.2%

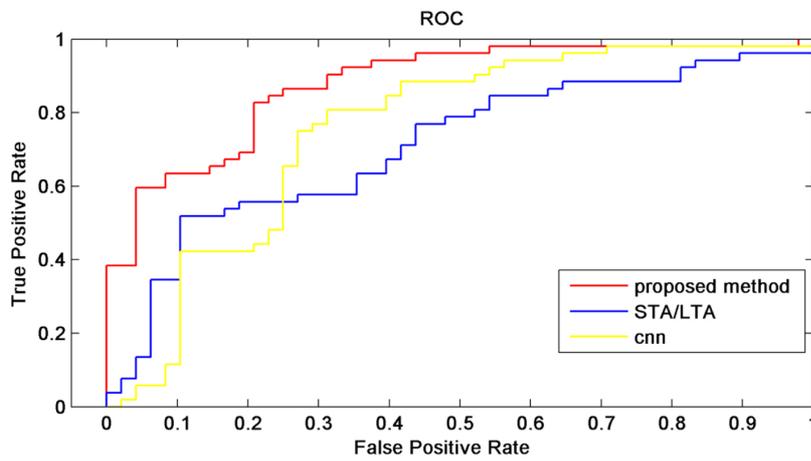


Figure 7. Courbes ROC de la méthode proposée, STA / LTA et CNN.

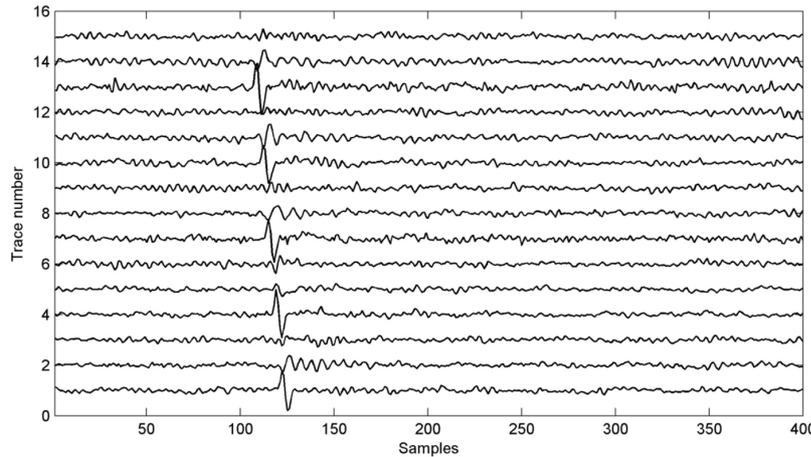


Figure 8. L'enregistrement microsismique de terrain.

À partir des tableaux 1, 2 et 3, nous pouvons voir que la précision de détection diminue à mesure que le SNR et l'amplitude du signal diminuent. La méthode proposée permet d'obtenir une précision plus élevée à faible SNR par rapport au STA / LTA et au CNN. La précision de la méthode proposée est de 80% lorsque le SNR est de -8 dB. La précision de STA / LTA et CNN n'est que de 60% à 76%. Nous dessinons des courbes de caractéristiques de fonctionnement du récepteur (ROC) de 100 groupes d'échantillons lorsque le SNR est de -8 dB et que l'amplitude de l'ondelette de Ricker est de 0,2 comme le montre la figure 7. Une méthode fonctionne plus précisément si sa courbe ROC est plus proche du coin supérieur gauche. L'AUC est la valeur de l'aire sous la courbe ROC qui est utilisée pour conclure si un classificateur est excellent. Lorsqu'une valeur AUC d'une méthode est plus grande, cette méthode a une précision plus élevée. Les valeurs AUC des méthodes proposées, STA / LTA et CNN, sont respectivement de 0,8570, 0,7091 et 0,7536. Comme le montre la figure 7, la courbe ROC de la méthode proposée est plus proche du coin supérieur gauche et sa valeur AUC est plus grande que celle des deux autres méthodes. Ainsi, la méthode proposée a une précision plus élevée que les deux autres méthodes. Pour vérifier davantage l'efficacité de la méthode proposée, nous ajoutons un bruit réel avec différents niveaux aux signaux microsismiques synthétiques ci-dessus pour composer 300 échantillons de groupe: l'amplitude des signaux microsismiques synthétiques est de 1 dans les 100 premiers groupes, 0,5 dans les 100 seconds groupes, et 0,2 dans les 100 derniers groupes. Les résultats de détection de la méthode proposée, STA / LTA et CNN sont énumérés dans le tableau 4. Nous pouvons voir que la méthode proposée a toujours la plus grande précision dans tous les cas.

Tableau 4. Précision de détection du signal synthétique contaminé par un bruit microsismique réel

Amplitude	1	0.5	0.2
La méthode proposée	100%	95%	82%
STA / LTA	93%	71%	60%
CNN	95.6%	77%	70%

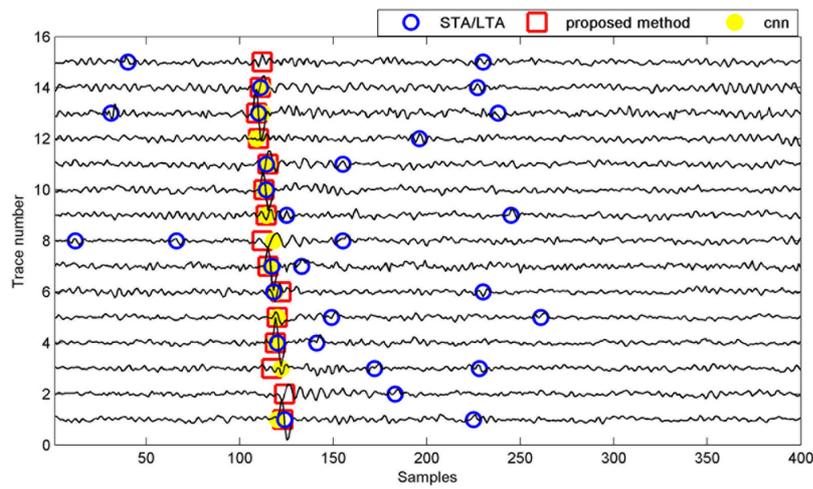


Figure 9. Résultats de détection de la figure 8 par la méthode proposée, STA / LTA et CNN.

4.3. Données sismiques microsismiques et désertiques réelles

Afin de prouver la validité de la méthode proposée dans le traitement de données réelles, nous avons sélectionné pour l'analyse et le traitement un enregistrement microsismique réel avec 15 traces dans une certaine région de la Chine, comme le montre la figure 8 (cet enregistrement est également utilisé dans certains enregistrements publiés articles tels que Li et al. [2018], Zhu et al. [2016]). Nous pouvons voir que les amplitudes des signaux dans certaines traces sont trop faibles pour être détectées. Ces signaux faibles ont également des formes d'onde similaires à celles du bruit. Les résultats de détection de la méthode proposée, STA / LTA et CNN sont présentés à la figure 9. La méthode STA / LTA ne fonctionne pas bien dans cette situation. Certains bruits avec une amplitude et une forme d'onde similaires sont identifiés comme un signal par la méthode STA / LTA. CNN identifie les signaux de faible amplitude comme du bruit et ne peuvent donc pas être détectés. La détection de signaux microsismiques faibles à haute fréquence est très difficile pour les méthodes basées sur l'énergie ou l'apprentissage profond. Ces signaux faibles peuvent être détectés avec précision par la méthode proposée qui fonctionne bien dans un SNR bas.

Ce défi apparaît également dans le dossier du désert. Le bruit aléatoire dans un enregistrement sismique du désert est concentré sur les bandes de basses fréquences qui chevauchent souvent le signal sismique. Cela pose de grandes difficultés dans la détection des signaux sismiques du désert. La figure 10 montre un véritable record du désert avec 30 traces. Nous pouvons voir que le bruit du désert a une basse fréquence et une forme d'onde similaire à celle du signal. Dans cet exemple, nous comparons également la méthode proposée avec le STA / LTA et CNN, et leurs résultats de détection sont présentés dans la figure 11. Les performances du STA / LTA diminuent considérablement pour la détection de signaux sismiques faibles dans le désert. La plupart des signaux peuvent être détectés par le CNN, mais certaines séquences de bruit sont également identifiées comme des signaux. La méthode proposée a la plus grande précision par rapport à la STA / LTA et CNN.

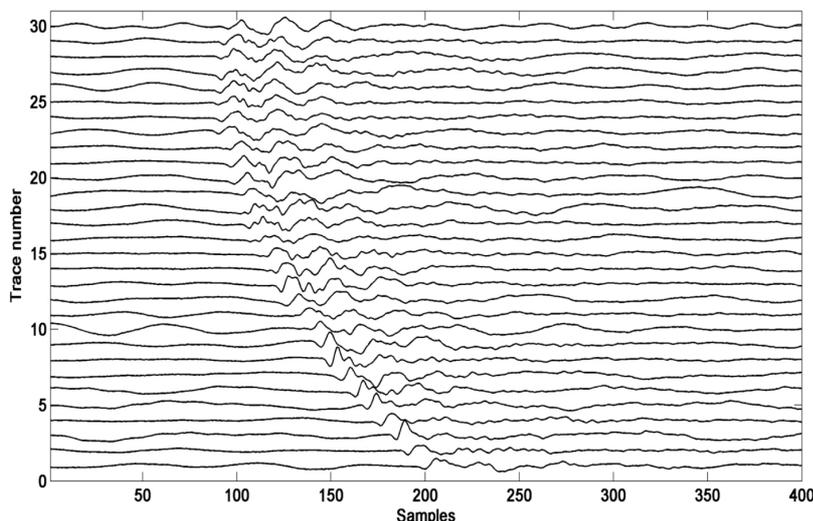


Figure 10. Le record du désert sur le terrain.

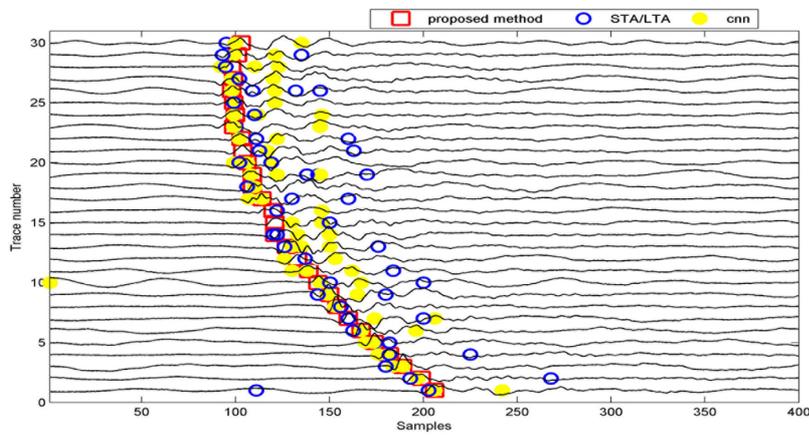


Figure 11. Résultats de détection de la figure 10 par la méthode proposée, STA / LTA et CNN.

5. Conclusion et discussion

Dans cet article, l'entropie d'énergie de cisaillement est utilisée comme caractéristique de signal pour détecter des signaux efficaces. Cette fonction permet de mieux différencier le signal du bruit qu'un simple calcul d'énergie. Le classificateur SVM est formé avec la fonction extraite. Il n'est pas nécessaire de sélectionner un seuil de détection de signal raisonnable avec l'utilisation de SVM. Les erreurs de détection dues à une sélection de seuil inappropriée peuvent être réduites. Le signal est détecté de manière intelligente par la méthode proposée. Un grand nombre d'expériences démontrent le potentiel et la supériorité de cette méthode proposée pour la détection de signaux microsismiques et sismiques de faible amplitude. Cela nécessite beaucoup moins de jeux de données de formation que les méthodes de détection de signal basées sur CNN.

Dans la méthode proposée, l'entropie d'énergie de cisaillement est utilisée comme une caractéristique pour former un classificateur SVM pour la détection de signal. Un classificateur plus efficace peut être obtenu en combinant plus de fonctionnalités. De plus, une sélection de fonction noyau plus raffinée dans le SVM, telle qu'une fonction noyau mixte, peut encore augmenter la précision de classification.

Remerciements

Cette recherche est financée par les Fondations nationales des sciences naturelles de Chine (sous subventions 41730422, 41974143, 41574096). Nous remercions les efforts du professeur Ghislain de Marsily. Nous remercions également les réviseurs anonymes pour leurs commentaires et suggestions.

Bibliographie

- [Adankon and Cheriet, 2009] M. M. Adankon; M. Cheriet Model selection for the LS-SVM. Application to handwriting recognition, *Pattern Recognit.*, Volume 42 (2009) no. 12, pp. 3264-3270
- [Allen, 1982] R. V. Allen Automatic phase pickers their present use and future prospects, *Bull. Seismol. Soc. Am.*, Volume 72 (1982) no. 6, pp. 225-242
- [Easley et al., 2008] Q. Easley; D. Labate; W. Q. Lim Sparse directional image representations using the discrete Shearlet transform, *Appl. Comput. Harmonic Anal.*, Volume 25 (2008) no. 1, pp. 25-46 | MR 2419703 | Zbl 1147.68794
- [Gao et al., 2018] D. W. Gao; Z. H. Zhao; H. Zhang Micro-seismic real-time monitoring technology for staged fracturing wells in Jiaoshiba Block, *Well Test.*, Volume 27 (2018) no. 5, pp. 42-49
- [Guo and Labate, 2006] K. Guo; D. Labate The theory of wavelets with composite dilations, *Harmon. Anal. Appl.*, Volume 10 (2006), pp. 231-318 | Zbl 1129.42433
- [Guo and Labate, 2007] K. Guo; D. Labate Optimally sparse multidimensional representation using the discrete cisaillement transform, *SIAM J. Math. Anal.*, Volume 39 (2007) no. 1, pp. 298-318
- [Hu et al., 2013] R. H. Hu; P. H. Lou; D. B. Tang Fault diagnosis of rolling bearings based on EMD and parameter adaptive support vector machine, *Comput.-Integr. Manuf. Syst.*, Volume 19 (2013) no. 2, pp. 438-447
- [Jia et al., 2016] R. S. Jia; H. M. Sun; Y. J. Peng; Y. Q. Liang; X. M. Lu Automatic event detection in low SNR microseismic signals based on multi-scale permutation entropy and a support vector

machine, *J. Seismol.*, Volume 21 (2016) no. 4, pp. 1-14

- [Leonard, 2000] M. Leonard Comparison of manual and automatic onset timepicking, *Bull. Seismol. Soc. Am.*, Volume 90 (2000), pp. 1384-1390
- [Li and Li, 2016] G. H. Li; Y. Li Random noise of seismic exploration in desert modeling and its applying in noise attenuation, *Chin. J. Geophys.*, Volume 59 (2016) no. 2, pp. 682-692
- [Li et al., 2018] Y. Li; Y. Wang; H. B. Lin; T. Zhong First arrival time picking for microseismic data based on DWSW algorithm, *J. Seismol.*, Volume 22 (2018) no. 4, pp. 833-840
- [Lim, 2010] W. Q. Lim The discrete cisaillement transform: a new directional transform and compactly supported cisaillement frames, *IEEE Trans. Image Process.*, Volume 19 (2010) no. 5, pp. 1166-1180 | MR 2723739 | Zbl 1371.94233
- [Mallat and Hwang, 1992] S. Mallat; W. L. Hwang Singularity detection and processing with wavelets, *IEEE Trans. Inf. Theory*, Volume 38 (1992), pp. 617-643 | MR 1162218 | Zbl 0745.93073
- [Maxwell and Urbancic, 2001] S. C. Maxwell; T. I. Urbancic The role of passive microseismic monitoring in the instrumented oil field, *The Leading Edge*, Volume 20 (2001), pp. 636-639
- [Rezek and Roberts, 1998] I. A. Rezek; S. J. Roberts Stochastic complexity measures for physiological signal analysis, *IEEE Trans. Biomed. Eng.*, Volume 45 (1998) no. 9, pp. 1186-1191
- [Samaneh et al., 2018] R. P. Samaneh; M. Milos; P. Supathorn A novel framework for intelligent signal detection via artificial neural networks for cyclic voltammetry in pyroprocessing technology, *Ann. Nucl. Energy*, Volume 111 (2018), pp. 242-254
- [Wang, 2009] Z. Wang Detection of weak earthquake signals based on wavelet analysis, *Prog. Explor. Geophys.*, Volume 32 (2009) no. 3, pp. 182-185
- [Widodo and Yang, 2007] A. Widodo; B. S. Yang Support vector machine in machine condition monitoring and fault diagnosis, *Mech. Syst. Signal Process.*, Volume 21 (2007) no. 6, pp. 2560-2574
- [Xiong et al., 2018] Z. H. Xiong; M. P. Nash; E. Cheng; V. V. Fedorov; M. K. Stiles; J. C. Zhao ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network, *Physiol. Meas.*, Volume 39 (2018) no. 9, 094006 pages
- [Yuan et al., 2018] S. Yuan; J. Liu; S. Wang; T. Wang; P. Shi Seismic waveform classification and first-break picking using convolution neural networks, *IEEE Geosci. Remote Sens. Lett.*, Volume 15 (2018) no. 2, pp. 272-276
- [Zhang and Van der Baan, 2018a] C. Zhang; M. Van der Baan A denoising framework for microseismic and reflection seismic data based on block matching, *Geophysics*, Volume 83 (2018a) no. 5, p. V283-V292
- [Zhang and Van der Baan, 2018b] C. Zhang; M. Van der Baan Multicomponent microseismic data denoising by 3D cisaillement transform, *Geophysics*, Volume 83 (2018b) no. 3, p. A45-A51
- [Zhang and Van der Baan, 2019] C. Zhang; M. Van der Baan Strong random noise attenuation by cisaillement transform and time-frequency peak filtering, *Geophysics*, Volume 84 (2019) no. 6, p. V319-V331
- [Zhang et al., 2004] X. G. Zhang; J. J. Xu; G. Y. Ge Defects recognition on X-ray images for weld inspection using SVM, *Proceedings of 2004 International Conference on Machine Learning and Cybernetics*, Volume 6, IEEE, 2004, pp. 3721-3725
- [Zhao et al., 2016] X. Zhao; Y. Li; G. H. Zhuang; C. Zhang; X. Han 2-D TFPF based on Contourlet transform for seismic random noise attenuation, *J. Appl. Geophys.*, Volume 129 (2016), pp. 158-166
- [Zhu et al., 2016] D. Zhu; Y. Li; C. Zhang Automatic time picking for microseismic data based on a fuzzy C-means clustering algorithm, *IEEE Geosci. Remote Sens. Lett.*, Volume 13 (2016) no. 12, pp. 1900-1904