

JOURNAL INTERNATIONAL DE TECHNOLOGIE, DE L'INNOVATION,
DE LA PHYSIQUE, DE L'ENERGIE ET DE L'ENVIRONNEMENT

**Approche par neurolattice
pour la prédiction des angles
de talus naturel des matières divisées**

Luc Émile Brunet, Jean-Marin Brunet



ISSN : 2428-8500

DOI : 10.18145/jitipee.v5i1.219

Approche par neurolattice pour la prédiction des angles de talus naturel des matières divisées

L. E. Brunet⁽¹⁾, J-M Brunet⁽²⁾

⁽¹⁾ R&D Mediation, 71 av. de Lattre de Tassigny, Bourges, France

⁽²⁾ R&D Mediation Rhône-Alpes, 45 rue de Serbie, 73000 Chambéry, France

luc.brunet@rd-mediation.fr

Jean-Marin.brunet@rd-mediation.fr

Résumé - *La modélisation du comportement des matériaux granulaires est délicate par les approches classiques, car ceux-ci se comportent selon des caractéristiques hybrides entre les fluides et les solides. Cette complexité permet d'envisager une modélisation combinant des automates cellulaires et des réseaux de neurones profonds. Cette approche permet de pré - entraîner des réseaux de neurones avec des données artificielles plus facile et moins coûteuses à obtenir que des données réelles, dans l'optique de pouvoir réduire le temps d'apprentissage et la quantité de données réelles nécessaires à l'ajustement fin du modèle. Nous avons évalué cette approche neurolattice sur la prédiction des angles de talus naturels qui représente un problème important dans les industries manipulant des solides pulvérulents (géologie, agroalimentaire, pharmacie, cosmétique...). Un automate cellulaire implanté dans une classe openAI GYM a servi de base d'apprentissage à un réseau de neurones LSTM (Long Short Term Memory) pour relier la distribution granulométrique à un angle de talus après écoulement. Le réseau de neurones obtenu prédit les angles de talus naturel avec une précision de plus de 98%.*

Mots clés : Deep learning – Matériaux granulaires – Formulation

DOI : 10.18145/jitipee.v5i1.219

Acronymes

API	Application Programming Interface
DEM	Discrete Elements Method
GPU	Graphic Processing Unit
GRU	Gated Recurrent Units
LSTM	Long Short Term Memory
RN	Réseau de neurones
TPU	Tensorial Processing Unit

Introduction

L'angle de talus naturel est l'angle formé par des matériaux granulaires lorsqu'ils forment un entassement. Très variable- par exemple de 15° pour l'argile humide à 40° pour la cendre- cette caractéristique de ces matériaux présente de nombreux intérêts industriels.



Figure 1: Tas de Maïs (source Wikipedia)

En effet, les procédés de formulation nécessitant d'admettre des matériaux granulaires dans un silo ou un réacteur sont présents dans des secteurs aussi variés que l'agroalimentaire (Figure 1), la cosmétique, l'industrie du bâtiment, la pharmacie ou encore la métallurgie ... En outre, l'angle de talus naturel, aussi appelé angle de repos, conditionne des phénomènes complexes tels que l'érosion des côtes ou la constitution de formations géologiques (Figure 2).

Cette caractéristique, importante pour l'industrie, est difficilement modélisable et prédictible car elle dépend d'un nombre important de paramètres dont certains sont eux-mêmes difficilement quantifiables (propriétés de surface, rugosité, dimension fractale par exemple), ce qui est source de perturbation dans des domaines les plus divers de l'industrie. En géologie, les produits de l'érosion, une fois transportés, forment souvent des empilements sédimentaires susceptibles d'être déstabilisés si la pente de l'empilement est plus forte que l'angle de repos et ainsi présenter des dangers. C'est une problématique fréquemment rencontrée lors du phénomène d'érosion des côtes.



Figure 2: Formation géologique Isfjord (Norvège) (source Wikipedia)

Pour l'industrie pharmaceutique, la formulation de certains médicaments à base de poudres nécessite de remplir des cavités à l'aide d'un « racleur ». La vitesse du racleur définit la qualité du remplissage de la cavité, et ce sont les propriétés de la poudre qui déterminent la vitesse optimale de remplissage [1].

Dans l'industrie agroalimentaire, les problématiques sont principalement liées au stockage et au déchargement des grains dans les silos [2] ou lors du traitement.

L'industrie extractive requiert quant à elle la manipulation, le déplacement, et le stockage de minerais. Souvent organisé en piles, le stockage consomme de la place et peut occasionner d'importants problèmes de sécurité. Le transport de ces matériaux, par exemple en bande convoyeuse, est également source de pannes.

Pour les fonderies, lors de l'utilisation de lingotières par exemple, du métal fondu est injecté par le bas d'un moule pour permettre son remplissage. Une poudre est dispersée sur le dessus du liquide pour l'isoler thermiquement de l'air et éviter l'oxydation et les impuretés. Cette poudre est contenue dans des sacs suspendus au-dessus du liquide ou posés au fond du moule, puis dispersée lors de la combustion des sacs. Les problèmes d'étalement de la poudre sont à l'origine des problèmes de qualités finales des lingots [1].

Mieux appréhender la dynamique des matériaux granulaires menant à la formation de talus est donc une nécessité pour de nombreuses industries et nécessite de réaliser des prédictions, aujourd'hui exclusivement basées sur la modélisation numérique, et largement perfectibles.

L'empilement des matériaux granulaires échappe en effet largement à la modélisation classique, notamment du fait que les grains ou poudres se comportent à la fois comme des solides et des liquides [3], et que les interactions (collisionnelles) complexes entre les grains peuvent être assimilées à celles de gaz [1]. Ainsi, les interactions complexes entre les particules sont régies par de nombreux facteurs micro et macroscopiques, ce qui conduit certains scientifiques à affirmer que, en fonction de la résolution des appareils de mesure, un simple grain contient presque toutes les lois de la nature à des échelles de plus en plus petites [4].

Deux approches sont principalement utilisées pour la simulation des écoulements granulaires. La première, dite "discrète" et communément appelée DEM (Discrete Element Method), est la plus répandue. Elle a été introduite en 1971 par Cundall et consiste à décrire l'écoulement grain

par grain, en considérant des interactions complexes, des formes et des géométries complexes, le déplacement de chaque particule étant calculé en résolvant la seconde loi de Newton pour chaque particule [3].

Cette méthode a l'avantage de pouvoir tenir compte des propriétés microscopiques des matériaux granulaires, mais est limitée par un coût important lié à la caractérisation de chaque grain par un laboratoire et l'impossibilité de l'appliquer à l'échelle industrielle ou environnementale où le nombre de particules est trop important et où une simple considération microscopique ne permet pas de rendre compte efficacement du comportement macroscopique de l'écoulement granulaire.

La seconde méthode, dite « continue » ou appelée CMM, consiste à assimiler le matériau granulaire comme un corps continu soumis à des lois spécifiques et dont l'écoulement peut être résolu par les équations de la mécanique des fluides. Cette méthode implique des approximations et de la modélisation physique, mais, corrigée par des données réelles, peut donner de bons résultats. Elle exclut en revanche les considérations microscopiques (forme et géométrie des particules notamment) alors qu'il s'agit de paramètres importants, ce qui peut conduire à de mauvais résultats [5] remettant en cause son applicabilité générale [6], surtout dans le cadre de processus industriels où la sécurité et l'efficacité sont essentielles.

Les professionnels ne disposent donc aujourd'hui que d'outils de modélisation approximatifs dont l'applicabilité générale ne peut être démontrée et qui ne peuvent tenir compte de tous les facteurs internes ou externes à l'origine de la complexité des interactions régissant les écoulements granulaires.

Cette complexité même fait penser à deux outils spécialement adaptés à la modélisation des phénomènes complexes à propriétés émergentes : les automates cellulaires (et plus exactement les modèles de particules sur treillis) et les réseaux de neurones profonds.

1. Stratégie de recherche

1.1. Les datasets rares

Si certains processus industriels peuvent générer une masse importante de données temps réel, la densité d'information contenue peut être relativement faible en quantité, et relativement faible en corrélation avec le résultat final et les caractéristiques physiques du produit manufacturé. Les technologies de deep learning sont de grandes consommatrices de données et leur essor est partiellement la conséquence de la disponibilité massive de contenus, notamment multimédia. Le secteur de l'intelligence artificielle est confronté à plusieurs challenges stratégiques : le premier est de traiter les relations de causes à effets. Dans la modélisation classique, des données d'entrées (input) sont soumises à un modèle (algorithme) (Figure 3) qui calcule un résultat (output) (Figure 3). Dans l'approche de deep learning, l'apprentissage à partir des données et de résultats connus permet l'élaboration d'un modèle pour une architecture de réseau de neurones optimale (Figure 4).

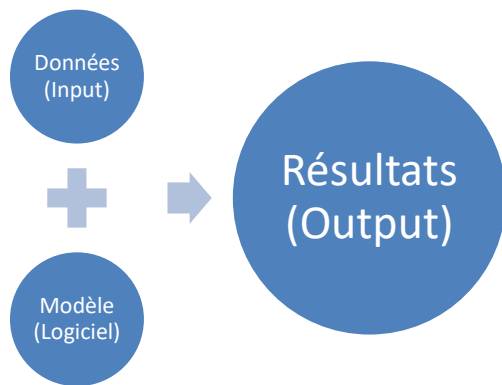


Figure 3: Processus classique de la modélisation

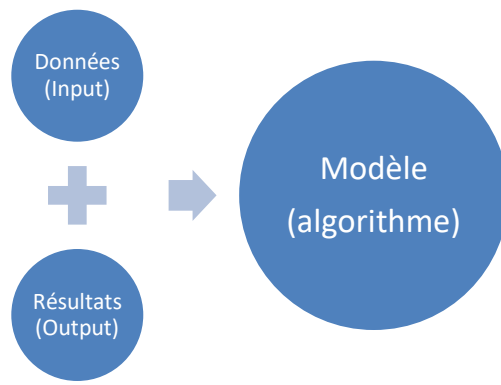


Figure 4: Processus deep learning

Cette différence fondamentale d'approche fait passer le processus d'investigation scientifique d'un état où l'on fait une hypothèse (de calcul) puis une confrontation à l'expérience, à un état où l'on cumule les expériences pour en déduire un modèle sous la forme d'un réseau de poids synaptiques et d'un couple preprocessing/postprocessing. Ce modèle de neurones artificiels a une structure très différente des modèles classiques. Ces derniers sont fait à partir d'équations (Navier-Stokes, Fourier, Schrödinger..) parfois complexes résolues par des approximations numériques raffinées, mais il est toutefois possible d'établir des relations entre des paramètres (causes, par exemple une augmentation de température) et des effets (par exemple une augmentation de pression). Les modèles Deep Learning sont parfois considérés comme des boîtes noires pour lesquelles les relations entre les causes et les effets sont des combinaisons tellement complexes qu'il n'est pas possible de les traduire en équation autrement qu'en appliquant les opérations tensorielles contenues dans le réseau de neurones. Un des enjeux à court terme est de pouvoir expliquer un résultat fourni par un réseau de neurones. Des algorithmes tels que APRIORI [7] peuvent aider à reconstruire ces règles, mais l'enjeu est beaucoup plus largement celui de la découverte d'associations. Dans le cas des réseaux convolutifs, il est parfois possible en observant les différentes couches de l'espace latent de reconnaître des schémas conduisant à l'identification d'une image [8]. Mais ces observations sont toutefois insuffisantes pour toutes les applications dans lesquelles l'apprentissage se fait sur des données non photographiques ainsi que lorsque la sécurité ou la fiabilité sont vitales mettant en jeu la vie humaine.

Le second enjeu est la réduction de la quantité de données nécessaires à un apprentissage de qualité.

La plupart des applications actuelles sont de fortes consommatrices de données. En deep reinforcement learning, les succès de AlphaGO sur le jeu de GO sont obtenus en faisant jouer des IA les unes contre les autres et s'auto enseigner les règles du jeu et les stratégies gagnantes. Cela n'est possible que dans la mesure où l'acquisition de données (ici le résultat d'un tour dans un jeu) ne consomme que peu de ressources. C'est pourquoi d'autres approches dites de transfert learning (apprentissage par transfert) se développent actuellement [9][10]. Dans le domaine médical notamment [11], les données peuvent également être rares quand elles sont le fruit de l'observation de cas plutôt que d'expérimentation. Le transfert learning consiste à préentraîner un réseau de neurones sur un corpus de données à faible coût d'acquisition, comme par exemple

celles issues d'un modèle analytique ou numérique puis à terminer l'entraînement sur des données plus coûteuses. Le préentraînement permet de disposer d'un modèle ayant une idée grossière mais cohérente du résultat pour réduire la quantité de données requises pour entraîner un modèle exact.

Il est possible de décomposer le transfer learning en deux process de recherche :

- Ab initio: il s'agit de mener de bout en bout la construction d'un modèle préentraîné à partir de données sorties d'une source à bas coût (comme par exemple un modèle numérique ou une forme semi-empirique)
- Reuse : il s'agit d'assembler des réseaux préentraînés par d'autres. Il en existe beaucoup dans le domaine de la classification d'images[12], de sons [13] ou dans le traitement du langage[14], mais beaucoup moins dans le domaine des procédés.

Les techniques de one-shot ou few-shot learning dont les applications en imagerie sont spécialement en développement aujourd'hui peuvent sans doute également contribuer au traitement de la rareté des données[15].

Nous pouvons résumer le processus de développement de la manière présentée en Figure 5. Plusieurs contraintes sont inhérentes à ce mode d'apprentissage : les modèles utilisés pour donner une « culture » au réseau de neurones pré-entraînés doivent être relativement fiables et rapides, et ne pas nécessiter de réglages de paramètres de convergence différents à chaque cas pour limiter l'intervention humaine.

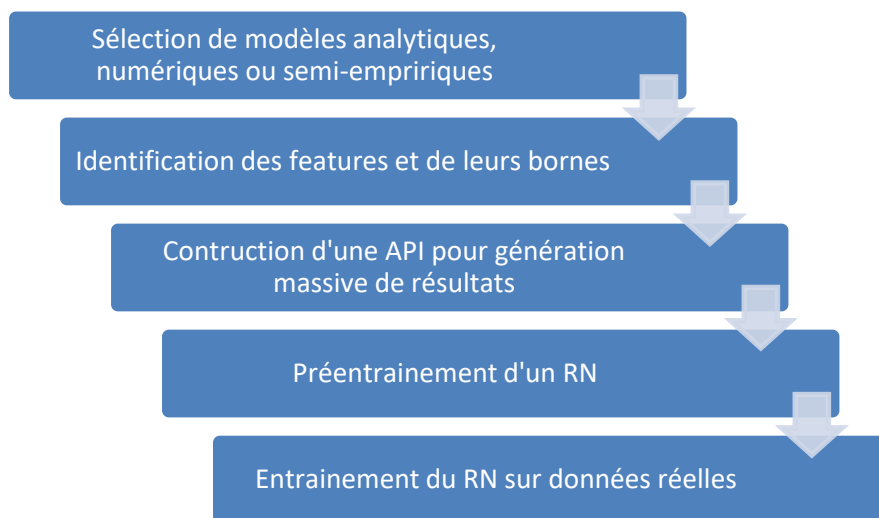


Figure 5: Process de transfer learning

Par ailleurs, la précision des prédictions du réseau de neurones pré-entraîné doit être excellente sur les données de validation issues des mêmes modèles.

Pour ce projet, deux éléments principaux sont à coupler : d'une part un automate cellulaire suffisamment simple pour être exécutable dans des temps de calcul actuellement acceptables et suffisamment complexe pour représenter un empilement réel, d'autre part un réseau de neurones profond capable d'apprendre de cet automate les relations complexes entre les événements qui ont abouti à la formation de l'empilement et l'angle de talus naturel finalement observé.

1.2. Particules on lattice

Utiliser des équations de physique pour pré-entraîner des réseaux de neurones revient à leur faire apprendre les hypothèses que sous-tendent ces équations dont certaines sont difficiles à admettre dans un contexte réel : continuité et linéarité des milieux, indépendance des variables, dérivabilité...

Dans les années 70, l'essor des calculateurs a permis l'émergence des automates cellulaires, dont le plus connu est le « jeu de la vie » d'abord développé par Ulam et Von Neuman pour une application à des cristaux dans les années quarante [16]. Le fait que des règles extrêmement simples conduisent à des comportements complexes et à des émergences a fasciné les premiers informaticiens.

Les automates cellulaires établissent des comportements individuels de particules réparties sur un treillis. Ces comportements dépendent des particules voisines et sont pris en compte à chaque itération. Assez rapidement, des structures stables, cycliques, chaotiques, fractales peuvent apparaître alors que le contenu informationnel des règles est assez faible. Ces comportements sont également observés dans la Nature. Si on désigne souvent par « hiver de l'IA » la période de moindre intérêt pour l'intelligence artificielle entre 1987 et 1993, les automates cellulaires ont eux aussi connu un « hiver », liés d'une part à leur besoin gigantesque en mémoire et à leur indépendance vis-à-vis de la variable temps remplacée par des « steps » qui les limitaient aux phénomènes non transitoires[17].

La convergence entre les technologies de deep learning et les automates cellulaires est aujourd'hui possible grâce aux accélérations permises par les GPU (Graphic Processing Units) et les unités de calcul tensoriel (TPU)[18] permettant même des simulations tridimensionnelles[19], voire une utilisation de ces automates, qui peuvent être des machines de Turing, pour constituer des éléments de deep learning[20]. Les automates cellulaires à mailles hexagonales ont rencontré des succès dans le domaine de la dynamique des fluides [21] et leur couplage avec les technologies de deep learning a été réalisé dans le domaine de la géophysique[22].

Dans cette publication, le modèle de particules sur treillis développé consiste en l'enchevêtrement de n réseaux correspondants à autant de classes granulaires de particules (Figure 6).

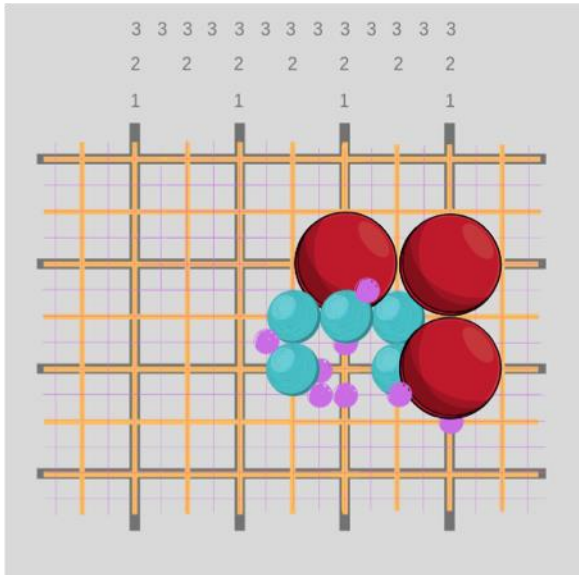


Figure 6: Insertion des particules sur les réseaux

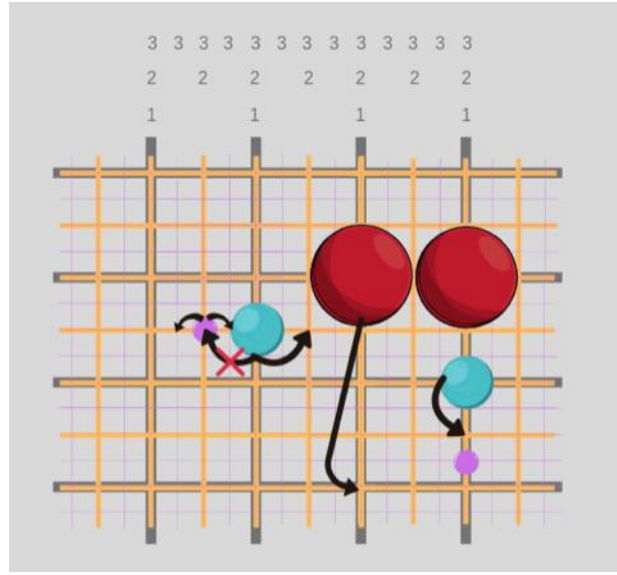


Figure 7: Représentation des règles de l'automate cellulaire

Les particules peuvent évoluer selon des règles simples, comme dans le cas des automates cellulaires de Conway.

Ces règles (Figure 7) sont :

- Les particules sont introduites dans le treillis selon une distribution de n classes déterminées à l'avance
- Une particule de classe n ne peut se situer que sur un nœud du réseau de classe n
- Une particule ayant un site vide en dessous le rejoint avec une probabilité de 1
- Une particule empêchée de descendre se déplace latéralement avec une probabilité p . Cette probabilité est fixée expérimentalement de manière à confiner l'empilement granulaire dans les limites du treillis.

Le dataset a été créé par une implémentation openAI GYM [23] dans l'optique future d'une utilisation en deep reinforcement learning pour le pilotage des coupes granulométriques en temps réel[24]. Il a été implémenté en calcul parallèle sur une instance multiprocesseur et un GPU GTX 1070Ti.

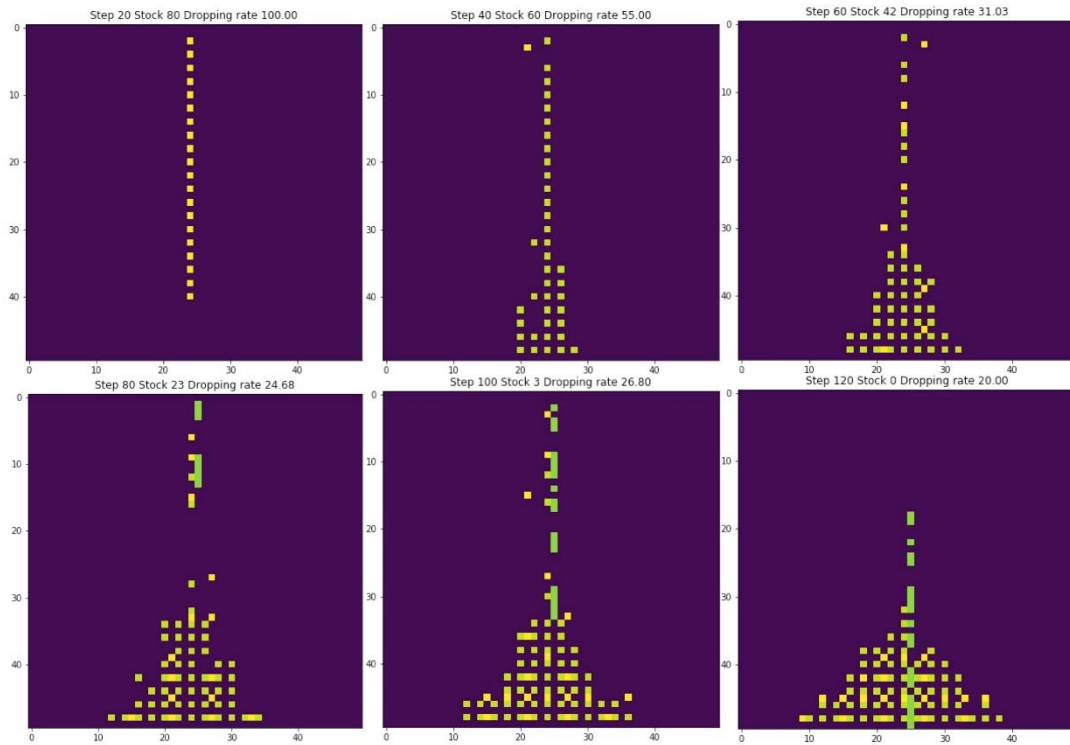


Figure 8: Étape de la constitution d'un empilement

Le modèle s'arrête quand l'ensemble atteint une fraction de déplacement inférieure à une valeur déterminée à l'avance. L'angle est alors calculé (Figure 8).

Les coupes granulométriques sont représentées comme des séquences de leur fraction dans le mélange, soit $[0.1, 0.8, 0.1]$ pour une séquence à trois classes.

Il est à noter que la stratégie d'implantation des treillis dans une instance openAI Gym permet aussi d'utiliser les résultats obtenus pour piloter par deep reinforcement learning un empilement obtenu par mariage de granulométrie ce qui présente un grand intérêt industriel.

1.3. Réseau de neurones

Le dataset est constitué de cas d'entraînement correspondant à une séquence de taille de grains croissante. Ces séquences peuvent être représentées par une couche Long Short Term Memory (LSTM)[25][26] de Tensorflow[27]. Nous faisons ici l'hypothèse que des séquences de granulométrie à l'intérieur de la distribution globale peuvent influencer l'angle de talus naturel. Proposés en 1997 par Hochreiter et Schmidhuber, les réseaux de neurones LSTM sont des réseaux de neurones récurrents qui permettent de prendre en compte des rétroactions entre les neurones. Particulièrement adaptés à la représentation de séquences, ils sont notamment une technologie clef du traitement du langage naturel. L'architecture classique est constituée d'une cellule et de trois portes (regulators) : une porte d'entrée, une porte de sortie et une porte d'oubli (Figure 9).

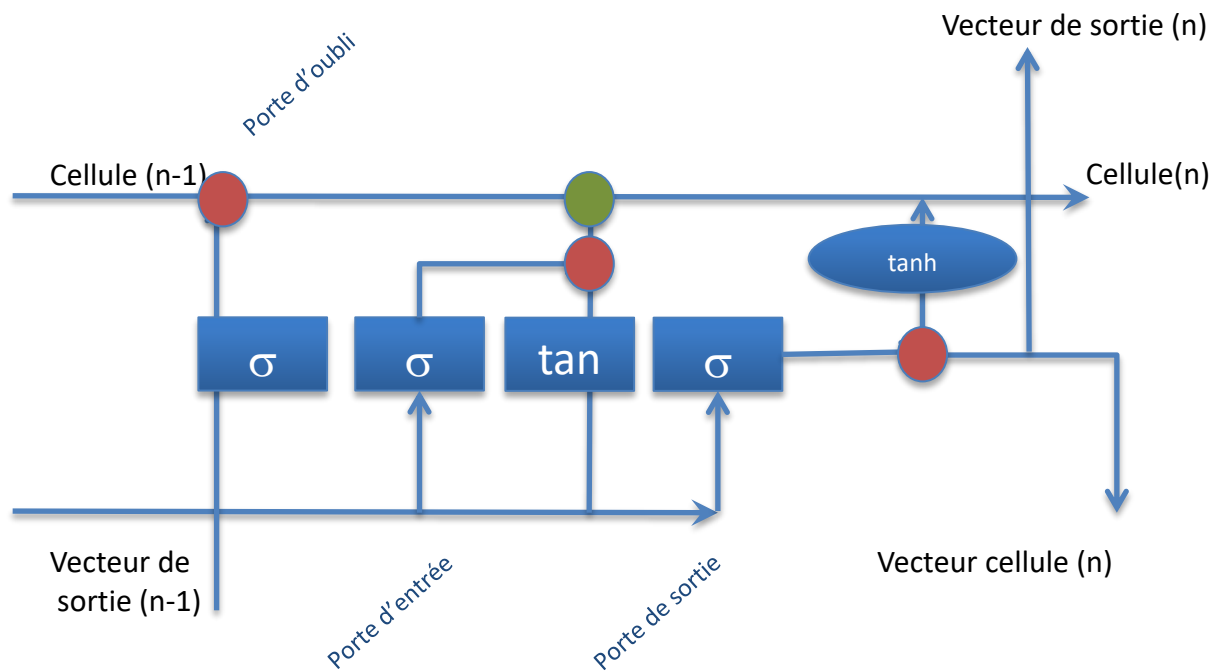


Figure 9: Représentation schématique d'une cellule LSTM (σ représente une activation sigmoïde, \tanh une activation hyperbolique, cercles rouges un produit matriciel d'Hadamard, cercle vert une addition)

Les réseaux de neurones LSTM montrent une bonne efficacité pour des séquences assez longues, mais ne conservent pas une mémoire à très long terme. Ils conviennent toutefois à la représentation de séquences de granulométrie.

2. Résultats

Deux configurations sont ici présentées : l'une a été élaborée sur un lattice de 4000x4000 avec une granulométrie resserrée, l'autre avec une grille de 1000x1000, une granulométrie plus large et un nombre d'instances plus élevé.

Les calculs sont réalisés sur une instance virtuelle de 16 cœurs CPU et 1920 cœurs GPU. Dans notre cas, les temps de calculs cumulés entre l'élaboration des datasets et l'apprentissage du réseau de neurones vont de quelques heures à 48h pour des treillis de moins de 10000x10000 mailles.

2.1. Modèle à 8 tailles de grains

L'automate cellulaire a été utilisé pour générer un dataset de 1600 expériences pour 8 classes de particules dans un treillis de 4000x4000 mailles et 4000 particules introduites selon la distribution des classes.

Les granulométries étant une séquence, une architecture de réseau de neurones LSTM a été choisie (Tableau 1) après plusieurs essais pour obtenir une bonne qualité d'apprentissage (dans cette version il n'y a pas de dropout). L'entraînement du réseau a été fait sur 1391 expériences et la validation sur 199.

Loss : Mean squared error	Input	• 8 neurones en séquence de granulométrie
Initializer : Normal	Hidden	• 32 neurones en type LSTM
Activation : Linear	Hidden	• 16 neurones en type Dense
Optimizer : Adam	Output	• 1 neurone (activation linéaire)
Epochs : 2000		
Batch size : 5		

Tableau 1: Principaux paramètres du modèle de deep learning

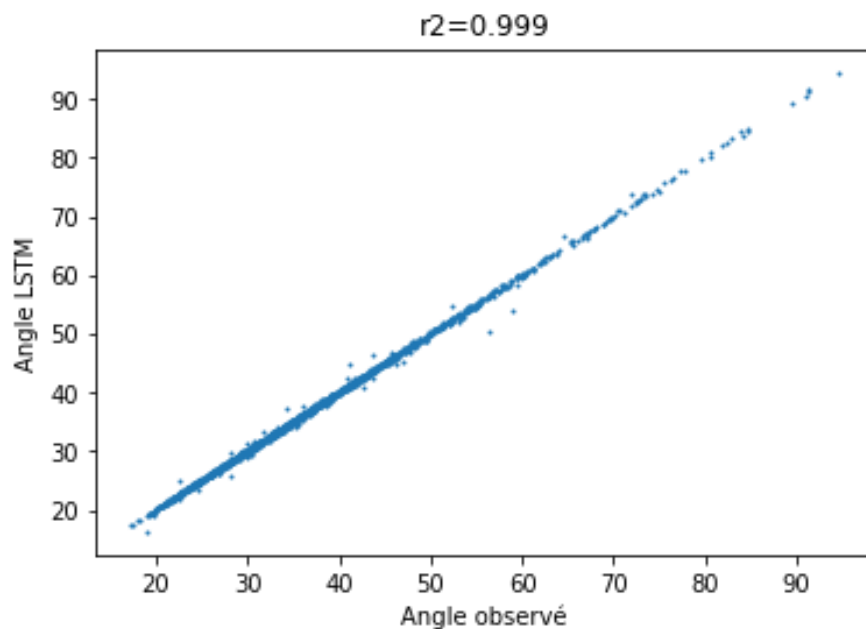


Figure 10: Prédiction des angles dans un modèle à 8 tailles

La prédiction des angles sur le dataset d'entraînement montre une bonne corrélation ($R^2=0.999$) de même que pour les données de validation non utilisées pour l'apprentissage ($R^2=0.985$) (Figure 10).

2.2. Modèle à 16 tailles de grains

Un modèle de plus petite taille (1000x1000) et 2000 grains de 16 tailles a été implémenté pour un temps de calcul complet de 45h.

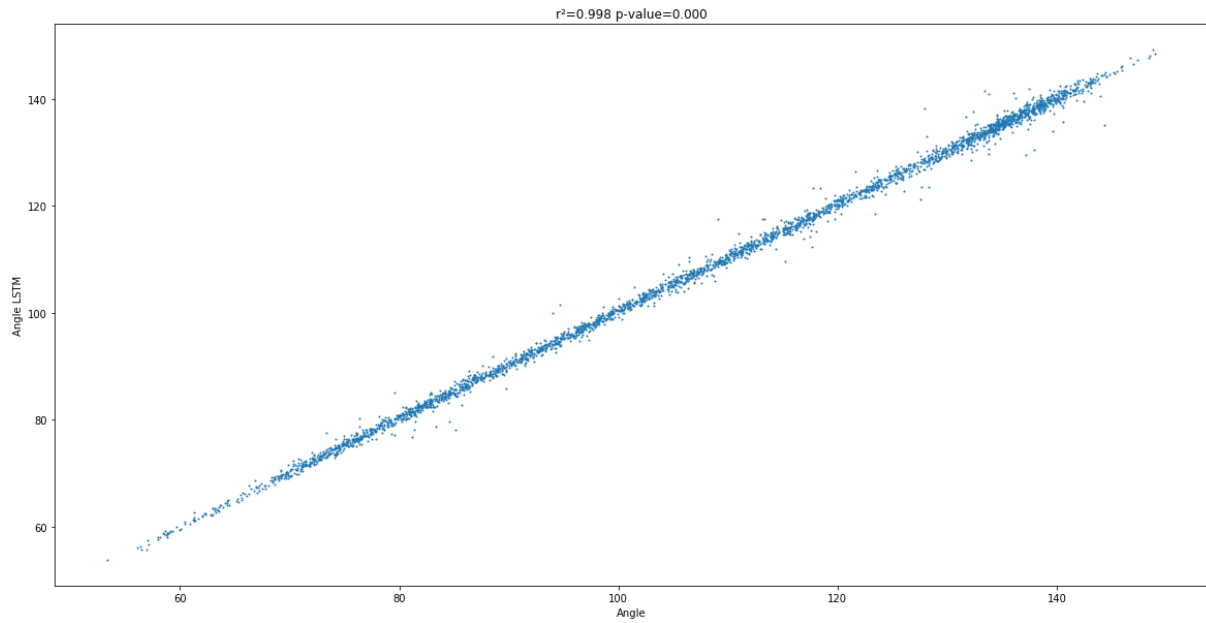


Figure 11: Prédiction des angles dans un modèle à 16 tailles

Le modèle montre une bonne corrélation pour les données d'entraînement. Sur des données non connues du modèle, la précision demeure convenable (Figure 12) avec des coefficients de corrélation de plus de 98%.

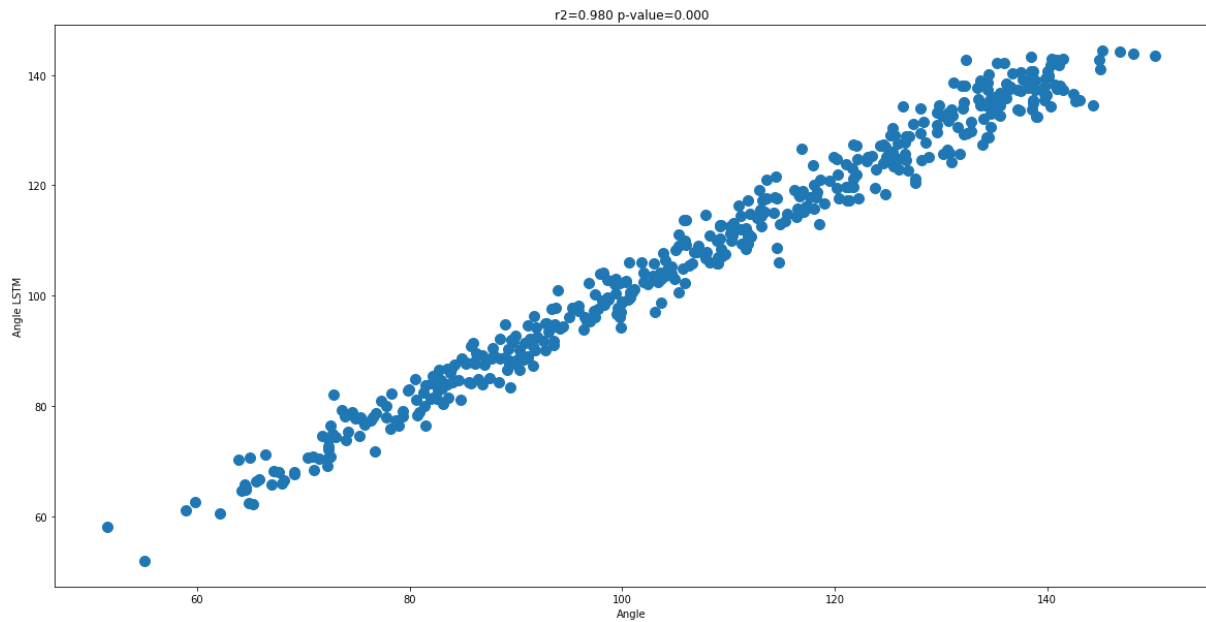


Figure 12: Prédiction des angles sur un échantillon de validation

Le nombre de neurones des couches cachées influence la précision du modèle (Tableau 2).

Input Layer	Hidden layers	R ² (training)	R ² (données nouvelles)
16	32,16	0,998	0,980
16	8,8	0,988	0,986
16	4,4	0,964	0,968

Tableau 2: Influence des paramètres d'apprentissage sur le coefficient de corrélation R2

Il est à noter que les résultats sont toujours meilleurs que la simple régression multilinéaire qui présente une distribution plus complexe et fourchée pour les angles faibles (Figure 13).

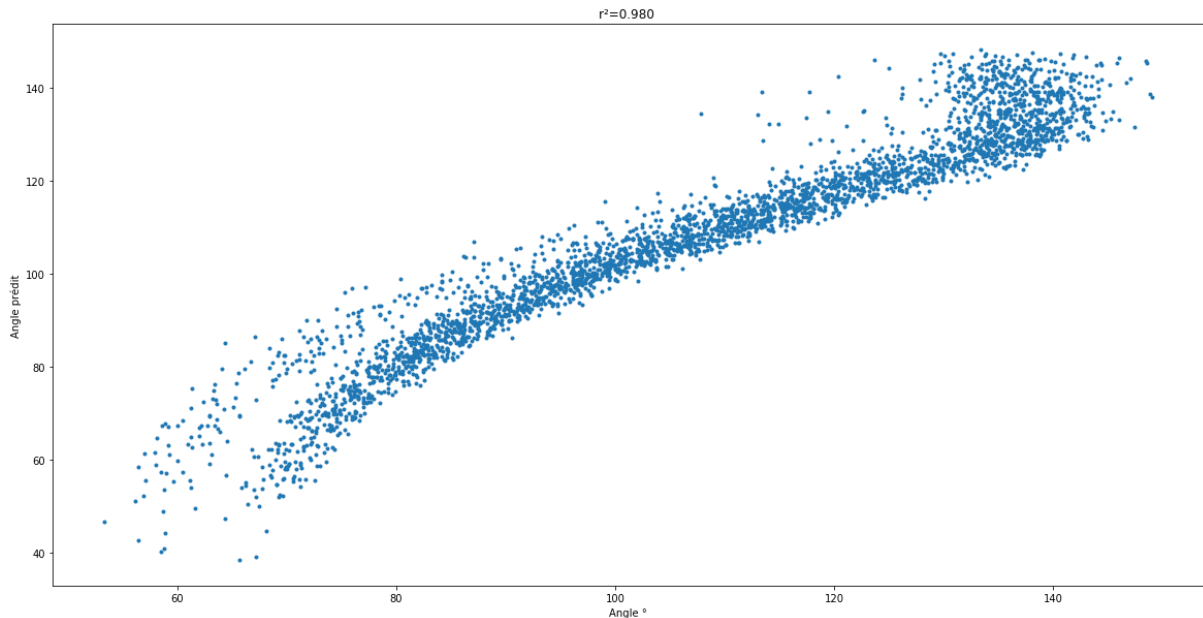


Figure 13: Comparaison avec une régression multilinéaire

Conclusion-Perspectives ultérieures

La technologie neurolattice ici présentée doit faire l'objet d'évaluations complémentaires dans deux contextes. D'une part l'implémentation du neurolattice dans une classe GYM permet d'envisager le pilotage de granulométrie mixte en temps réel pour obtenir des propriétés de poudres correspondant à des contraintes industrielles données, notamment en utilisant du deep reinforcement learning soit à partir des modèles de particules sur réseaux, soit à partir des réseaux LSTM préentraînés; d'autre part en réentraînant les réseaux LSTM préentraînés sur les modèles de particules sur réseaux avec des données réelles dans une procédure de transfer learning.

Le réseau LSTM a donné des résultats intéressants mais pour une distribution de taille de grains relativement modeste. L'utilisation de cellules GRU (Gated Recurrent Units) pourrait donner, en raison de leur plus grande simplicité, des performances meilleures notamment pour les datasets de petite taille, issus d'observations réelles notamment[28].

L'automate cellulaire d'apprentissage utilisé ici est relativement simple et pourrait être amélioré, et en premier lieu en passant à une représentation tridimensionnelle. Les approches DEM pourront également permettre de réaliser un réseau de neurones préentraîné plus proche dans ses prédictions de la réalité physique.

Remerciements

Ces travaux ont été soutenus par EURENCO.

Références

- [1] S. Riber, "Numerical methods for the simulation of continuum granular flow models," Université de recherche Paris Sciences et Lettres, Thèse, 2017.
- [2] J. Cante *et al.*, "PFEM-based modeling of industrial granular flows", *Comput. Part. Mech.*, vol. 1, no. 1, pp. 47–70, May 2014.
Doi: 10.1007/s40571-014-0004-9
- [3] C. J. Coetzee, "The modelling of granular flow using the particle-in-cell method," Stellenbosch : University of Stellenbosch, Thesis, 2004.
- [4] F. Radjai, J.-N. Roux, and A. Daouadji, "Modeling Granular Materials: Century-Long Research across Scales", *J. Eng. Mech.*, vol. 143, no. 4, p. 04017002, Apr. 2017.
Doi: 10.1061/(ASCE)EM.1943-7889.0001196

- [5] R. Gao, X. Du, Y. Zeng, Y. Li, and J. Yan, "A new method to simulate irregular particles by discrete element method", *J. Rock Mech. Geotech. Eng.*, vol. 4, no. 3, pp. 276–281, Sep. 2012.
Doi: 10.3724/SP.J.1235.2012.00276
- [6] A. Mangeney, L. Staron, D. Volfson, and L. Tsimring, "Comparison Between Discrete and Continuum Modeling of Granular Spreading," in *Proceedings of the 10th WSEAS International Conference on Applied Mathematics*, Stevens Point, Wisconsin, USA, pp. 63–70, 2006.
- [7] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. of 20th Intl. Conf. on VLDB*, 1994, pp. 487–499.
- [8] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps," *ArXiv13126034 Cs*, Dec. 2013.
- [9] S. J. Pan and Q. Yang, "A Survey on Transfer Learning" *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
Doi: 10.1109/TKDE.2009.191
- [10] J. Ngiam, D. Peng, V. Vasudevan, S. Kornblith, Q. V. Le, and R. Pang, "Domain Adaptive Transfer Learning with Specialist Models", *ArXiv:1811.07056*, Nov. 2018.
- [11] M. Christopher *et al.*, "Performance of Deep Learning Architectures and Transfer Learning for Detecting Glaucomatous Optic Neuropathy in Fundus Photographs," *Sci. Rep.*, vol. 8, no. 1, p. 16685, Nov. 2018.
Doi: 10.1038/s41598-018-35044-9
- [12] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, 2009.
Doi: 10.1109/CVPR.2009.5206848
- [13] L. Deng, M. L. Seltzer, D. Yu, A. Acero, A. Mohamed, and G. Hinton, "Binary Coding of Speech Spectrograms Using a Deep Auto-Encoder", p. 4.
- [14] R. Al-Rfou, B. Perozzi, and S. Skiena, "Polyglot: Distributed Word Representations for Multilingual NLP," in *Proceedings of the Seventeenth Conference on Computational Natural Language Learning*, Sofia, Bulgaria, 2013, pp. 183–192. arXiv:1307.1662
- [15] O. Vinyals, C. Blundell, T. Lillicrap, koray kavukcuoglu, and D. Wierstra, "Matching Networks for One Shot Learning," in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds. Curran Associates, Inc., 2016, pp. 3630–3638.
- [16] J. Von Neumann and A.W. Burks, "Theory of self-reproducing automata." Urbana, University of Illinois Press, 1966.
- [17] H. Gutowitz, "Cellular Automata: Theory and Experiment." MIT Press, 1991.

- [18] S. Szkoda, Z. Koza, and M. Tykierko, “Accelerating cellular automata simulations using AVX and CUDA” *ArXiv12082428 Phys.*, Aug. 2012.
- [19] S. Chen, “Three dimensional Cellular Automaton – Finite Element (CAFE) modeling for the grain structures development in Gas Tungsten / Metal Arc Welding processes”, École nationale supérieure des mines de Paris, Thèse, 2014.
- [20] W. Gilpin, “Cellular automata as convolutional neural networks”, *ArXiv180902942*, Sep. 2018.
- [21] Dieter A. Wolf-Gladro, “Lattice-Gas Cellular Automata and Lattice Boltzmann Models - An Introduction”, Springer, 2005
<https://www.springer.com/us/book/9783540669739>
- [22] Xiaodong Song, Ganlin Zhang, “Modeling spatio-temporal distribution of soil moisture by deep learning-based cellular automata model”, *Journal of Arid Land*, October 2016, Volume 8, Issue 5, pp 734–748
Doi: 10.1007/s40333-016-0049-0
- [23] J. Francis, “Introduction to reinforcement learning and OpenAI Gym”, *O’Reilly Media*, 13-Jul-2017. <https://www.oreilly.com/learning/introduction-to-reinforcement-learning-and-openai-gym>
- [24] K. Arulkumaran, “A Brief Survey of Deep Reinforcement Learning” , *ArXiv:1708.05866v2*, 2017.
- [25] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory”, *Neural Computation*, MIT Press Journals, Vol 9,N°8, pp1735-1780, 1997.
Doi: 10.1162/neco.1997.9.8.1735
- [26] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [27] M. Abadi *et al.*, “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems”, *ArXiv160304467 Cs*, Mar. 2016.
- [28] M. Wielgosz, A. Skoczen, and M. Mertik, “Recurrent Neural Networks for anomaly detection in the Post-Mortem time series of LHC superconducting magnets”, *ArXiv:1611.06241*, p. 28, 2017.