

Présentation de stage

Détection d'arcs électriques par des méthodes d'intelligence artificielle

Céline Lorentz

Tuteurs : Patrick Schweitzer, Christophe Bonnet
M2 CSMI - Université de Strasbourg
Institut Jean Lamour - Nancy

26 Août 2022

Plan de la présentation

- 1 Introduction
- 2 Techniques expérimentales de génération d'arcs
- 3 Création des bases de données
- 4 Modèles de deep learning pour la détection d'arcs
- 5 Comparaison des performances entre algorithmes
- 6 Vérification de la pertinence de la base de données retenue
- 7 Conclusion

Plan de la présentation

- 1 Introduction
- 2 Techniques expérimentales de génération d'arcs
- 3 Création des bases de données
- 4 Modèles de deep learning pour la détection d'arcs
- 5 Comparaison des performances entre algorithmes
- 6 Vérification de la pertinence de la base de données retenue
- 7 Conclusion

Présentation du laboratoire



Institut Jean Lamour

- Laboratoire de recherche fondamentale et appliquée en science des matériaux
- Unité mixte du CNRS et de l'Université de Lorraine
- Matériaux, métallurgie, nanosciences, plasmas, surfaces, électronique

Présentation du laboratoire



- Nancy (campus Artem)
- 4 départements scientifiques /
25 équipes / 8 centres de
compétences
- Classé ZRR (Zone à Régime
Restrictif)

Présentation de l'équipe à laquelle je suis rattachée

Équipe Mesures et Architectures Électroniques (MAE)

- Dirigée par M. Schweitzer (l'un de mes deux tuteurs avec M. Bonnet)
- Rattachée au département N2EV

Deux axes de recherches complémentaires

- Capteurs, instrumentation et mesure
- Architectures de circuits et systèmes électroniques

Présentation du sujet



Qu'est-ce qu'un arc électrique ?

- Arc électrique = phénomène physique qui se produit lorsqu'un milieu généralement isolant est ionisé et laisse passer le courant
- Visible à l'oeil nu

Road map

- Étude bibliographique
- Développer des algorithmes d'intelligence artificielle qui permettent de détecter les défauts d'arc électrique dans les circuits continus (domaine aéronautique)
- Comparer les performances entre les différents algorithmes
- Aller en laboratoire faire de nouvelles acquisitions et créer une nouvelle base de données

État de l'art et travaux précédents

- Création d'une base de données

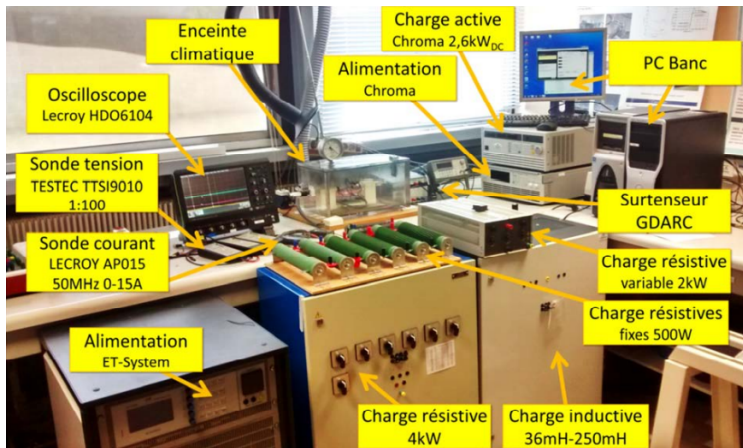
⇒ Nombreux problèmes de labellisation et labellisation à la main ⇒ trop long

- Si on veut rajouter un nouveau signal dans la base de données, il faut tout recommencer
- Création d'un premier code de classification

Plan de la présentation

- 1 Introduction
- 2 Techniques expérimentales de génération d'arcs
- 3 Création des bases de données
- 4 Modèles de deep learning pour la détection d'arcs
- 5 Comparaison des performances entre algorithmes
- 6 Vérification de la pertinence de la base de données retenue
- 7 Conclusion

Le banc expérimental



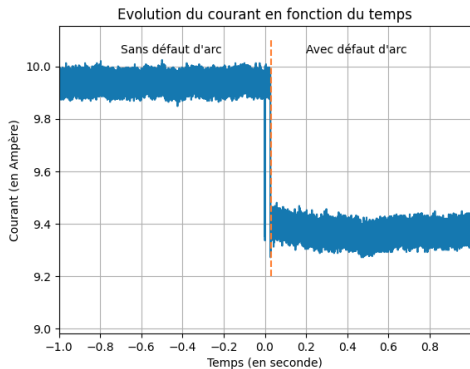
- Tension de ligne

- **Courant de ligne**

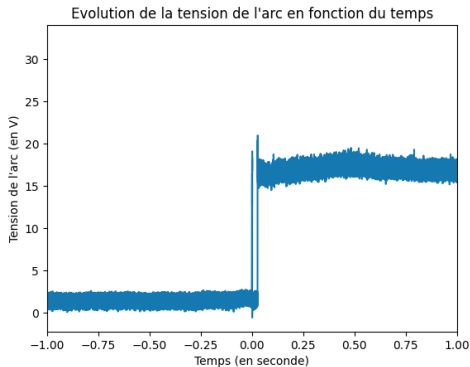
⇒ détection

- Tension d'arc

Affichage du signal : courant et tension



- Arc débute au temps $t = 0 \Rightarrow$ chute de courant



- À l'inverse, la tension augmente
- Bruit différent avant/après arc

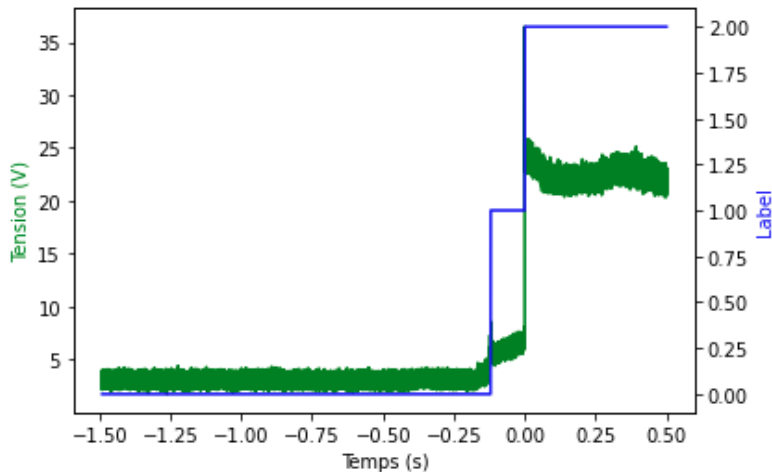
Plan de la présentation

- 1 Introduction
- 2 Techniques expérimentales de génération d'arcs
- 3 Création des bases de données**
- 4 Modèles de deep learning pour la détection d'arcs
- 5 Comparaison des performances entre algorithmes
- 6 Vérification de la pertinence de la base de données retenue
- 7 Conclusion

Labellisation des données

- À l'origine : 250 signaux et 40 leurres
- Lissage des signaux (à cause du bruit, des étincelles)
⇒ Permet de faire une labellisation automatisée
- Labellisation point par point

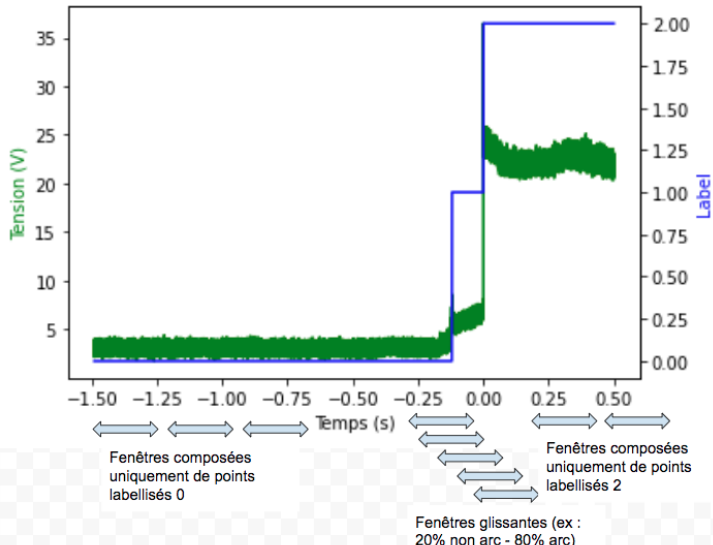
Exemple de labellisation point par point



Label	Signification
0	Pas d'arc
1	Début d'arc
2	Arc

- 1 fichier par signal

Sélection de segments



- Différentes labellisations

- Cas 1 : 3 classes (non arc et transition et arc)
- Cas 2 : 2 classes (non arc et transition)
- Cas 3 : 2 classes (non arc et transition+arc)
- Cas 4 : 2 classes (non arc et arc)

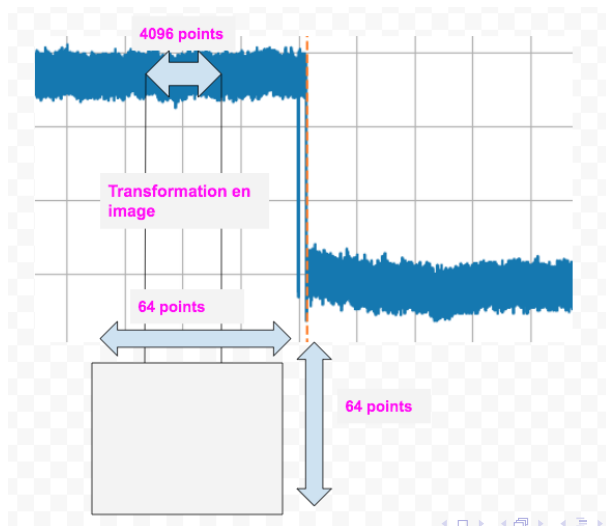
- Différentes tailles de segments

- De taille 28*28 (M28)
- De taille 64*64 (M64)
- De taille 128*128 (M128)

Présentation des bases de données créées

- À la fin : entre 12 000 et 19 000 segments par base de données
- Données sous forme de matrices 2D issues de segments du courant numérisé

Transformation en image

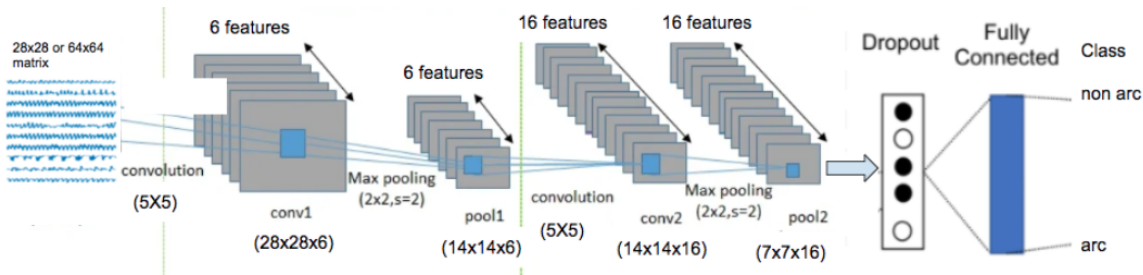


Plan de la présentation

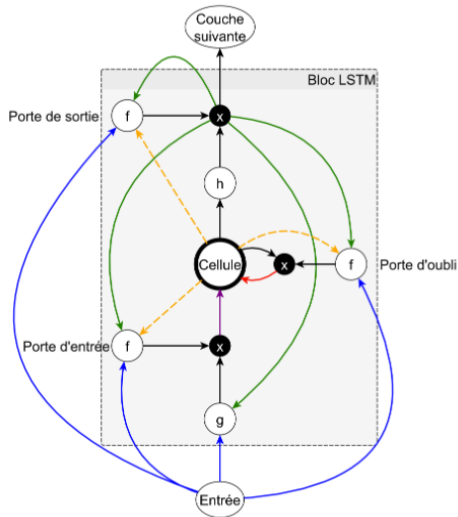
- 1 Introduction
- 2 Techniques expérimentales de génération d'arcs
- 3 Création des bases de données
- 4 Modèles de deep learning pour la détection d'arcs**
- 5 Comparaison des performances entre algorithmes
- 6 Vérification de la pertinence de la base de données retenue
- 7 Conclusion

Réseau de neurones convolutifs (CNN)

- Très puissants pour analyser des signaux (images, sons, vidéos)
- Réseau LeNet5(1998) = implantable dans un système électronique
- Convolution = application d'un filtre à une entrée



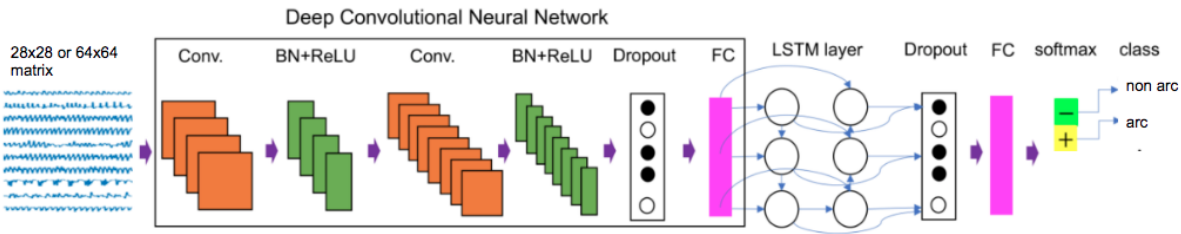
Réseau Long Short Term Memory (LSTM)



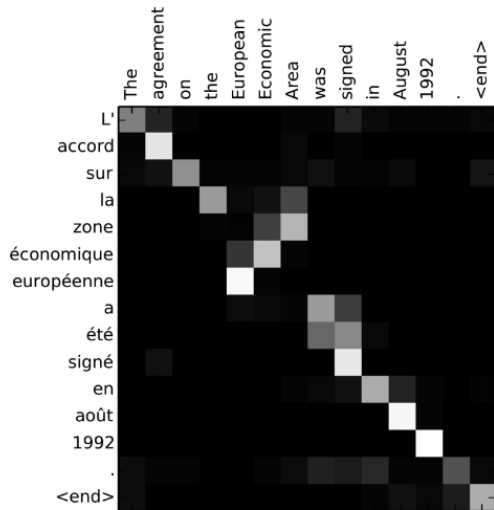
- LSTM = notion de mémoire
- Mémoire interne appelée cellule (valeur numérique)
- 3 portes (entrée, oubli, sortie)

Réseau CNN+LSTM

- Redimensionnement des layers nécessaire



Le principe de self-attention



- Mécanismes d'attention = traduction, langage
- Matrice d'attention = lien entre éléments
- Plus une cellule est claire et plus les éléments sont liés
- Auto-attention = interdépendance au sein d'une même séquence
- Transformer : blocs d'encodeurs et de décodeurs

Plan de la présentation

- 1 Introduction
- 2 Techniques expérimentales de génération d'arcs
- 3 Création des bases de données
- 4 Modèles de deep learning pour la détection d'arcs
- 5 Comparaison des performances entre algorithmes**
- 6 Vérification de la pertinence de la base de données retenue
- 7 Conclusion

Métriques de score

- Matrice de confusion :

	$\hat{-}$	$\hat{+}$
$-$	<i>TN</i>	<i>FP</i>
$+$	<i>FN</i>	<i>TP</i>

⇒ Fiabilité générale

- Précision = taux de prédictions correctes parmi les prédictions positives
- Rappel = taux d'individus positifs détectés par le modèle

Métriques de score

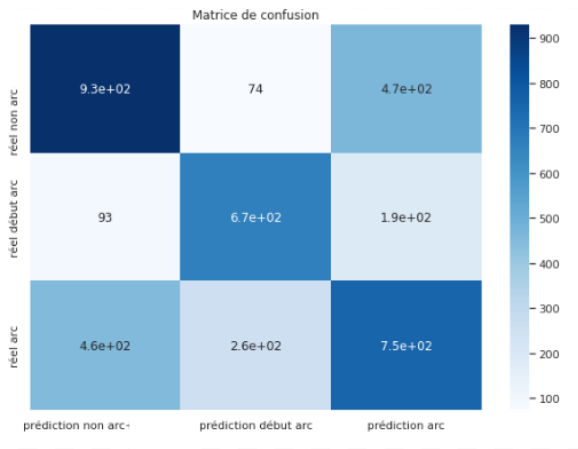
- Score F1 = moyenne harmonique de la précision et du rappel, défini par :

$$F_1 = \frac{2}{\frac{1}{\text{precision}} + \frac{1}{\text{rappel}}}$$

- Courbe ROC (Receiver Operating Characteristic) = taux de vrais positifs en fonction du taux de faux positifs

⇒ Aire sous la courbe = doit être la plus grande possible

Matrice de confusion pour le cas 1



⇒ Trop d'erreurs : la classification multi-classes n'est pas encore au point, on va se concentrer sur la classification binaire

Score F1

	C2	C3	C4
CNN	0.98	0.83	0.70
LSTM	0.94	0.76	0.67
CNN+LSTM	0.98	0.80	0.66
Attention	0.97	0.83	0.64
	M64	M64	M64

Table: Comparaison du score F1 pour nos différents algorithmes et nos différentes bases de données, $f_s=1\text{MHz}$

Sous-échantillonnage

- On garde le cas 2 (non arc et transition)
- Oscilloscope : $f_s=1\text{MHz}$
- Réduction de la fréquence d'échantillonnage = réduction du coût du projet
- On passe à $f_s=100\text{kHz}$ et on va voir si en enlevant autant d'info, cela marche toujours

Sous-échantillonnage

	C2	C2	C2
CNN	0.99	1.00	1.00
LSTM	0.98	1.00	1.00
CNN+LSTM	0.99	1.00	1.00
Attention	0.98	0.98	0.98
	M28	M64	M128

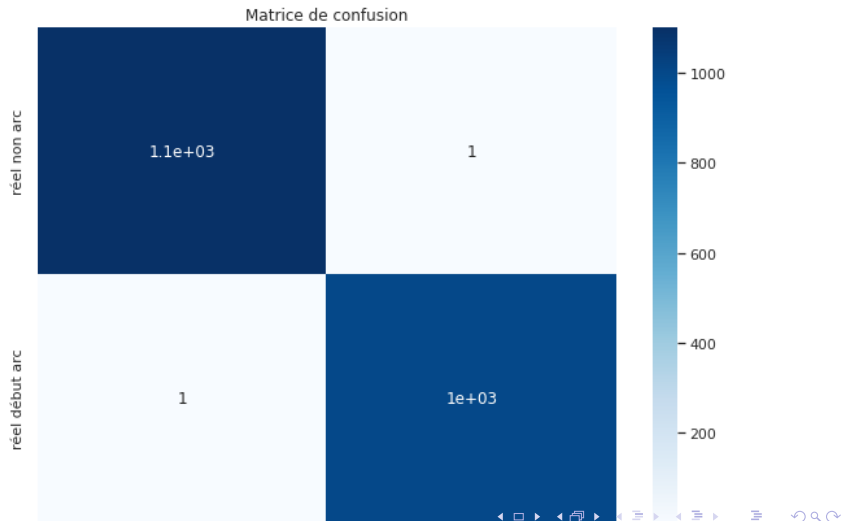
Table: Comparaison du score F1 pour nos différents algorithmes et nos différentes bases de données, $f_s=100\text{kHz}$

⇒ Cas 2, algorithme CNN+LSTM, segments $64*64$, $f_s=100\text{kHz}$

Plan de la présentation

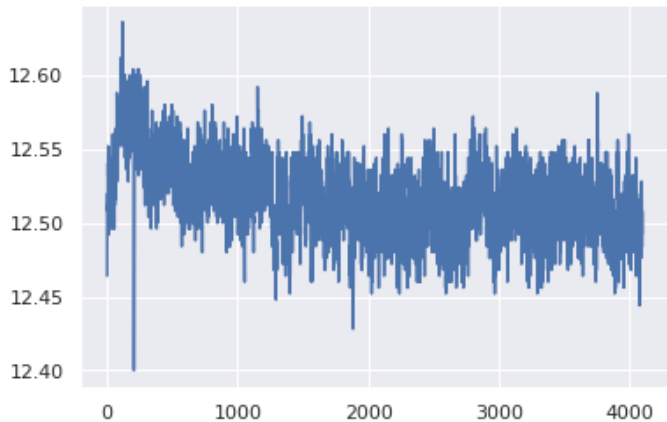
- 1 Introduction
- 2 Techniques expérimentales de génération d'arcs
- 3 Création des bases de données
- 4 Modèles de deep learning pour la détection d'arcs
- 5 Comparaison des performances entre algorithmes
- 6 Vérification de la pertinence de la base de données retenue**
- 7 Conclusion

Matrice de confusion

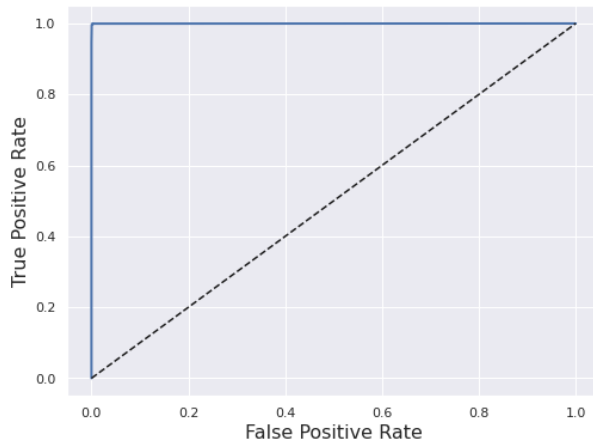


Erreur de l'algorithme

Début d'arc mais l'algo a prédit qu'il n'y avait pas d'arc 1590



Courbe ROC



⇒ Aire sous la courbe = 0.9997

Plan de la présentation

- 1 Introduction
- 2 Techniques expérimentales de génération d'arcs
- 3 Création des bases de données
- 4 Modèles de deep learning pour la détection d'arcs
- 5 Comparaison des performances entre algorithmes
- 6 Vérification de la pertinence de la base de données retenue
- 7 Conclusion**





Conclusion

- Mettre en oeuvre les connaissances apprises durant le master \Rightarrow stage très en accord avec ma formation
- Différents aspects d'un projet d'IA
- Création de la base de données du début à la fin (très formateur)
- Tests sur différentes bases de données \Rightarrow les algorithmes fonctionnent très bien mais pas sur toutes les bases de données
- **Meilleurs résultats** = base de données non arc-transition, sur des segments 64*64, fs=100kHz testée sur CNN+LSTM (10 545 segments)





Perspectives

- Faire des tests sur des signaux plus complexes
- Modifier les modèles pour qu'ils s'adaptent mieux à la classification multi-classes
- Tester d'autres modèles d'IA
- Implanter nos algorithmes dans des systèmes électroniques


Bibliographie I

-  Y. Liu, Y.-X. Huang, X. Zhang, W. Qi, J. Guo, Y. Hu, L. Zhang, and H. Su, "Deep c-lstm neural network for epileptic seizure and tumor detection using high-dimension eeg signals," *IEEE Access*, vol. 8, pp. 37495–37504, 2020.
-  T. Liu, J. Bao, J. Wang, and Y. Zhang, "A hybrid cnn-lstm algorithm for online defect recognition of co2 welding," *Sensors*, vol. 18, no. 12, 2018.
-  M. Z. Islam, M. M. Islam, and A. Asraf, "A combined deep cnn-lstm network for the detection of novel coronavirus (covid-19) using x-ray images," *Informatics in Medicine Unlocked*, vol. 20, p. 100412, 2020.
-  J.-B. Humbert, *Étude et détection des défauts d'arcs électriques dans un réseau électrique aéronautique 270V HVDC*.
Theses, Université de Lorraine, June 2018.

Bibliographie II

-  R. Herault and C. Chatelain, “Découvrez les cellules à mémoire interne : les LSTM.” <https://openclassrooms.com/fr/courses/5801891-initiez-vous-au-deep-learning/5814656-decouvrez-les-cellules-a-memoire-interne-les-lstm>, 2021.
-  A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” 2017.
-  S. V. Thiruloga, V. K. Kukkala, and S. Pasricha, “Tenet: Temporal cnn with attention for anomaly detection in automotive cyber-physical systems,” 2021.
-  J.-B. Cordonnier, A. Loukas, and M. Jaggi, “On the relationship between self-attention and convolutional layers,” 2019.

Bibliographie III

-  D. Sculley, G. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, M. Young, J.-F. Crespo, and D. Dennison, “Hidden technical debt in machine learning systems,” in *Advances in Neural Information Processing Systems* (C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, eds.), vol. 28, Curran Associates, Inc., 2015.