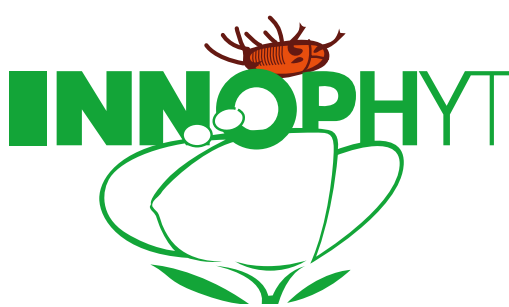


Approches connexionnistes pour la reconnaissance d’images d’insectes

Maxime Martineau, Romain Raveaux, Donatello Conte, Gilles Venturini

LI EA 6300, Université François Rabelais de Tours

maxime.martineau@univ-tours.fr



Introduction

Du point de vue de la biologie, l’identification de spécimens arthropodes (appelés communément insectes) est une source d’information riche sur la biodiversité des milieux naturels. Le pont entre l’informatique et la biologie est fait dès lors que les constats suivants sont posés : bien que des efforts aient été faits pour la rendre accessible, l’identification d’un spécimen demande une forte connaissance métier. La force de travail est de ce fait limitée. Inversement, la masse d’individus à identifier est importante. Le besoin d’automatisation est donc de plus en plus important. Une façon simple et intuitive d’automatiser cette tâche est de capturer une image de l’individu que l’on cherche à identifier.

Problème

On peut ainsi définir notre problème qui est la reconnaissance d’insectes à base d’image. Celui-ci s’inscrit donc dans le cadre de la reconnaissance d’objets à base d’images et plus largement de la reconnaissance d’images. Les particularités suivantes ont été dégagées. La nature des individus génère une forte variabilité du fait des mutations génétiques ainsi que des facteurs tels que l’âge. La capture de ces mêmes individus engendre également une variabilité importante : le fond peut être indéterminé si les images sont prises dans la nature (*field-based*) ou uniforme comme dans le cadre d’images basées laboratoire (*lab-based*). La figure 1 donne des exemples d’images utilisées dans la littérature.

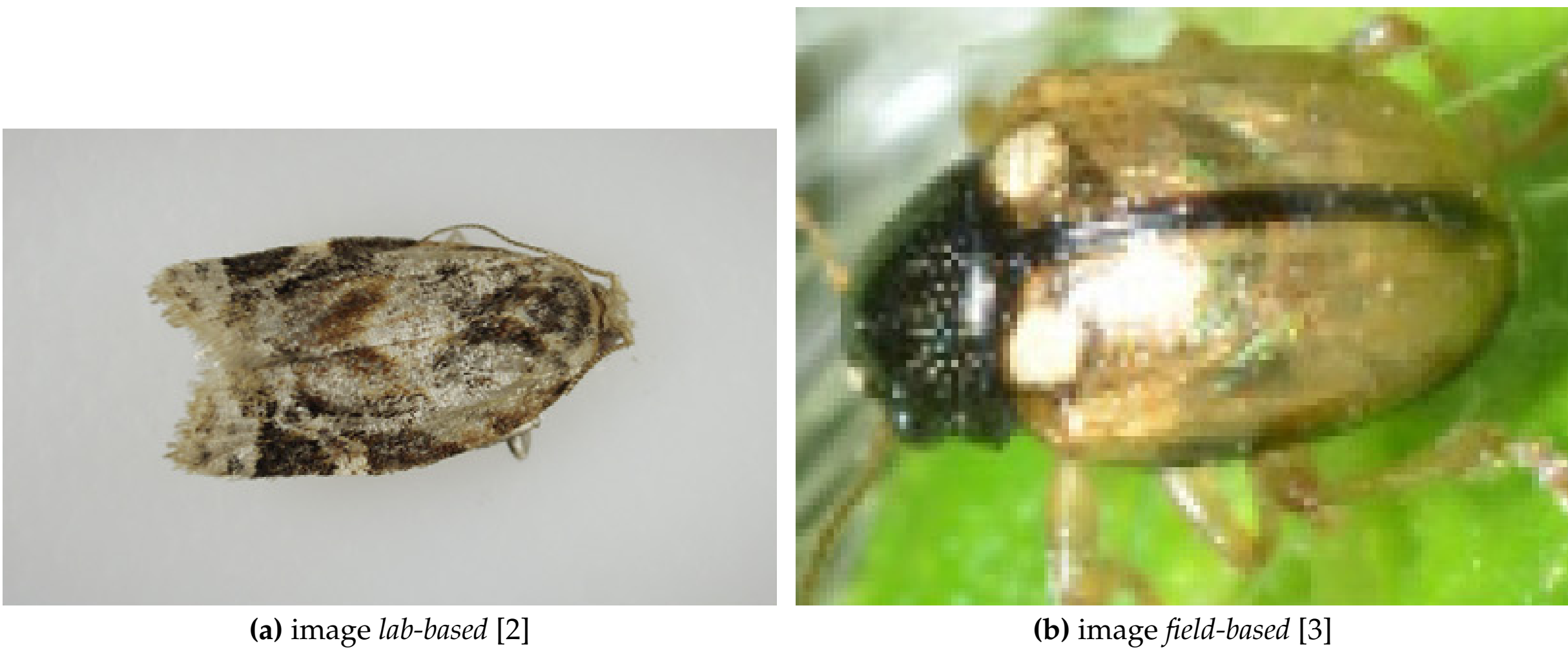


FIGURE 1 – Exemples d’images

Du fait de la grande diversité des insectes, le nombre de classes est potentiellement très grand. Il se peut également que de nouvelles classes soient rencontrées ou même que les individus rencontrés au sein d’une même classe soient assez différents tout au long de la phase opérative. Enfin, l’ensemble des classes d’insectes est en fait un arbre : on peut réaliser l’identification à plusieurs niveaux de granularité (au niveau de l’espèce, du genre, de la famille ou de l’ordre — du plus précis au plus global). L’usage d’un niveau ou d’un autre dépend du cadre de la reconnaissance, comme des organismes devant être reconnus.

Travaux antérieurs

Le développement de cette problématique a donné lieu à un article de revue de la littérature [1]. Celui-ci détaille suivant trois grands axes d’analyse (acquisition, représentation et classification) comment a été abordée la reconnaissance d’images d’insectes au fil des trois dernières décennies. Le focus est ici mis sur les approches de constructions des descripteurs. Trois grandes catégories d’approches se succèdent presque chronologiquement : les descripteurs “hand-crafted”, les descripteurs “mid-level” appris et les représentations hiérarchiques apprises.

Category Levels		
Handcrafted features	Domain-dependent	Wing’ Venations Geometry Shape Color Texture Raw Pixel
	Global and generic image features	SIFT Others
	Local features	BoW PCA MLP Sparse Coding
	Unsupervised representations	
	Supervised representations	
	Auto-encoder	
Hierarchical representations		

TABLE 1 – Types d’approches de représentation [1]

Le chemin dessiné par ces différents travaux pose une question : doit-on traiter les images d’insectes comme n’importe quelles images ou ont-elles des propriétés telles que des méthodes dédiées s’imposent pour les reconnaître ?

Méthode proposée

Le cadre de la présente étude est celui des images *lab-based* : les images sont prises dans des conditions de laboratoire, c’est-à-dire à l’aide d’un éclairage contrôlé et d’un fond uniforme. La base utilisée comprend trente classes différentes d’insectes. La technique utilisée se base sur l’architecture de réseaux de neurones profond VGG16. L’architecture a été apprise successivement sur les images du problème partant de poids aléatoires et de poids appris sur le problème ImageNet-1000, technique appelée *transfer learning*.

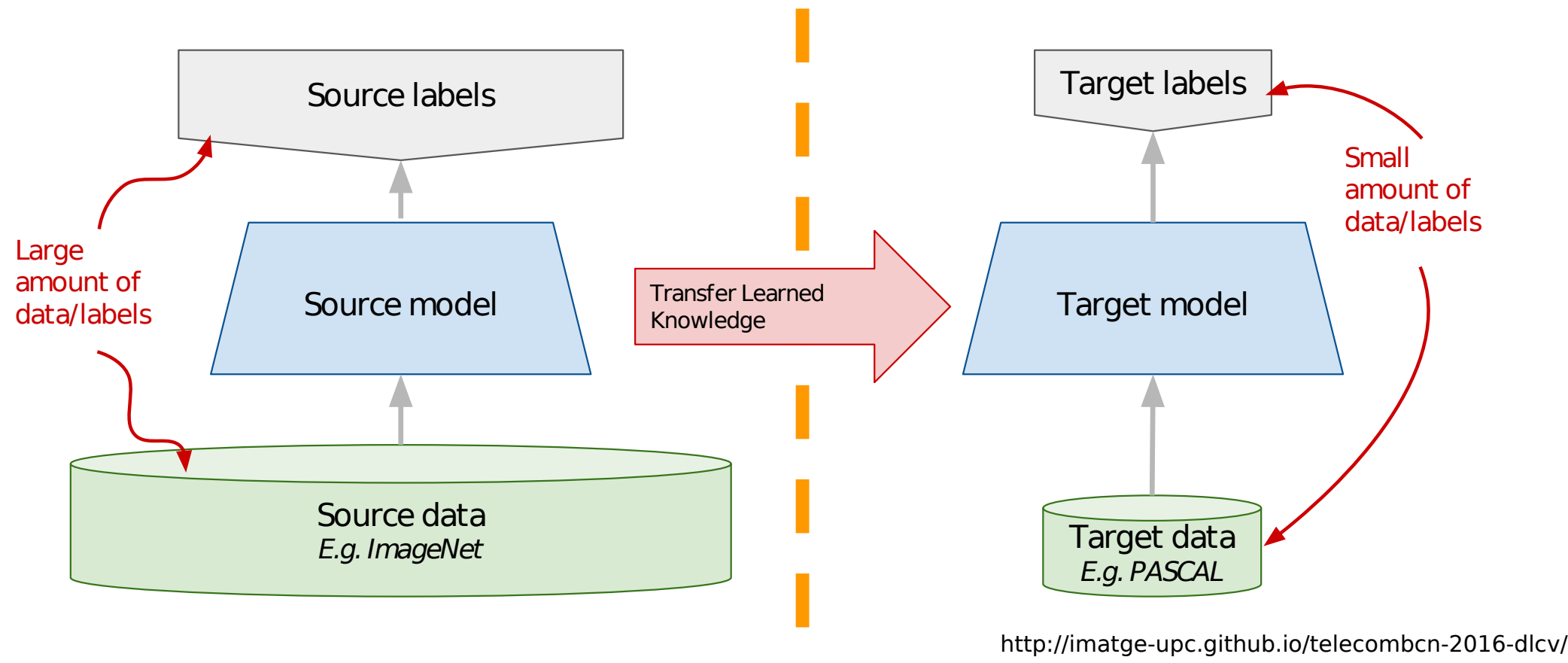


FIGURE 2 – Schéma explicatif sur le transfer learning

L’hypothèse de départ est que les connaissances apprises sur le problème ImageNet sont transférables au problème de reconnaissance d’images d’insectes. Lors de la phase de transfert des connaissances (à droite sur la FIGURE 2), le modèle peut être appris totalement ou partiellement : on pourra donc appliquer la rétropropagation sur un nombre plus ou moins grand de couches de VGG16, laissant les autres inchangées.

Expérimentations

- Pour tester cette hypothèse, deux bases de données ont été constituées :
- IRBI : base *lab-based* collectée dans le cadre du projet de recherche CARAMBA. 30 classes avec une moyenne de 102 images/classe ± 70
 - ImageNet-arthropods : base *field-based* obtenue à partir du synset arthropods d’ImageNet et réduite de façon à obtenir la même cardinalité moyenne par classe qu’IRBI. 443 classes avec une moyenne de 96 images/classe ± 78

Étude comparative

- Sur ces deux bases d’images ont été appliquées les méthodes suivantes :
- SIFTBoW, un dictionnaire de mots visuels SIFT adossé à un classifieur SVM optimisé en grille.
 - VGG16
 - frsc : appris à partir de poids aléatoires
 - fitu : appris à partir de poids appris sur ImageNet-1000
 - fitu/w : idem mais avec pondération pour palier les disparités de cardinalités.

Méthode	IRBI		ImageNet-arthropods	
	Top-1	Top-5	Top-1	Top-5
SIFTBoW	52.3 % $\pm$ 3.7	82.7 % $\pm$ 3.3		
VGG16-frsc	37.3 % $\pm$ 2.8	70.3 % $\pm$ 2.1	19.5 % $\pm$ 0.8	40.4 % $\pm$ 0.7
VGG16-fitu	67.4 % $\pm$ 2.6	89.2 % $\pm$ 2.6	39.9 % $\pm$ 0.6	65.8 % $\pm$ 0.4
VGG16-fitu/w	70.6 % $\pm$ 5.1	90.3 % $\pm$ 1.9	40.6 % $\pm$ 0.6	68.3 % $\pm$ 0.9

TABLE 2 – Taux de reconnaissances moyens sur une validation croisée 5-pli

Étude avancée

La transférabilité des connaissances apprises sur ImageNet-1000 avec VGG16 a été étudiée de manière approfondie en faisant varier le nombre d’exemples d’apprentissages utilisés et le nombre de couches apprises.

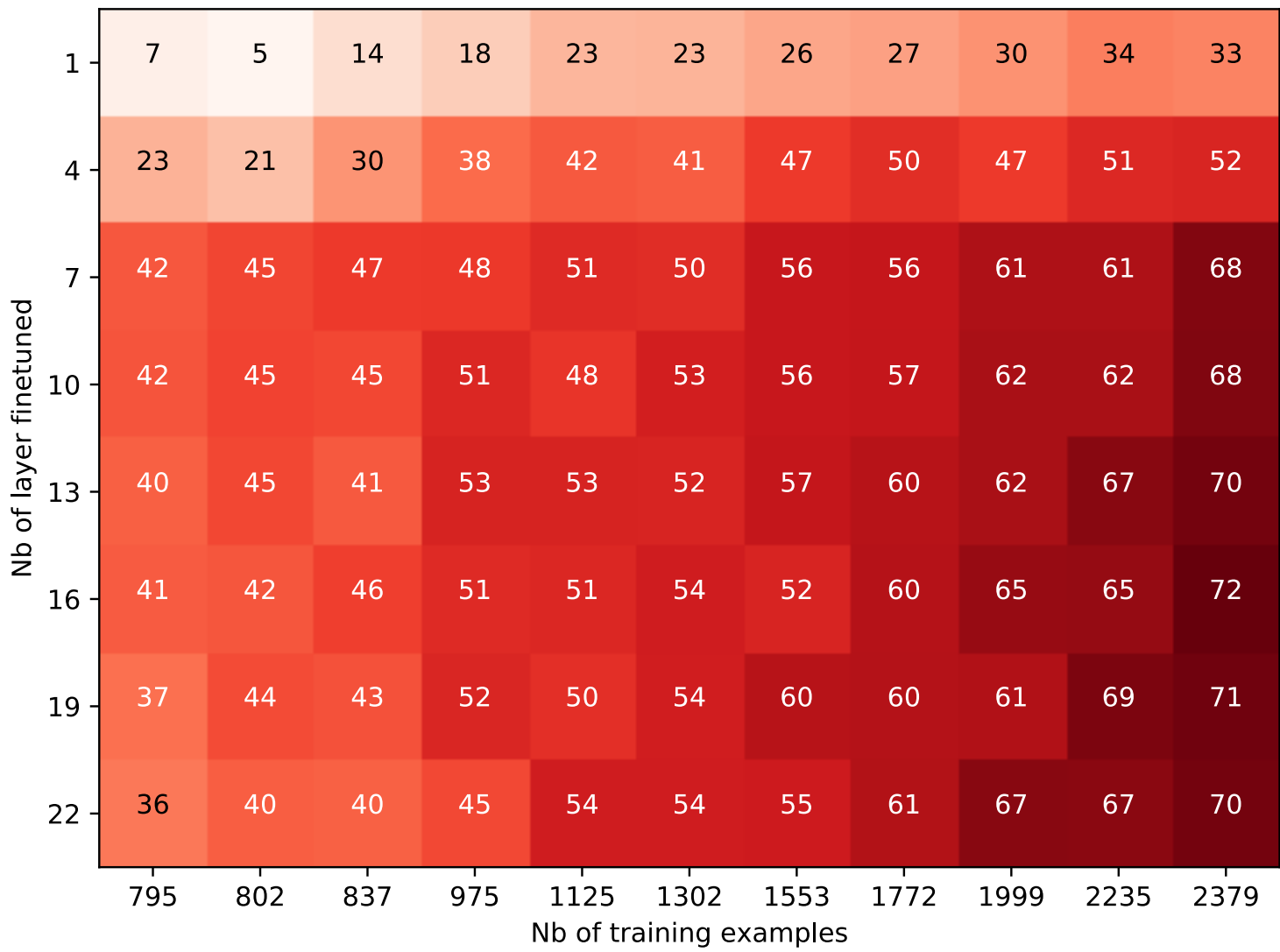


FIGURE 3 – Variations conjointes du nombre d’exemples d’apprentissages et du nombre de couches apprises

Conclusions

- L’architecture profonde VGG16 surpasse l’approche classique par dictionnaire de mots visuels SIFT,
- les connaissances apprises sur un problème généraliste tel qu’ImageNet-1000 sont transférables au problème de reconnaissance d’images d’insectes,
- et ce jusqu’à un niveau d’abstraction élevé.

Références

[1] Maxime Martineau, Donatello Conte, Romain Raveaux, Ingrid Arnault, Damien Munier, and Gilles Venturini. A survey on image-based insect classification. *Pattern Recognition*, 65 :273 – 284, 2017.

[2] Chenglu Wen, Daoxi Wu, Huosheng Hu, and Wei Pan. Pose estimation-dependent identification method for field moth images using deep learning architecture. *Biosystems Engineering*, 136 :117–128, August 2015.

[3] Chengjun Xie, Jie Zhang, Rui Li, Jinyan Li, Peilin Hong, Junfeng Xia, and Peng Chen. Automatic classification for field crop insects via multiple-task sparse representation and multiple-kernel learning. *Computers and Electronics in Agriculture*, 119 :123–132, November 2015.

Remerciements

Cette étude a été financée par la Région Centre-Val de Loire. Merci à Clément Chatelain (LITIS, INSA de Rouen) pour son aide.