Use of deep learning methods applied to the study of ticks and tick-borne diseases

Sk Imran Hossain, Yann Frendo, Valerie Poux, Isabelle Lebert, Olivier Lesens, Delphine Martineau, Engelbert Mephu Nguifo, Gwenaël Vourc'h, Jocelyn de Goër de Herve

Use of deep learning methods applied to the study of ticks and tick-borne diseases

I. Hossain, Y. Frendo, V. Poux, I. Lebert, D. Martineau, O. Lesens, E. Mephu-Nguifo, G. Vourc'h, J. De Goër

Session - Management and integration of agronomical, phenotypical and environmental data

Jocelyn DE GOËR
UMR EPIA

July 7th 2022

JOBIM 2022
05>08JUIL
Université de Rennes 1 | RENNES

Ticks and tick-borne diseases

- ✓ **Ticks are acarians, ectoparasites and hematophagous**
  - ✓ 12 genus, 700 species of hard ticks in the world
  - ✓ 4 genus in France
- ✓ **Ticks are vectors of diseases:**
  - ✓ Transmission during blood feeding

Before blood feeding

Blood feeding on human skin

After blood feeding

Metropolitan France

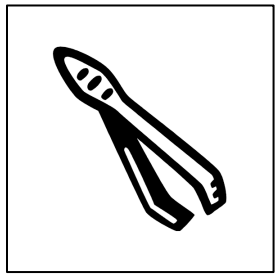
Southern France

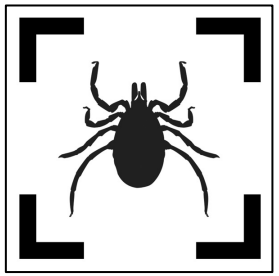
Ixodes	***Dermacentor***	***Rhipicephalus***	***Hyalomma***
			
- Babesia *- Powassan* *- Anaplasma* *- B. miyamotoi* *- Ehrlichia* *- F. tularensis* *- Bartonella* *- B. burgdorferi (Lyme)*	*- Babesia* *- Colorado Tick Fever* *- Rickettsiae* *- F. tularensis*	*- Babesia* *- Rickettsiae* *- Bartonella*	*- Crimean-Congo hemorrhagic fever*

- **Parasite**
- **Virus**
- **Bacteria**

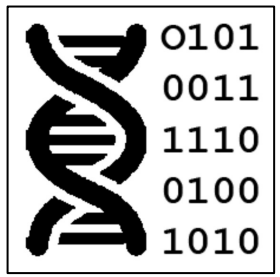
Ticks and tick-borne diseases

- **Study and surveillance of ticks and ticks-borne diseases**


Collection in the field

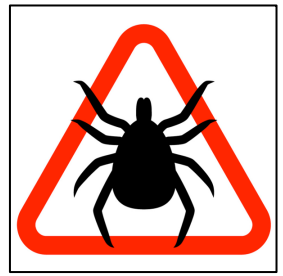

Identification


Pathogen detection



Bio-informatics


Statistics


Prevention

- **Lyme disease**
  - Lyme disease is the most reported tick-borne disease in US and Europe
    - Estimation of cases each year:
      - USA: 400 000 (source CDC)
      - France: 60 000 (source SPF)

Research projects

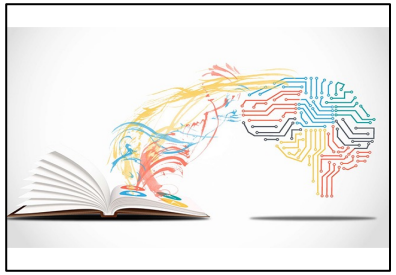
- ✓ **Two projects about Lyme disease and ticks surveillance**
  - ✓ **2018: DAPPEM Project**
    - ✓ Mobile application development for early skin signs of Lyme Disease (erythema migrans)
  - ✓ **2020: DCLIC Project**
    - ✓ Real-time ticks genus identification system
- ✓ **These projects are based on Artificial Intelligence methods**

Artificial Intelligence

- **Artificial intelligence:**
  - **AI symbolic** : based on formal and logical descriptions (algorithms)
    - Experts systems
  - **AI connectionist** : the machine is able to "learn" by itself from data
    - Deep Learning

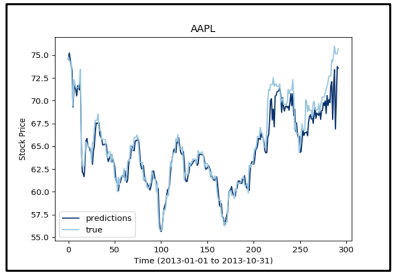
- **Since 2012, sufficient computing capacity for its application:**
  - GPU-based computing (Graphics Processing Units)

- **Deep Learning is used in many domains:**

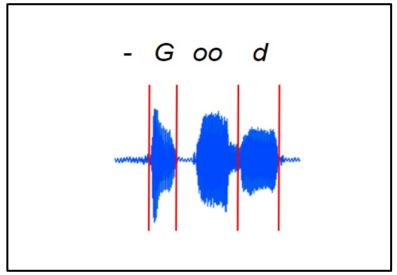


Text Mining

Bio-Informatics data analysis



Time series analysis



Sound analysis
Voice recognition

Medical diagnostic assistance

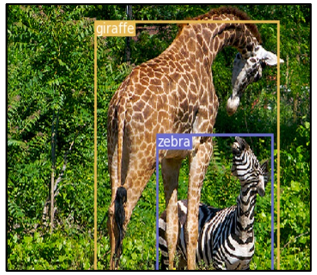
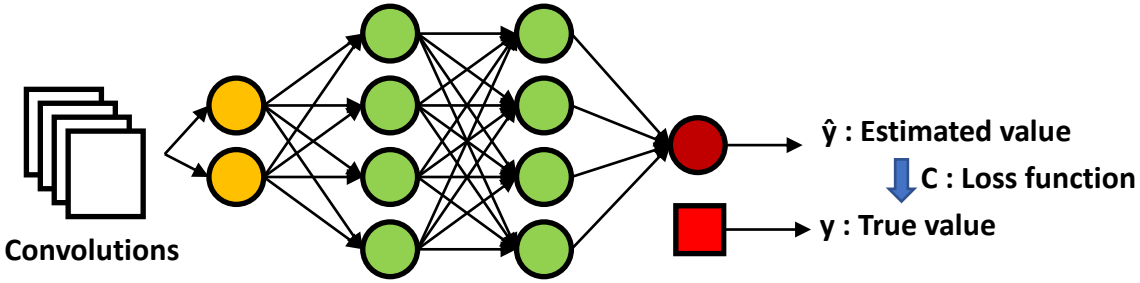


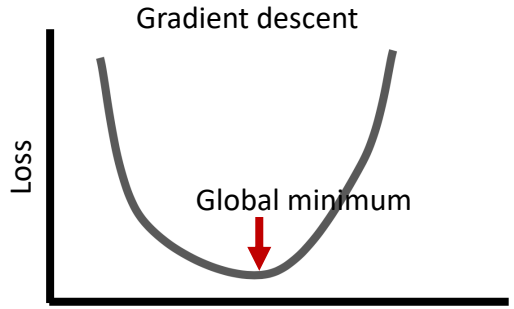
Image / video analysis

Artificial Intelligence

✓ Training of a neural network for image classification

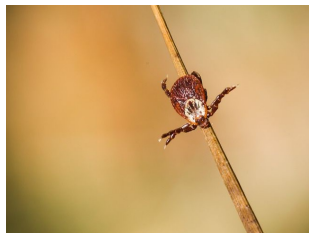
Image Datasets
Labeled and classified



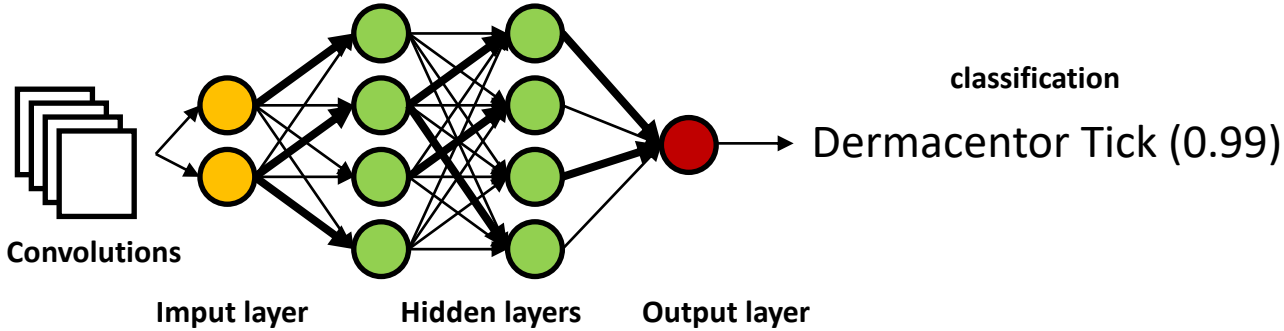


Calibration of the model with backpropagation method

✓ Query from image



Query image

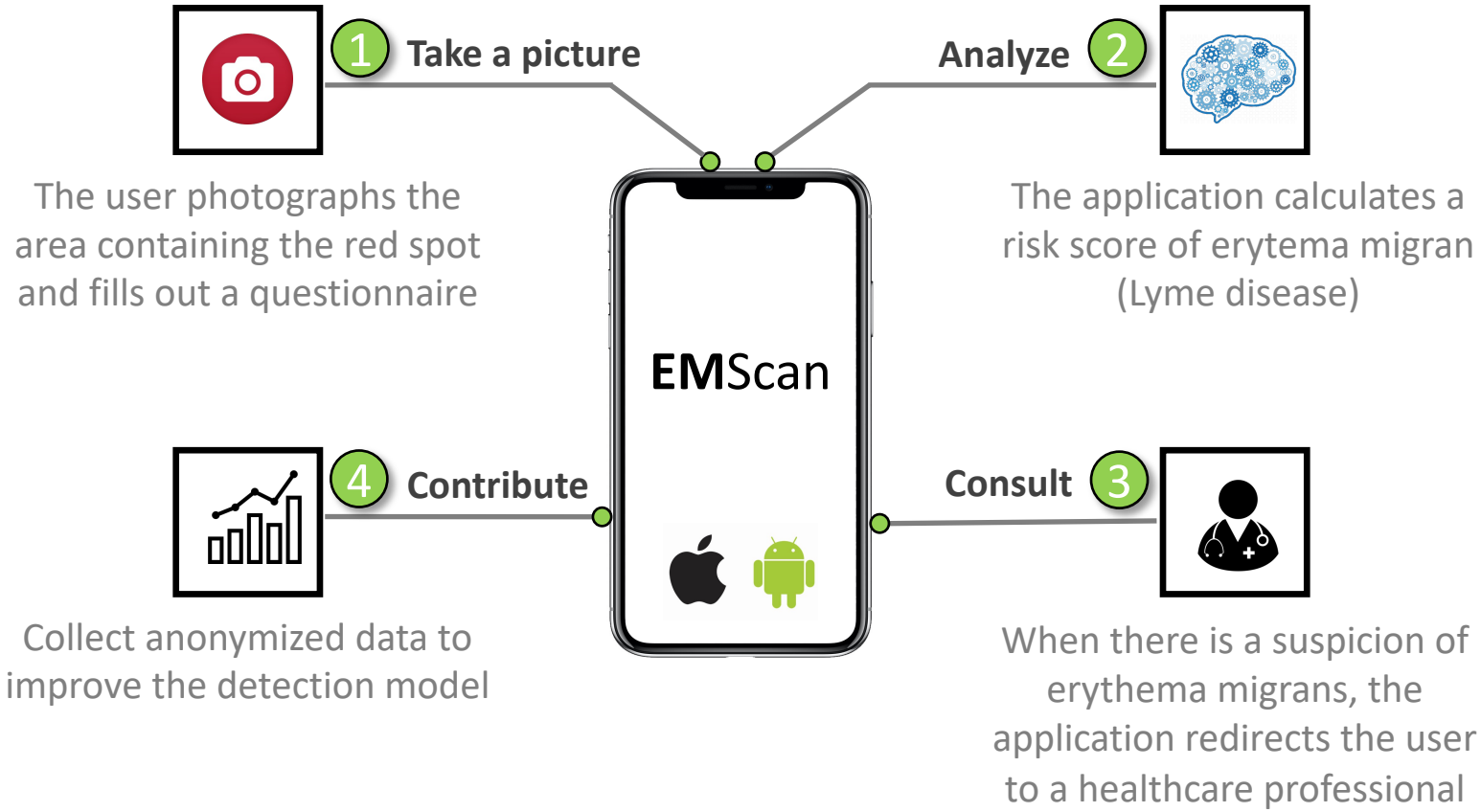


classification

Dermacentor Tick (0.99)

DAPPEM Project

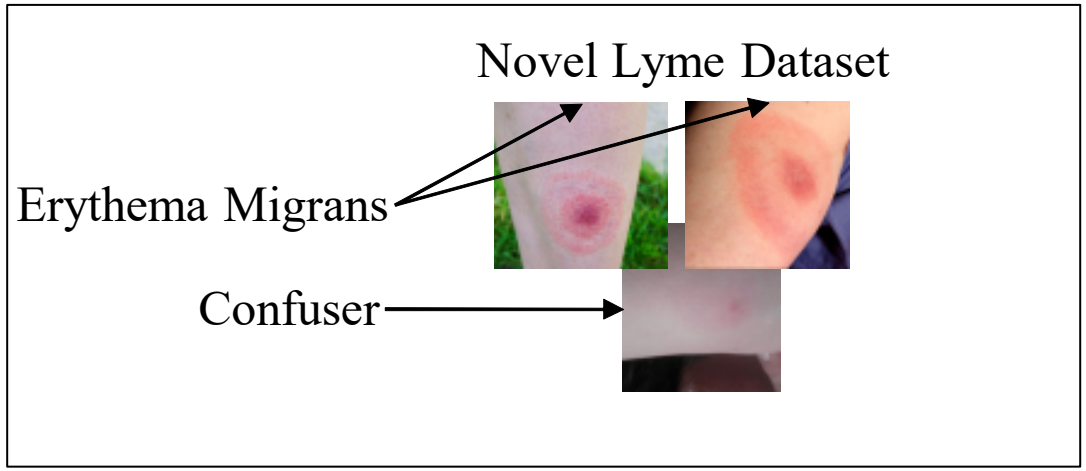
- ✓ **DAPPEM:** Développer une application permettant la détection des Erythèmes Migrants (premiers signes cutanés de la maladie de Lyme)
  - ✓ Target users: Health care workers and general public

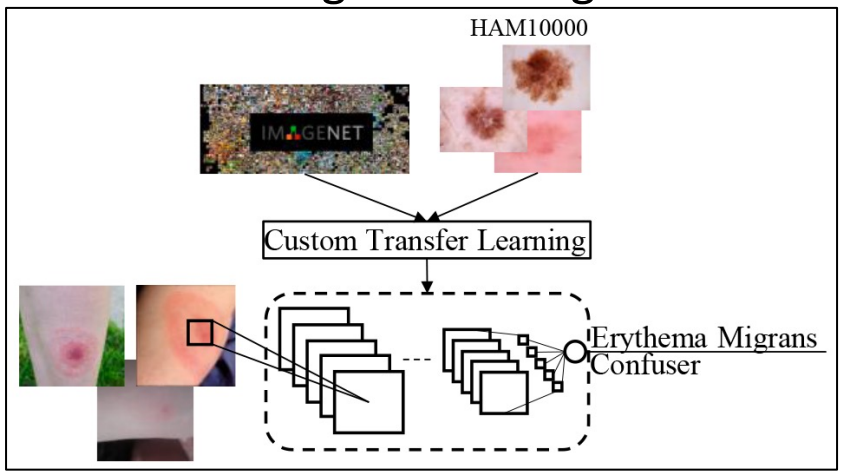


Erythema migrans
Early sign of Lyme disease

DAPPEM Project

- ✓ **Develop and train an artificial neural network to identify EMs from a photo**
  - ✓ Creation of a training dataset
    - ✓ 866 images of EM identifyed by doctors
    - ✓ 806 images of non EM (confuser class)
  - ✓ Training 23 CNN architectures
    - ✓ Transfert Learning
    - ✓ Adding other image dataset





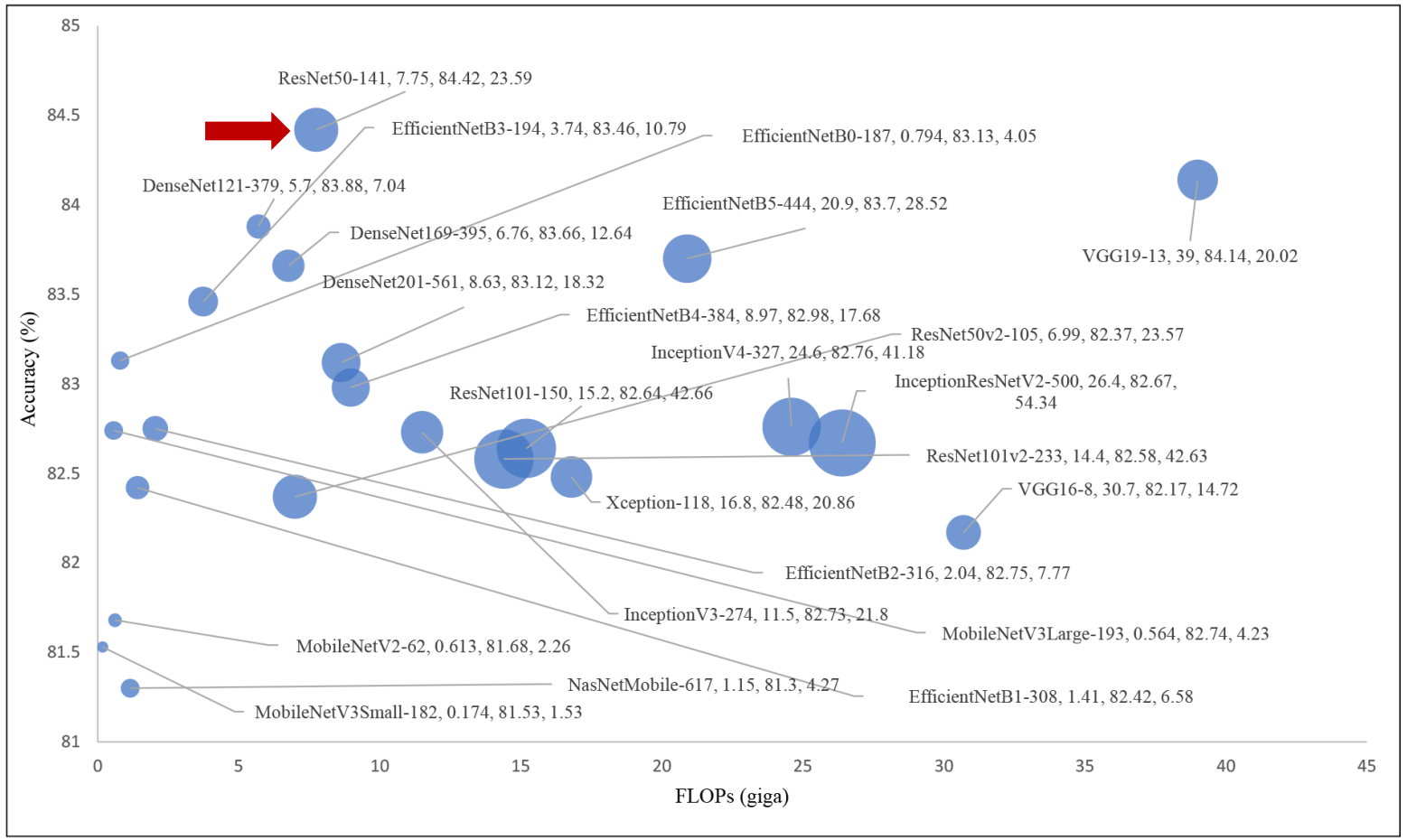
Work done by Imran Hossain, PhD student at LIMOS, supervised by E. Mephu-Nguifo and J. de Goër

Publication: Sk. Imran Hossain and al, *Exploring convolutional neural networks with transfer learning for diagnosing Lyme disease from skin lesion images*

Journal Comput. Methods Programs Biomed., 2022

DAPPEM Project

✓ 23 CNN trained models comparison



Hossain et al. 2022. Computer Methods and Programs in Biomedicine

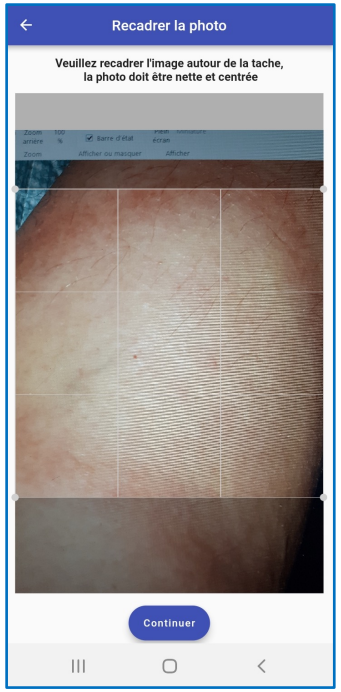
Selected model:

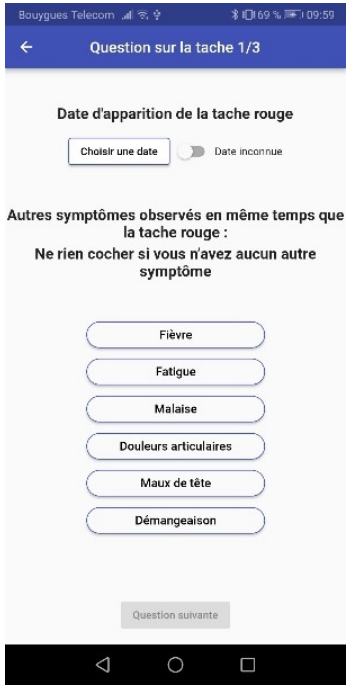
Architecture:	ResNet50-141
Accuracy:	84,4 %
Sensitivity:	87,9 %
Specificity:	80,7 %
Parameters:	23 million

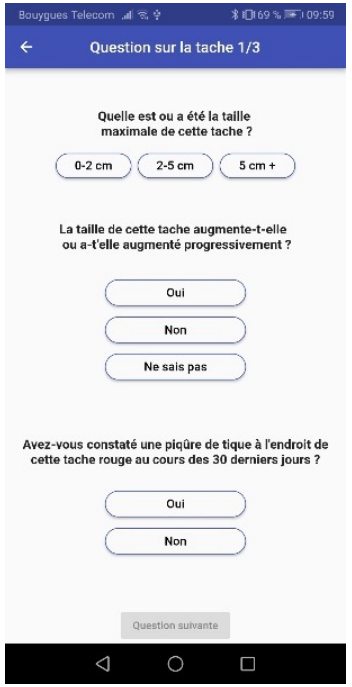
Size represents number of model parameters (million units)

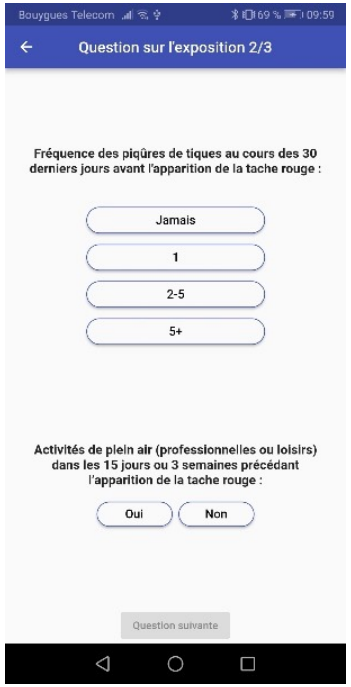
DAPPEM Project

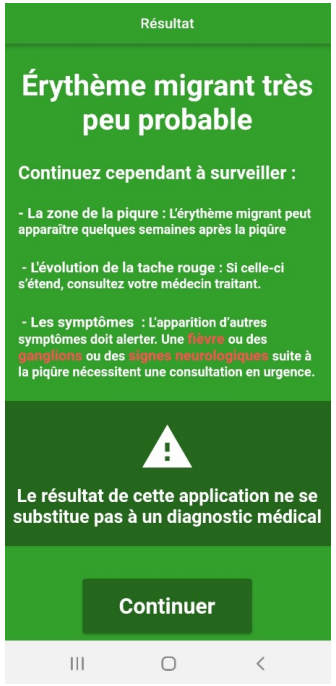
✓ EMScan mobile application mains screens:

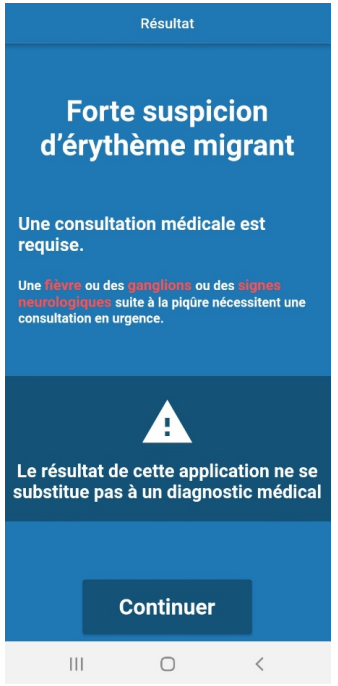












Take picture

Questionnaire

Result screens

DAPPEM Project

- **Research units:**
  - EPIA (UMR INRAE – VAS). J. de Goër, I. Lebert, G. Vourc'h
  - LMGE (UMR UCA-CNRS) / CHU : O. Lesens, D. Martineau
  - LIMOS (UMR UCA-CNRS-EMSE). E. Mephu-Nguifo, I. Hossain (Ph.D student), Y. Frendo
- **Non-academic partners:**
  - ONF, Office National des Forêts. C. Planchette
  - MSA, Mutualité Sociale Agricole. M. Ruols
  - CNPF, Centre National de la Propriété Forestière et CRPF. M. D. Mourisset
- **Funding:**
  - European Auvergne–Rhône-Alpes Regional Development Fund: 190k€
  - MSA 2019 et 2020: 30k€
- **EMScan application will be available in june 2023**
  - Under certification as medical device by ANSM

DCLIC Project

- ✓ **Propjet DCLIC:** Deep Convolutional Learning for *Ixodidae* Characterization

- ✓ **Objective of the project:**
  - ✓ Evaluate the use of Deep Learning to identify ticks genus from photo or video
    - ✓ Ticks present in France: *Ixodes, Dermacentor, Rhipicephalus et Hyalomma*

- ✓ **Precision**:
  - ✓ Real-time location and identification of one or more ticks at the same time
  - ✓ Tracking
  - ✓ Counting

DCLIC Project

✓ Creation of a dataset for training

- ✓ Photos with enough detail to be identified
- ✓ 1,636 photos of ticks manually classified by genus

391 *Ixodes*

253 *Dermacentor*

170 *Hyalomma*

50 *Rhipicephalus*

✓ Adding images without ticks

- ✓ 15,375 images of insects, acarians and spiders

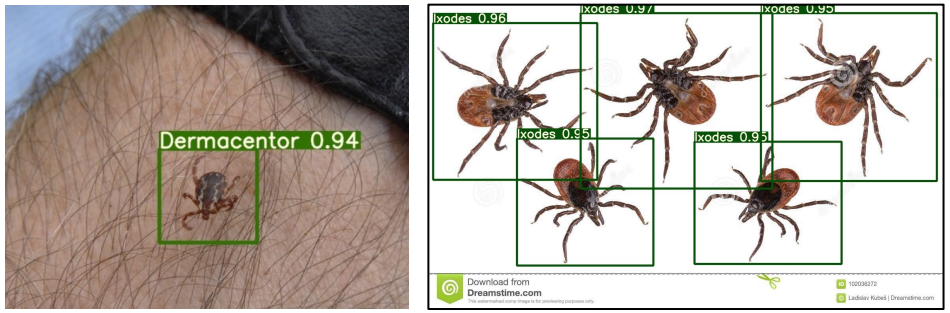
391	*Halyomorpha*	68	*Tetranychidae*	2 278	*Odonata*
2 107	*Lepidoptera*	2 049	*Hymenoptera*	2 388	*Hemiptera*
2 030	*Dipter*	2 111	*Coleoptera*	2 419	*Araneae*

Tetranychidae

Thomisidae

DCLIC Project

- **Characteristics of the model:**
  - Family of architecture: YoloV5 (https://github.com/ultralytics/yolov5)
  - Number of layers: 500
  - Number of parameters: 46 652 890
  - Size of images: 640x640px
  - Training time 500 epochs: 45h (2x NVIDIA RTX 8000)
    - 5 trainings were necessary
- **Results:**
  - Accuracy: 0.96
  - Real-time detection (@30fp/s) on a smartphone

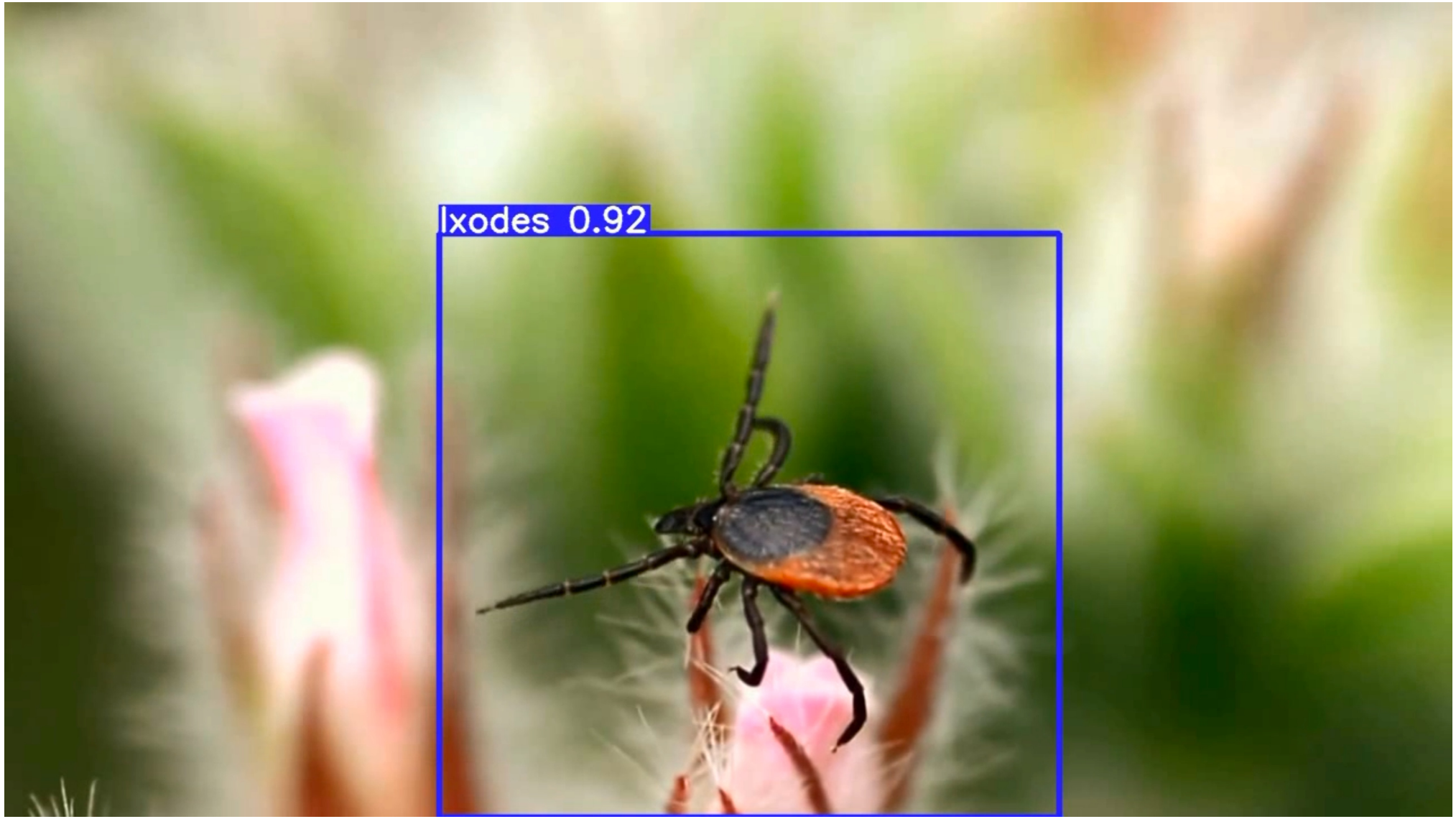


Simple and multiple identification

- **Limitations:**
  - Detection problem when ticks are too small on the image
    - Diversity of the training dataset
    - Limitations of current neural networks
  - A second experiment is in preparation

Tick not detected

DCLIC Project



DCLIC Project

- **Research units:**
  - UMR EPIA (UMR INRAE – VAS): J. de Goër, V. Poux, M. Rene-Martelet, G. Vourc'h
  - UMR BioEpar: O. Plantard
  - UMR ASTRES: F. Stachurski, L. Vial
  - Programme CiTIQUE: J. Durand, A. Brun-Jacob, J. Marchand, I. Carravieri, P. Frey-Klett

- **Funding:**
  - INRAE Animal Health Department: 10k€

- **Perspectives:**
  - Integration into the CiTIQUE participatory action research project
  - Share the model with the community

https://citique.fr

Conclusion

- ✓ **Deep Learning is a great tool**
  - ✓ Allowed significant improvements in image and analysis
- ✓ **Requires:**
  - ✓ Sufficient representative data
  - ✓ GPU computing resources
- ✓ Methods could be improved, especially in terms of the explicability of the results
- ✓ It is necessary to take into account the context

Tree or broccoli


Apple or owl


Chiwawa or muffin

Thank you !

Mini-symposium - Management and integration of agronomical, phenotypical and environmental data

Jocelyn DE GOËR

UMR EPIA

July 7th 2022

JOBIM 2022
05>08JUIL
Université de Rennes 1 | RENNES