Oral Presentation in LABELS Workshop at MICCAI 2018, Granada, Spain

Capsule Networks against Medical Imaging Data Challenges

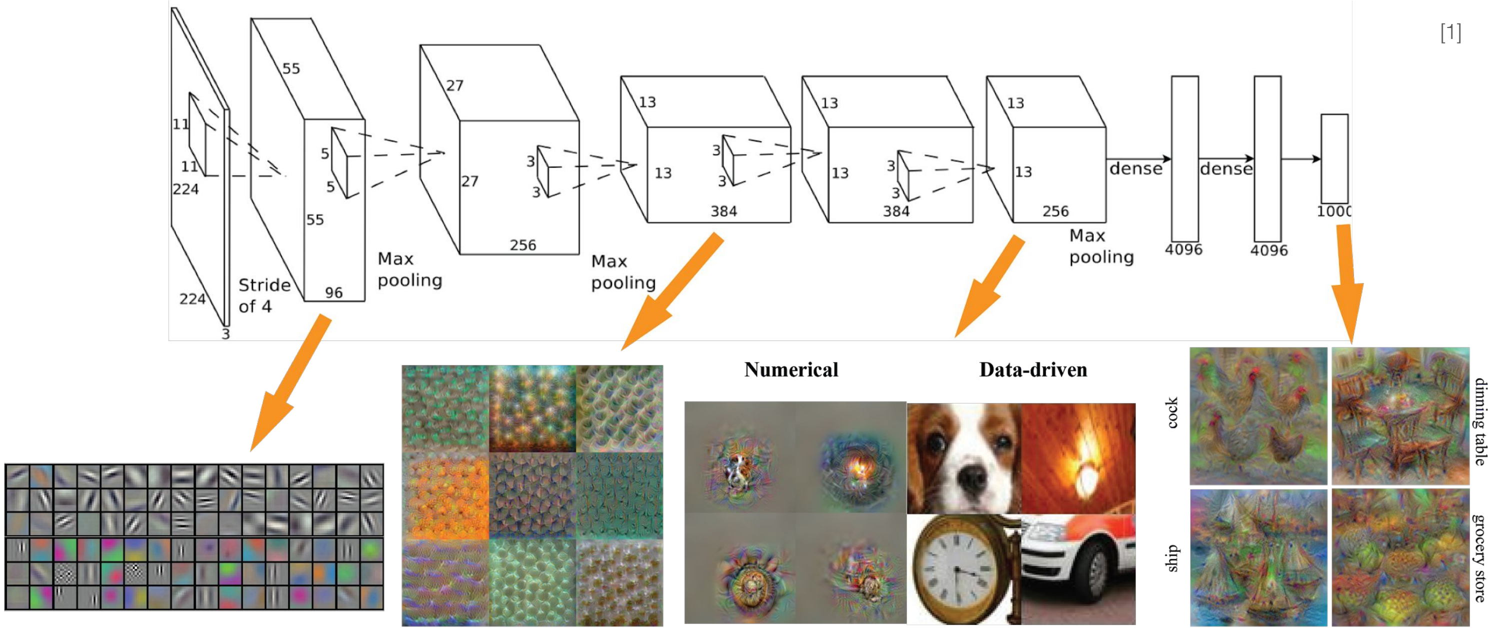
Amelia Jiménez-Sánchez
Universitat Pompeu Fabra

Shadi Albarqouni
Technische Universität München

Diana Mateus
École Centrale de Nantes

CONVOLUTIONAL NETWORKS: FEATURE EXTRACTION

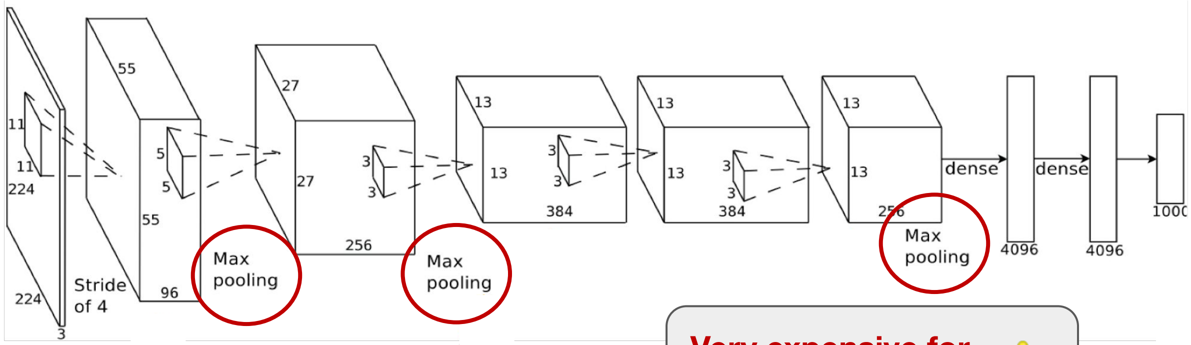
[1]



Conv 1: Edge+Blob **Conv 3: Texture** **Conv 5: Object Parts** **Fc8: Object Classes**

[1] http://vision03.csail.mit.edu/cnn_art/index.html

CONVOLUTIONAL NETWORKS: SHORTCOMINGS



[1]

- ConvNets are **not spatial invariant**, need to include: scale, rotations, translations

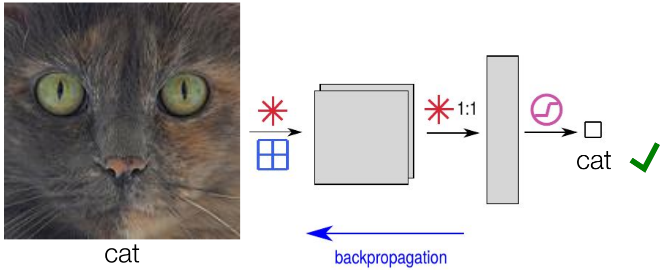
Very expensive for medical images

[2]

[1] http://vision03.csail.mit.edu/cnn_art/index.html [2] https://www.flickr.com/ #cat

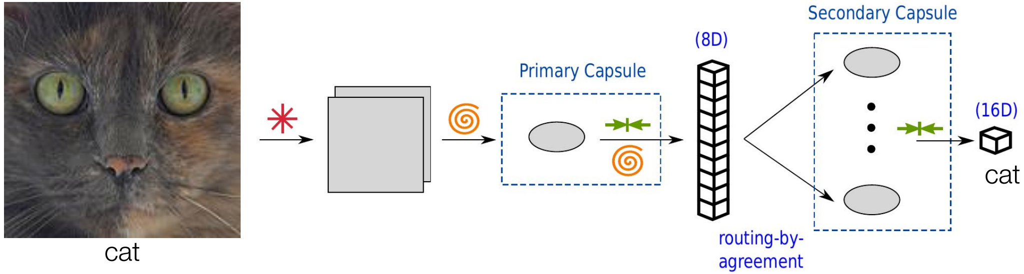
CONVOLUTIONAL *vs.* CAPSULE NETWORK

ConvNet



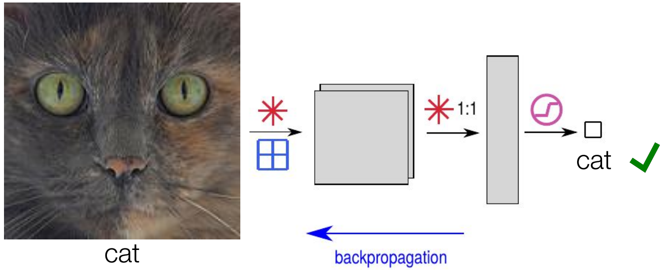
presence / absence

CapsNet



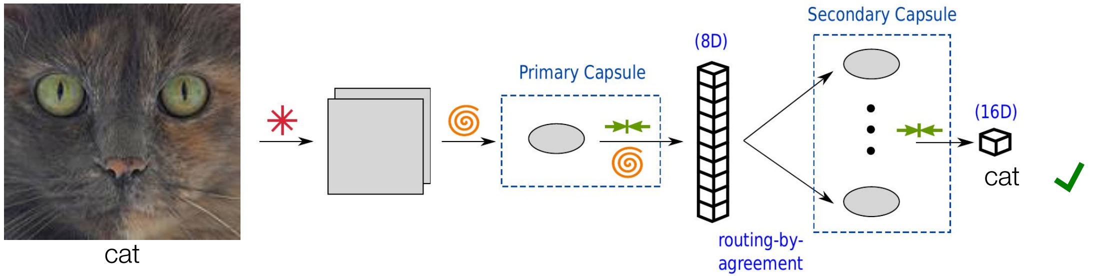
CONVOLUTIONAL *vs.* CAPSULE NETWORK

ConvNet



presence / absence

CapsNet

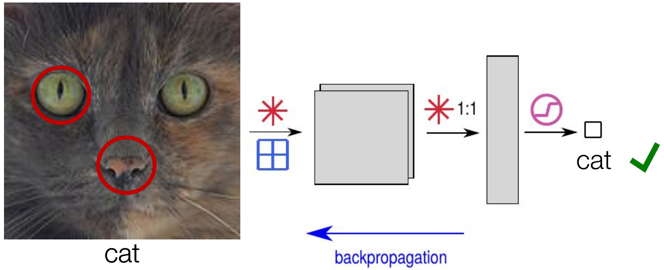


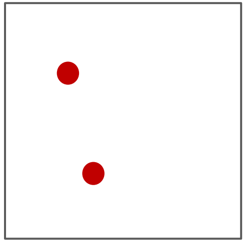
presence / absence (length of the vector)

+ pose: e.g. spatial location, scale, rotation, etc

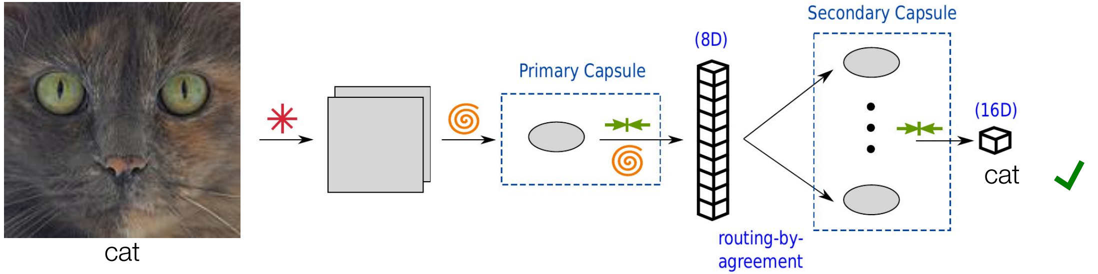
CONVOLUTIONAL *vs.* CAPSULE NETWORK

ConvNet



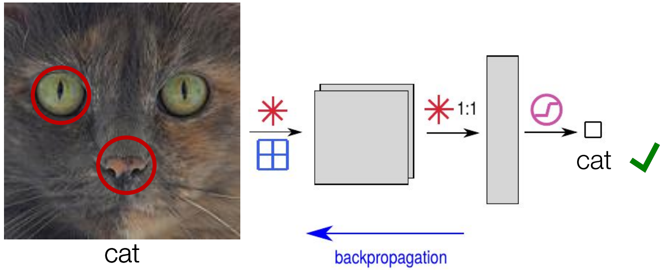


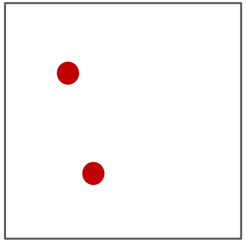
CapsNet



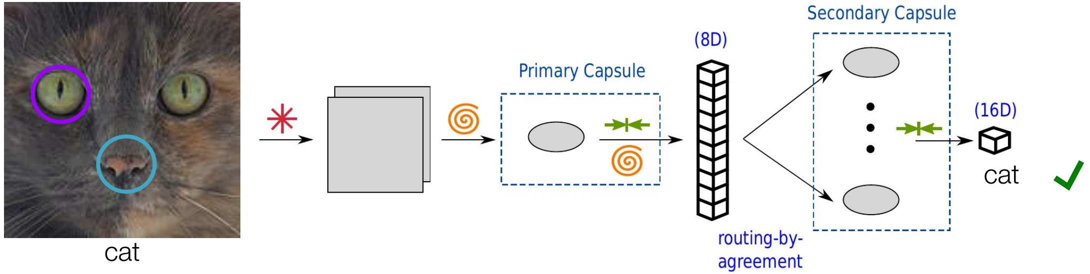
CONVOLUTIONAL *vs.* CAPSULE NETWORK

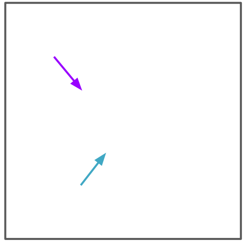
ConvNet





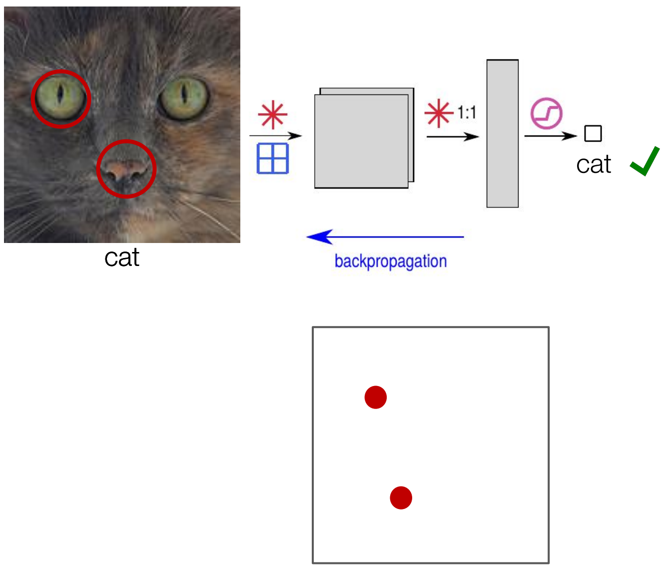
CapsNet



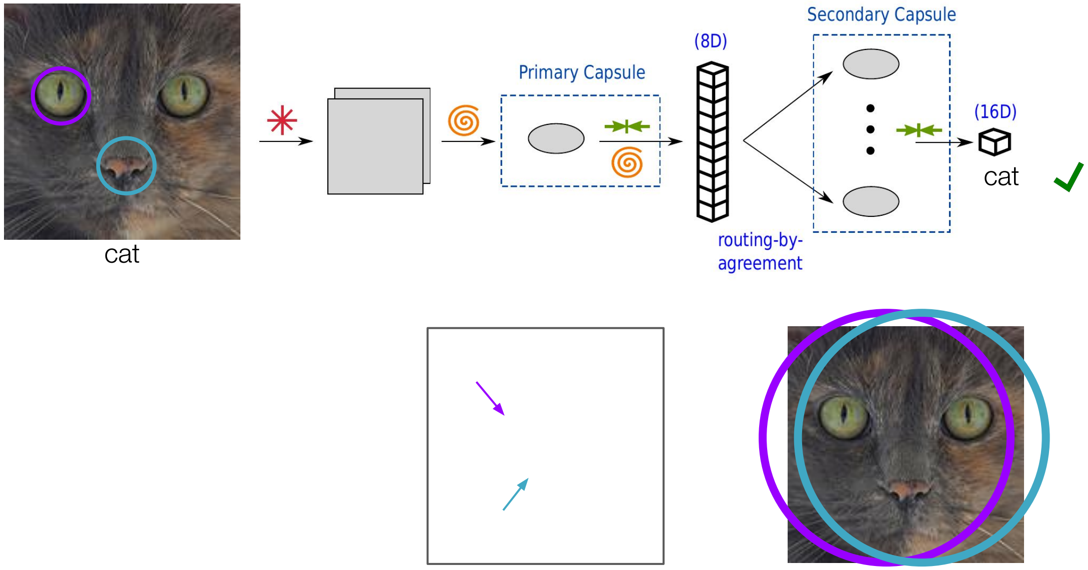


CONVOLUTIONAL *vs.* CAPSULE NETWORK

ConvNet



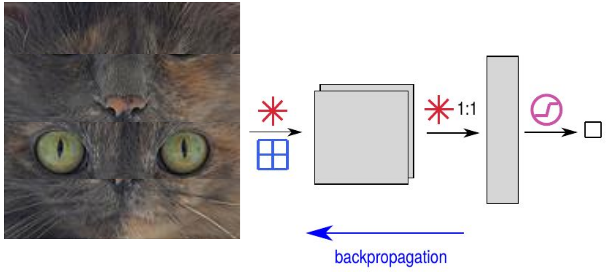
CapsNet



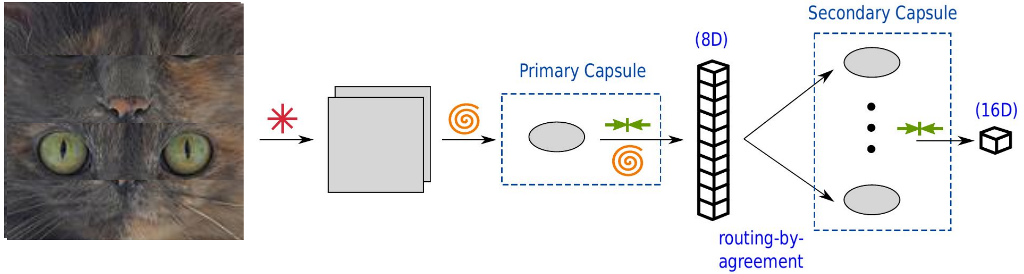
strong agreement

CONVOLUTIONAL *vs.* CAPSULE NETWORK

ConvNet

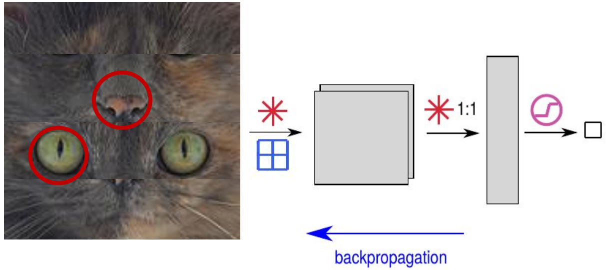


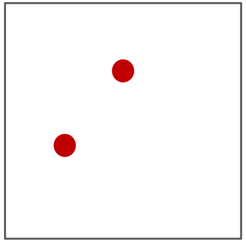
CapsNet



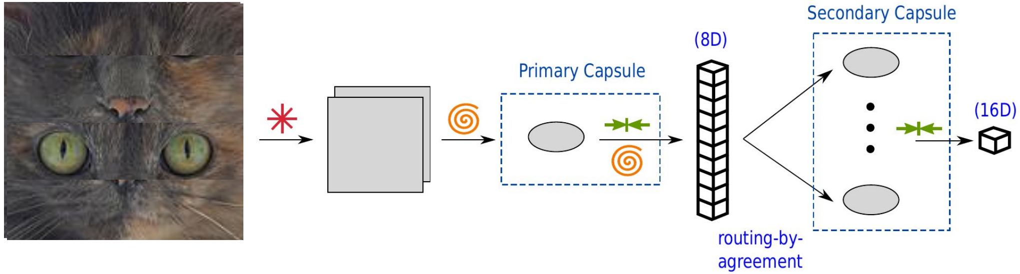
CONVOLUTIONAL *vs.* CAPSULE NETWORK

ConvNet



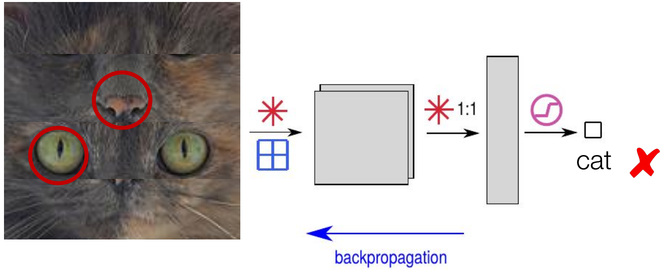


CapsNet

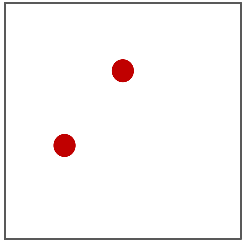


CONVOLUTIONAL *vs.* CAPSULE NETWORK

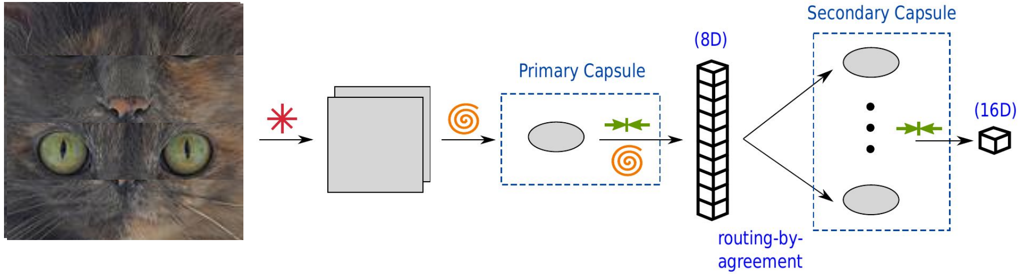
ConvNet



detect **"object" parts** (features) are present

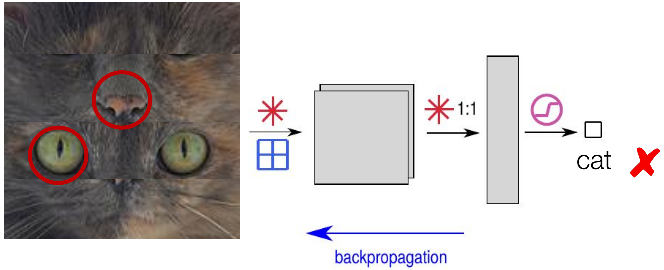


CapsNet

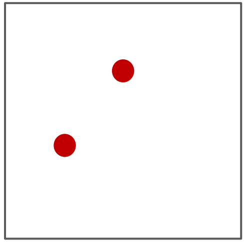


CONVOLUTIONAL *vs.* CAPSULE NETWORK

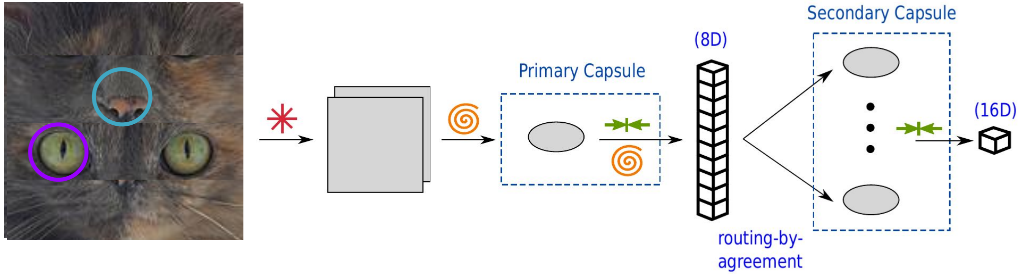
ConvNet

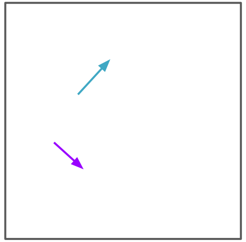


detect **“object” parts** (features) are present



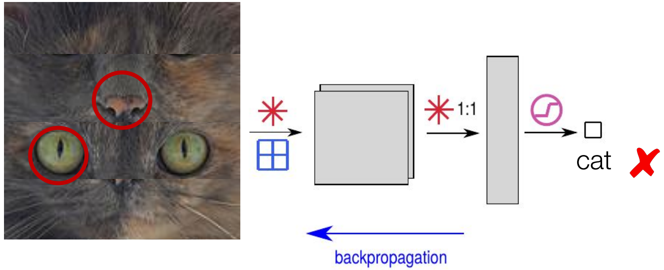
CapsNet



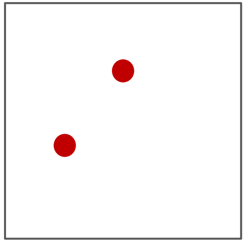


CONVOLUTIONAL *vs.* CAPSULE NETWORK

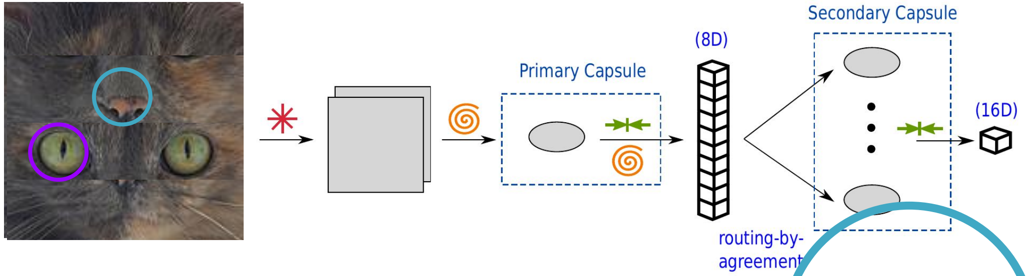
ConvNet

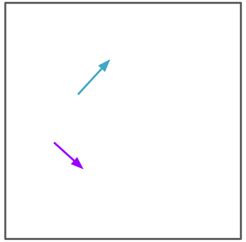


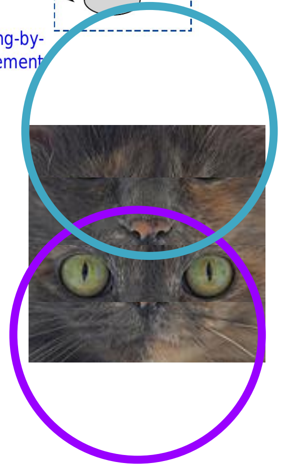
detect **"object" parts** (features) are present



CapsNet



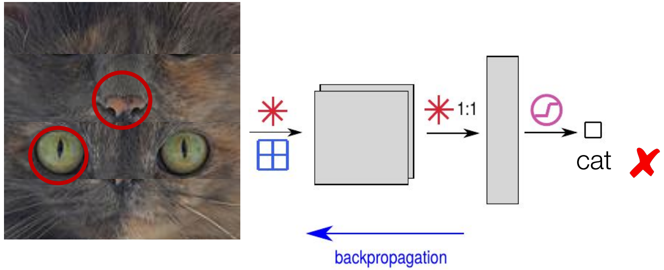




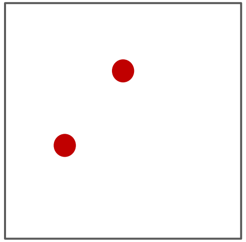
not agreement

CONVOLUTIONAL *vs.* CAPSULE NETWORK

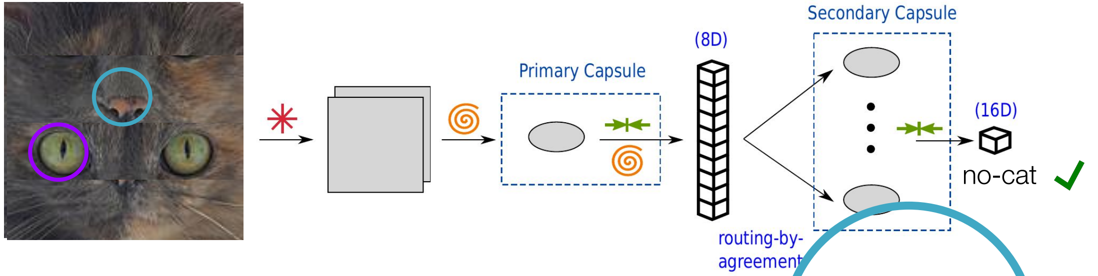
ConvNet

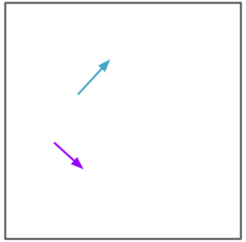


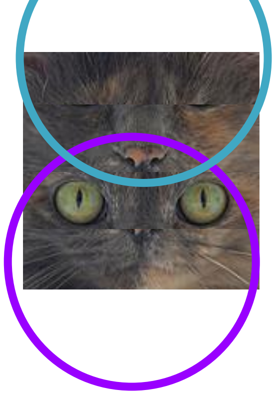
detect **“object” parts** (features) are present



CapsNet







not agreement

CONVOLUTIONAL *vs.* CAPSULE NETWORK

Summary of differences:

	ConvNets	**CapsNets**
Layer	pooling	---
Process	scalar	vector
Optimization	backpropagation	routing-by-agreement
Loss	cross-entropy	margin + reconstruction

CONVOLUTIONAL *vs.* CAPSULE NETWORK

Summary of differences:

	ConvNets	**CapsNets**
Layer	pooling	---
Process	scalar	vector
Optimization	backpropagation	routing-by-agreement
Loss	cross-entropy	margin + reconstruction

Margin loss: $||v_k|| > 0.9$ -> instance is present
$||v_k|| < 0.1$ -> instance is absent

HYPOTHESIS

CapsNets are designed to **learn** the **pose** of the instance along its **presence**. Consequently, less variations of the instance (**fewer** annotated images) are needed.

Medical datasets are often **small** and **highly imbalanced**.

HYPOTHESIS

We argue that CapsNet will perform better than ConvNets under medical data challenges.

HYPOTHESIS

We argue that CapsNet will perform better than ConvNets under medical data challenges.

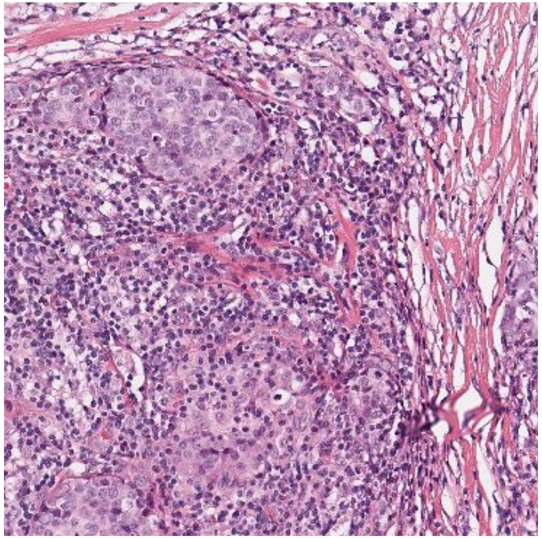
(1) How do networks behave under decreasing **amounts of training data**?

(2) Is there a change in their response to **class-imbalance**?

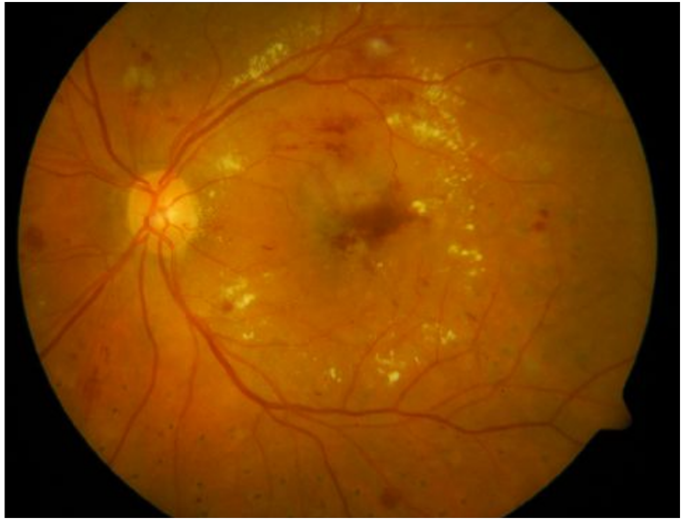
(3) Is there any benefit from **data augmentation** as a complementary strategy?

DATASETS

i) Mitosis detection (TUPAC16) [1]



ii) Diabetic retinopathy detection (DIARETDB1) [2]

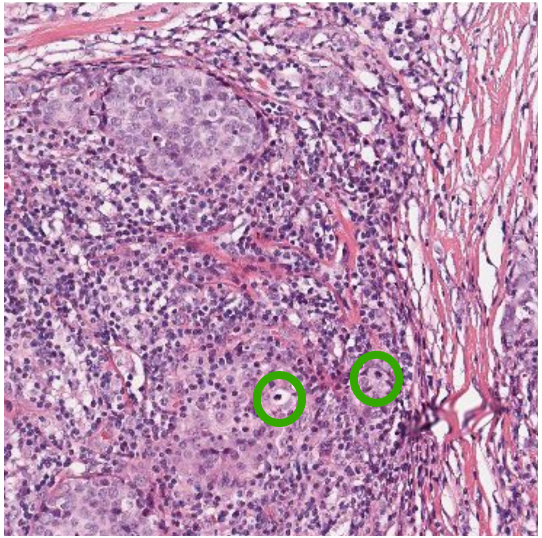


[1] Tumor Proliferation Assessment Challenge 2016
(TUPAC16 http://tupac.tue-image.nl/)

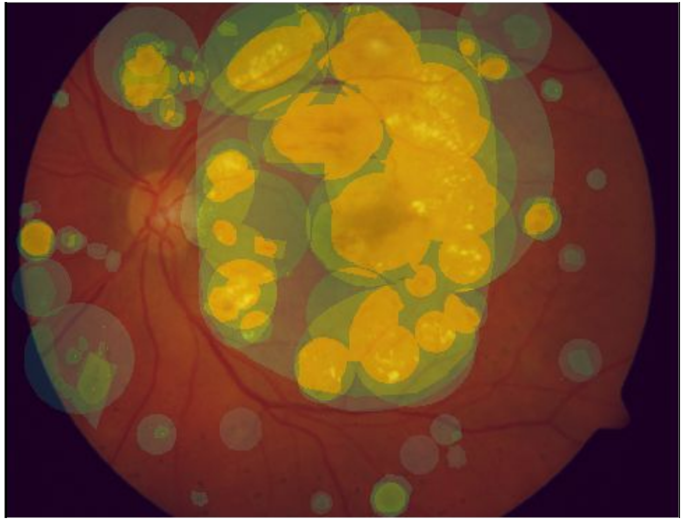
[2] Standard Diabetic Retinopathy Database - Calibration level 1
(DIARETDB1 http://www.it.lut.fi/project/imageret/diaretdb1/)

DATASETS

i) Mitosis detection (TUPAC16) [1]



ii) Diabetic retinopathy detection (DIARETDB1) [2]

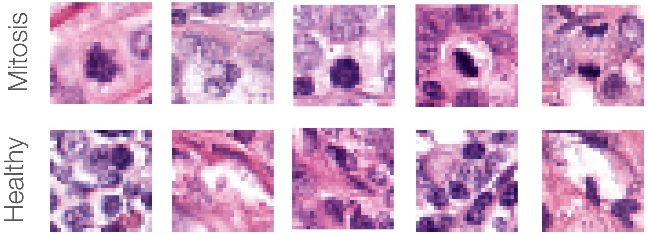


[1] Tumor Proliferation Assessment Challenge 2016
(TUPAC16 http://tupac.tue-image.nl/)

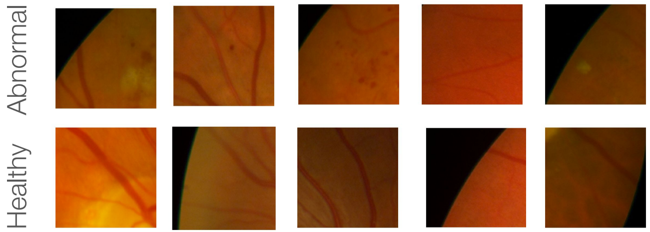
[2] Standard Diabetic Retinopathy Database - Calibration level 1
(DIARETDB1 http://www.it.lut.fi/project/imageret/diaretdb1/)

DATASETS

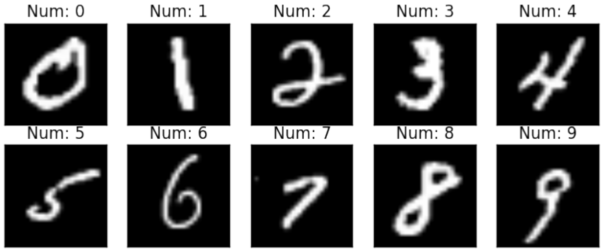
i) Mitosis detection (TUPAC16)



ii) Diabetic retinopathy detection (DIARETDB1)

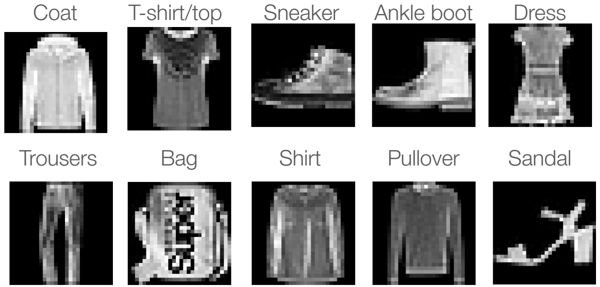


iii) Handwritten Digit Recognition (MNIST) [1]



[1] MNIST database of handwritten digits
(MNIST http://yann.lecun.com/exdb/mnist/)

iv) Clothes Classification (Fashion-MNIST) [2]

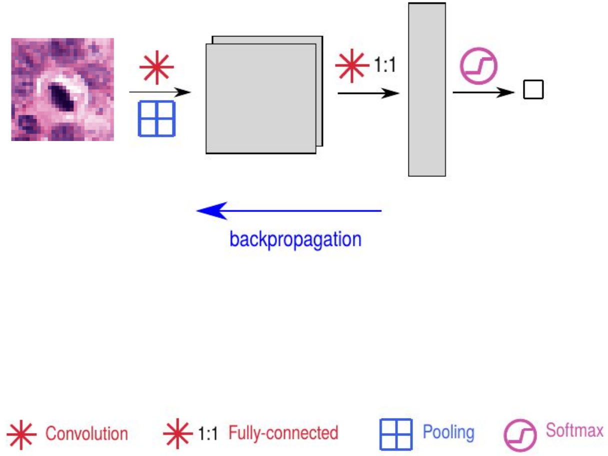


[2] Zalando's article images dataset (Fashion-MNIST
https://github.com/zalandoresearch/fashion-mnist)

ARCHITECTURES

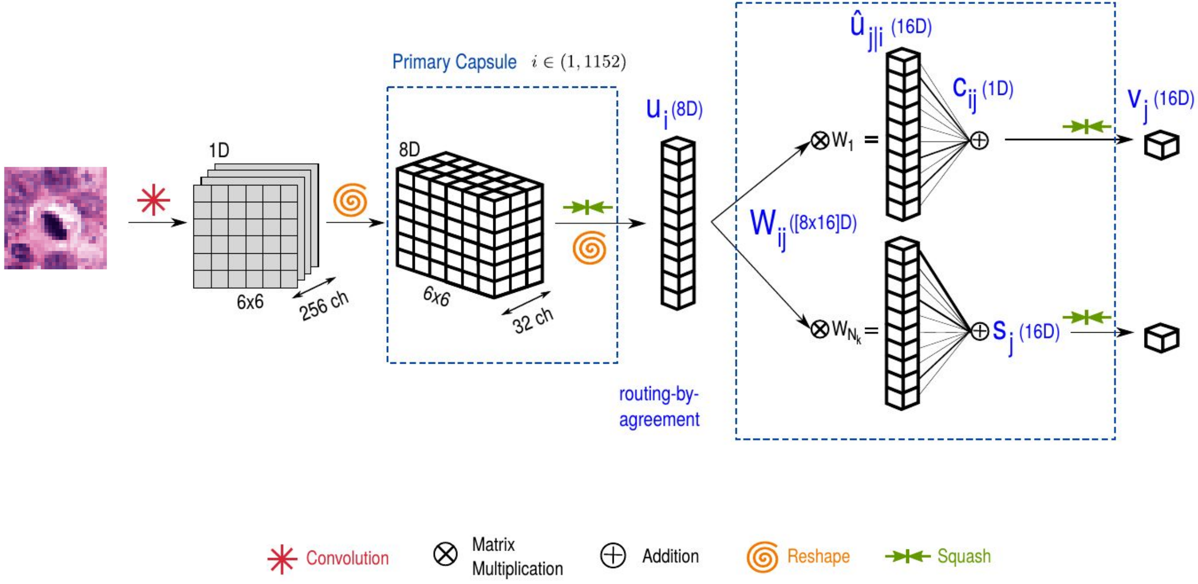
ConvNets (LeNet, Baseline)

params: 60 K, 35.4 M



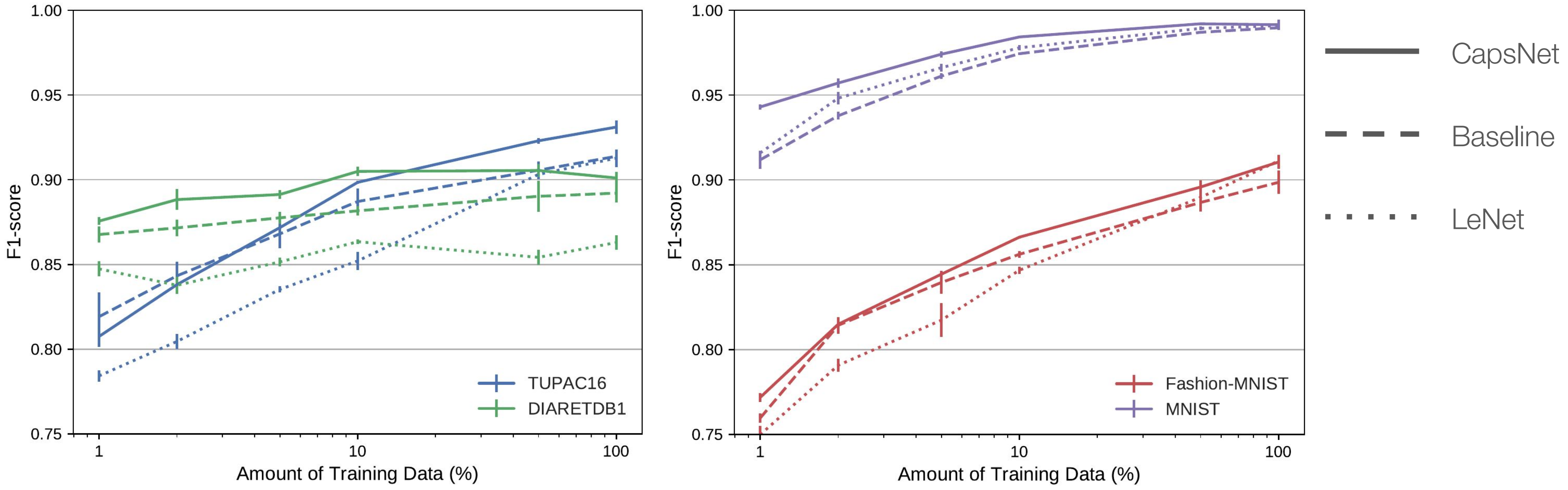
CapsNet

params: 8.2 M

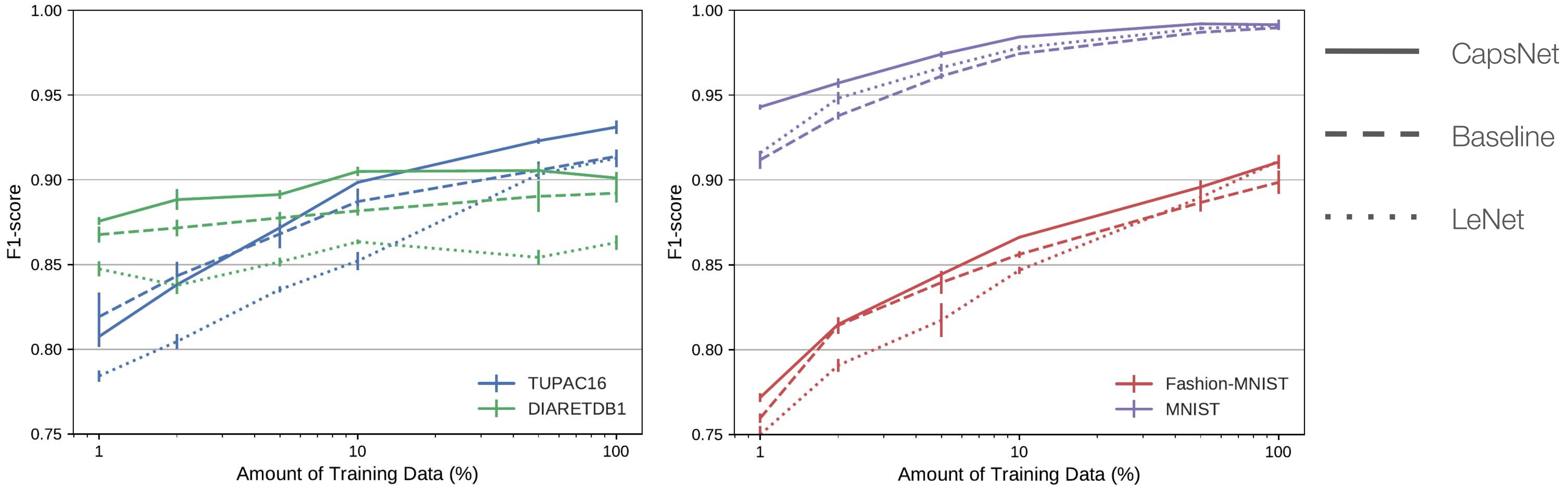


(1) How do networks behave under decreasing **amounts of training data**?

(1) How do networks behave under decreasing **amounts of training data**?

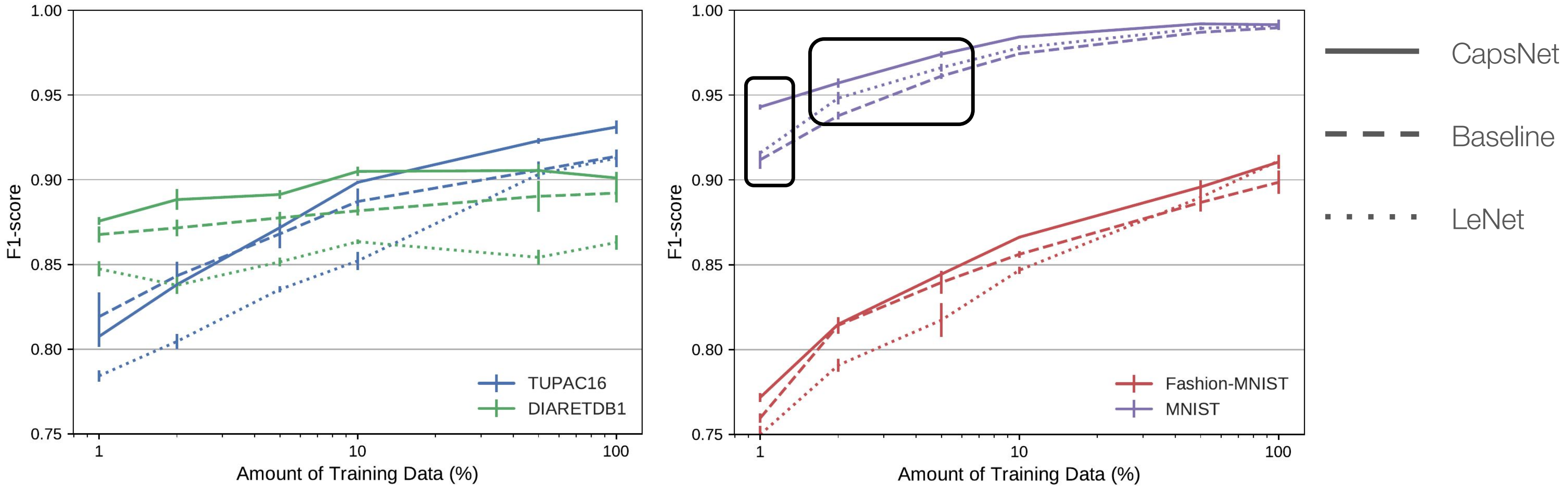


(1) How do networks behave under decreasing **amounts of training data**?



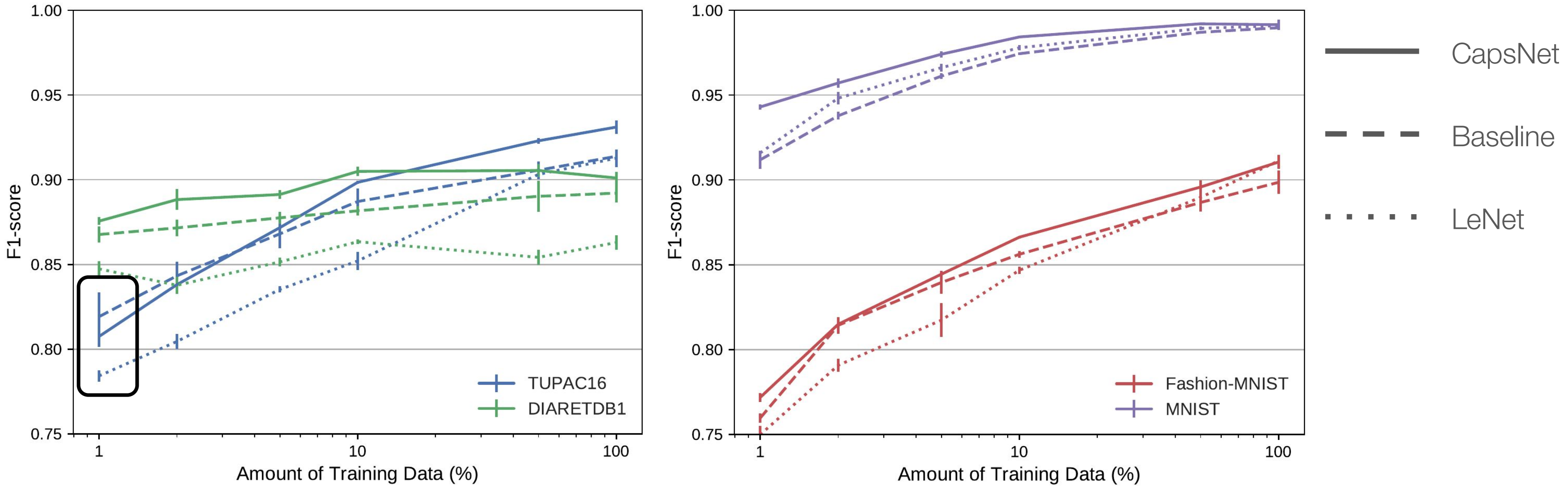
- CapsNet performs overall better than ConvNets (LeNet & Baseline).

(1) How do networks behave under decreasing **amounts of training data**?



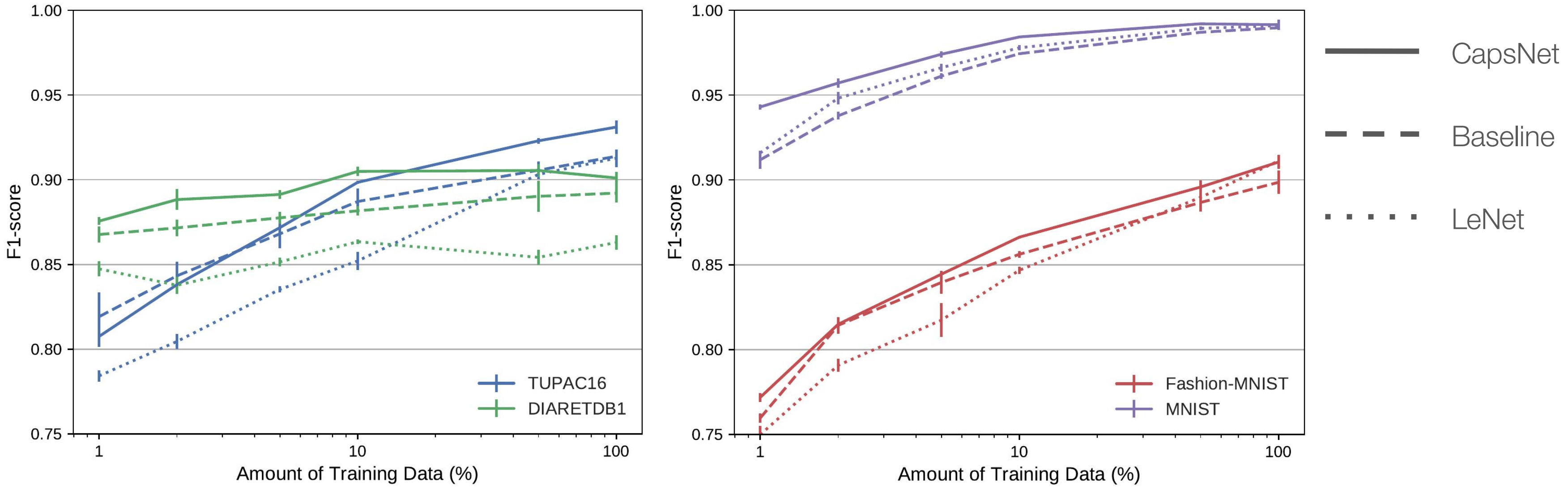
- CapsNet performs overall better than ConvNets (LeNet & Baseline).
- The gap is higher for small amount of data (MNIST).

(1) How do networks behave under decreasing **amounts of training data**?



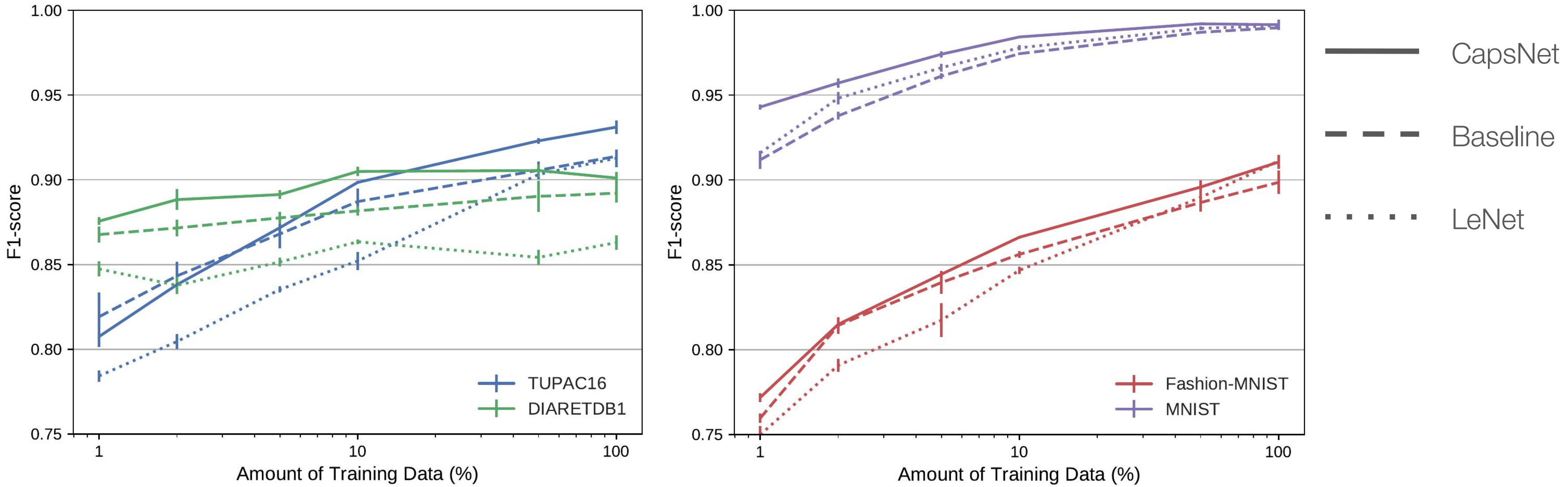
- CapsNet performs overall better than ConvNets (LeNet & Baseline).
- The gap is higher for small amount of data (MNIST).
- Improvement is limited in more complex dataset (TUPAC16).

(1) How do networks behave under decreasing **amounts of training data**?



- CapsNet performs overall better than ConvNets (LeNet & Baseline).
- The gap is higher for small amount of data (MNIST).
- Improvement is <u>limited</u> in more complex dataset (TUPAC16).
- All our experiments validated the significance test with a p-value < 0.05 (except for TUPAC16).

(1) How do networks behave under decreasing **amounts of training data**?



<u>Take home messages</u>:

- CapsNet requires **less images** for a better performance.
- Behaviour can change for different datasets.

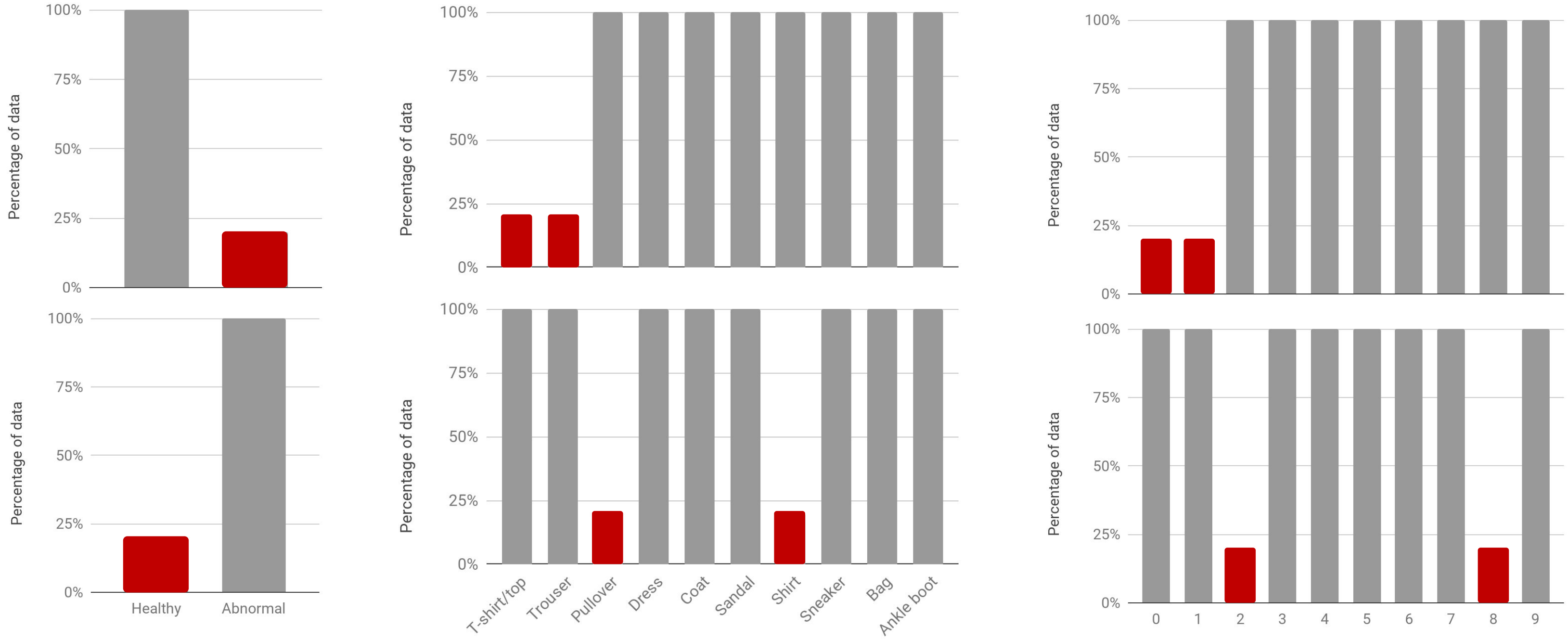
(2) Is there a change in their response to **class-imbalance**?

(2) Is there a change in their response to **class-imbalance**?

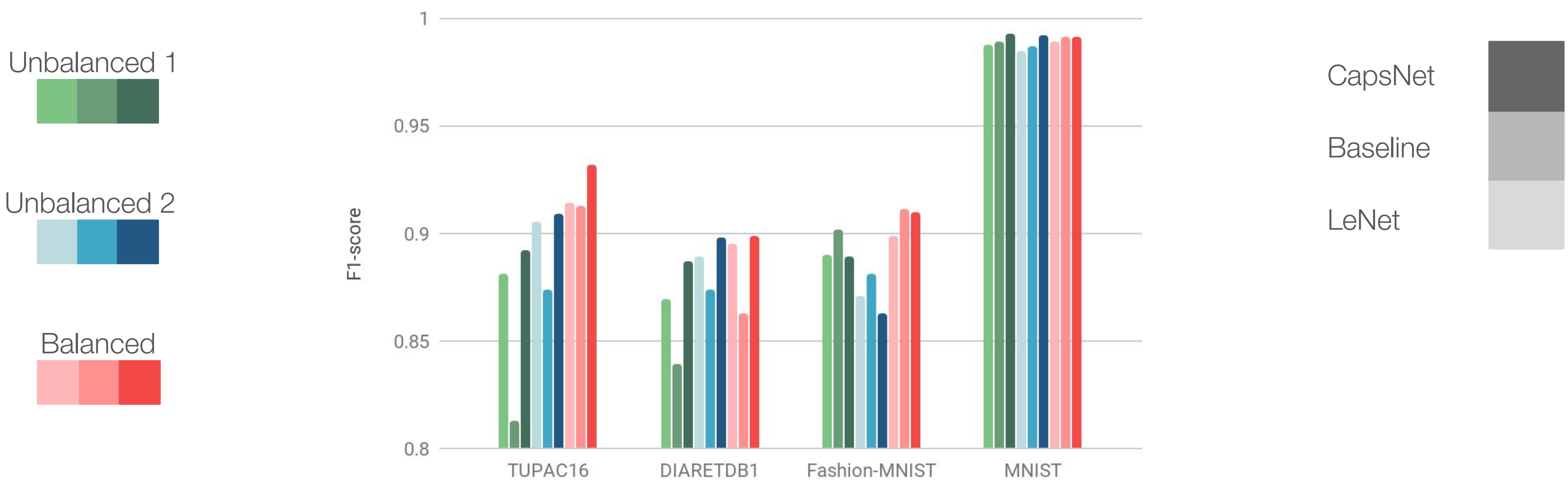
TUPAC16 & DIARETDB1 | Fashion-MNIST | MNIST

Unbalanced 1

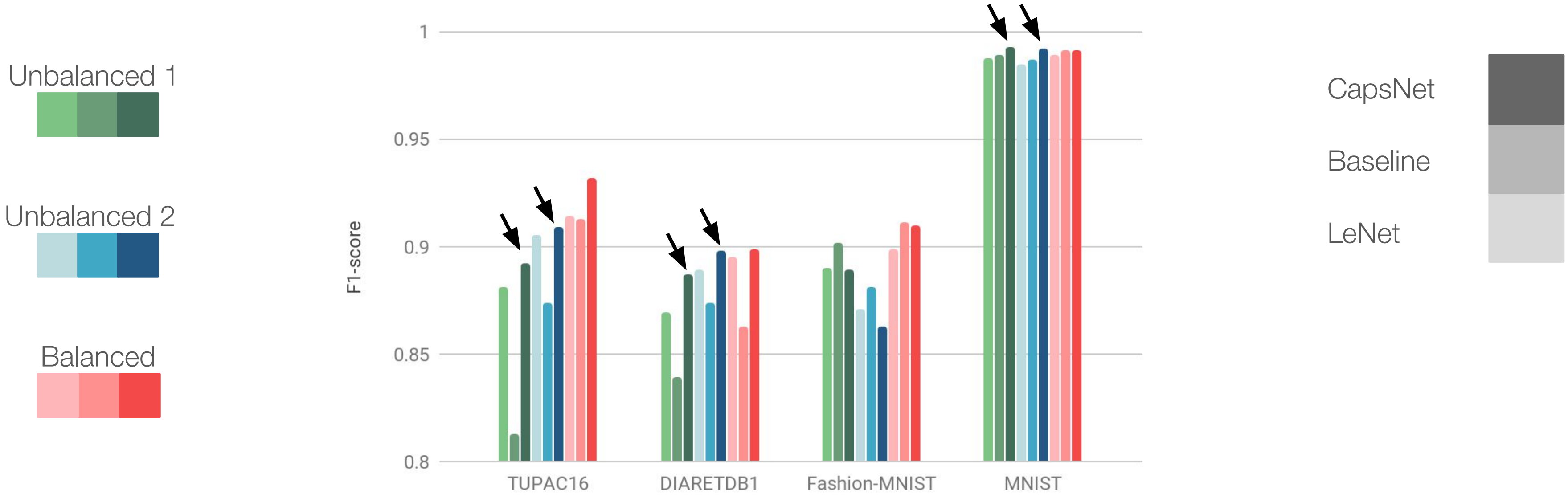
Unbalanced 2



(2) Is there a change in their response to **class-imbalance**?

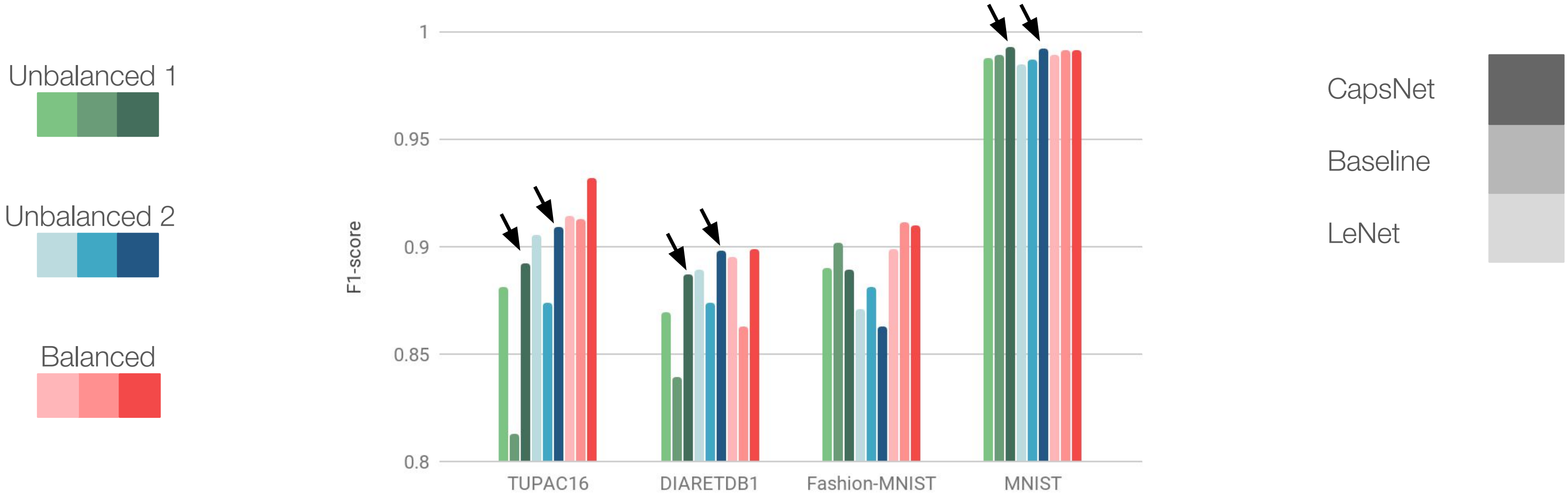


(2) Is there a change in their response to **class-imbalance**?



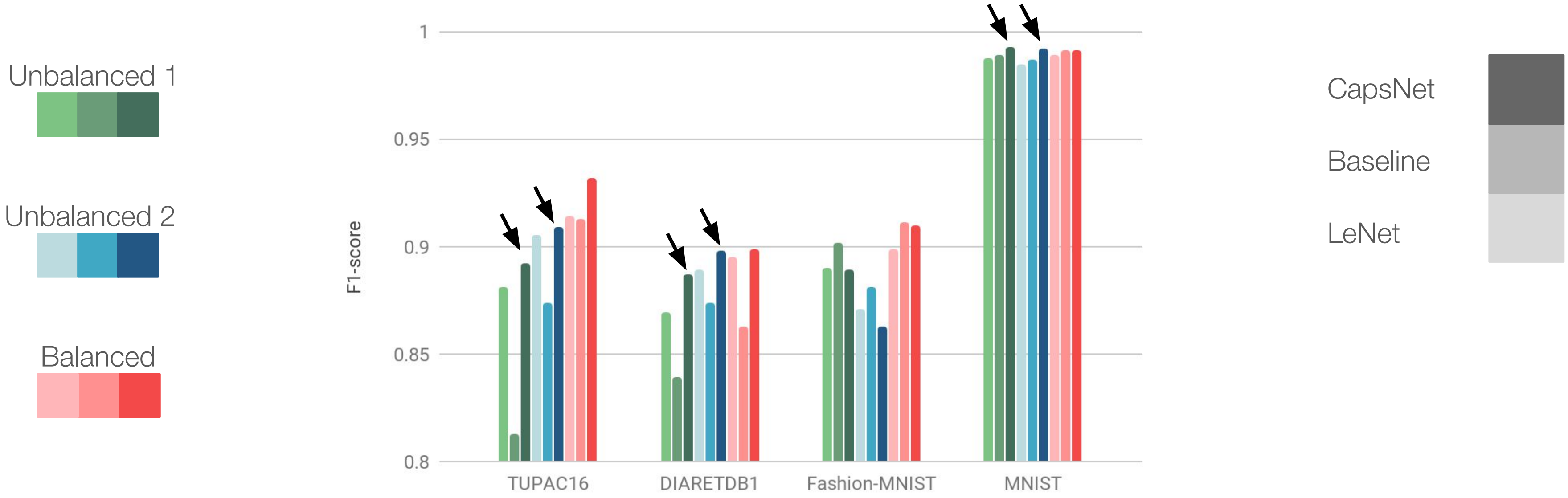
- CapsNet surpasses performance of ConvNets for all cases, except for Fashion-MNIST.

(2) Is there a change in their response to **class-imbalance**?



- CapsNet surpasses performance of ConvNets for all cases, except for Fashion-MNIST.
- At least one of the unbalanced cases verified the significance test for all datasets.

(2) Is there a change in their response to **class-imbalance**?

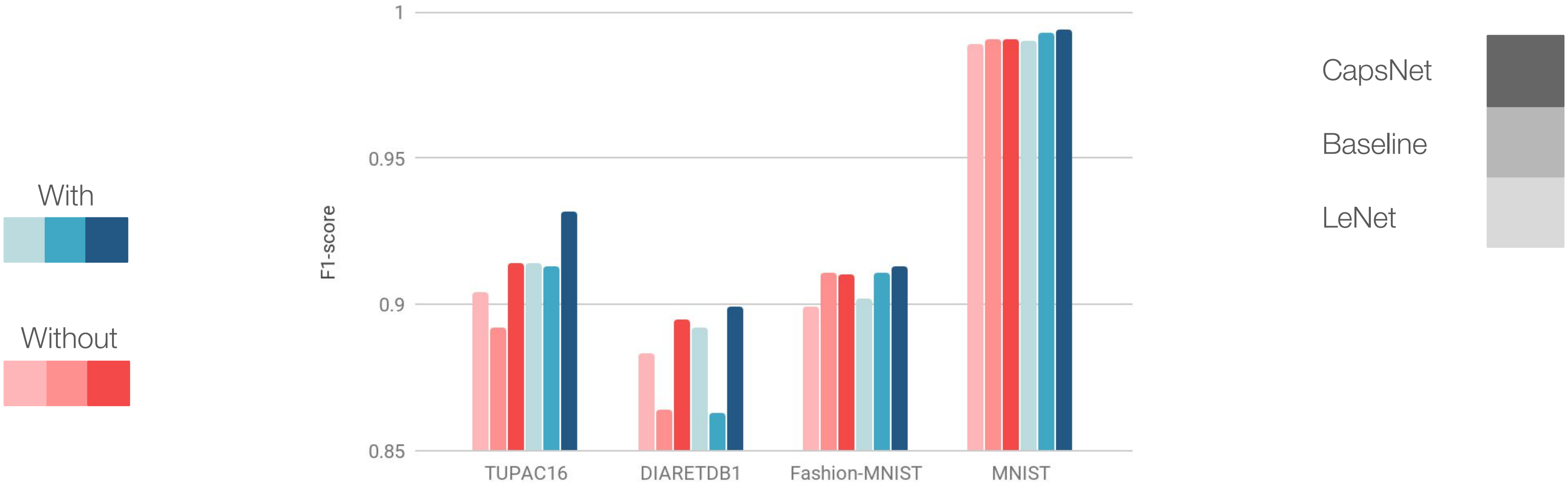


<u>**Take home message**</u>:

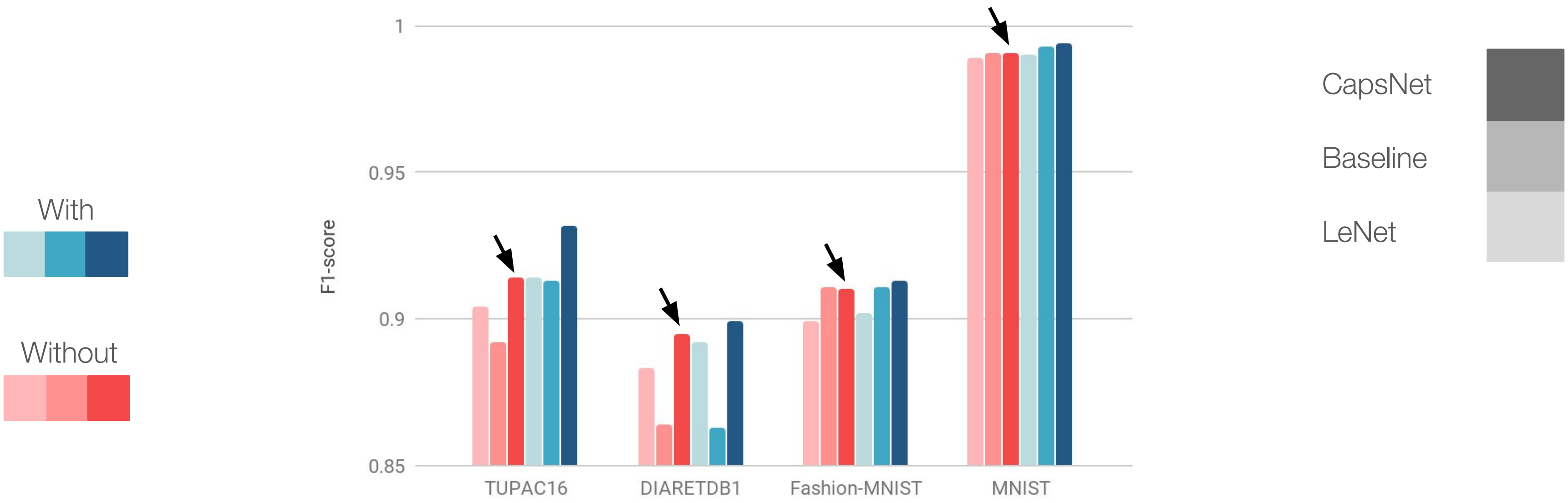
- CapsNet is **more robust** to imbalance in the class distribution.

(3) Is there any benefit from **data augmentation** as a complementary strategy?

(3) Is there any benefit from **data augmentation** as a complementary strategy?

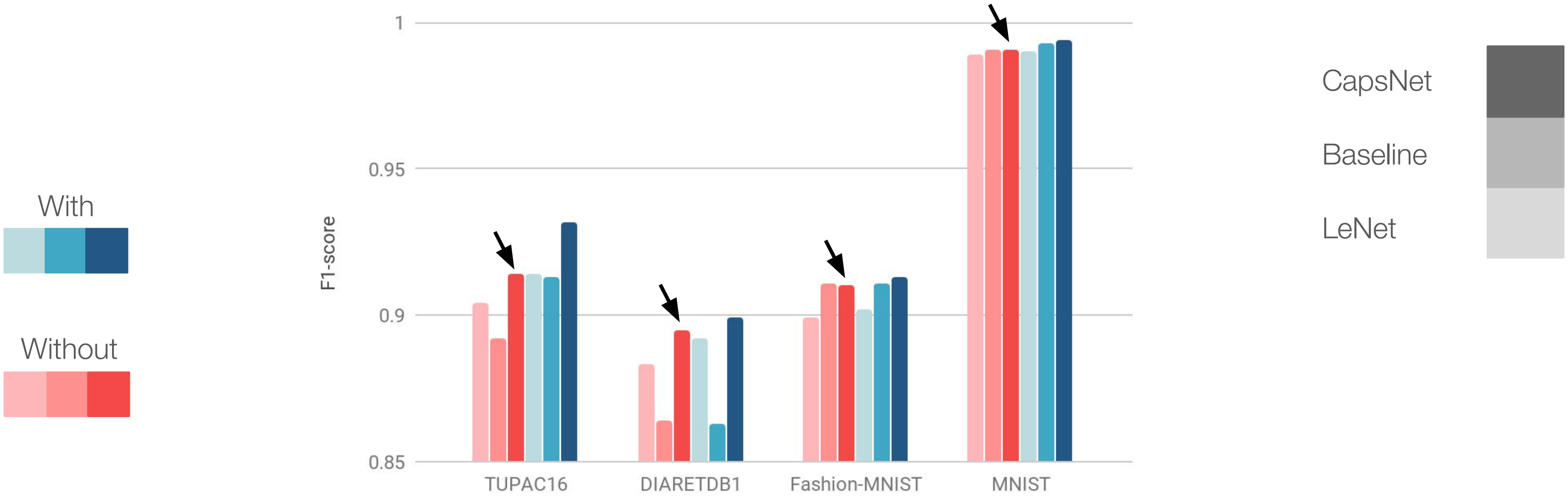


(3) Is there any benefit from **data augmentation** as a complementary strategy?



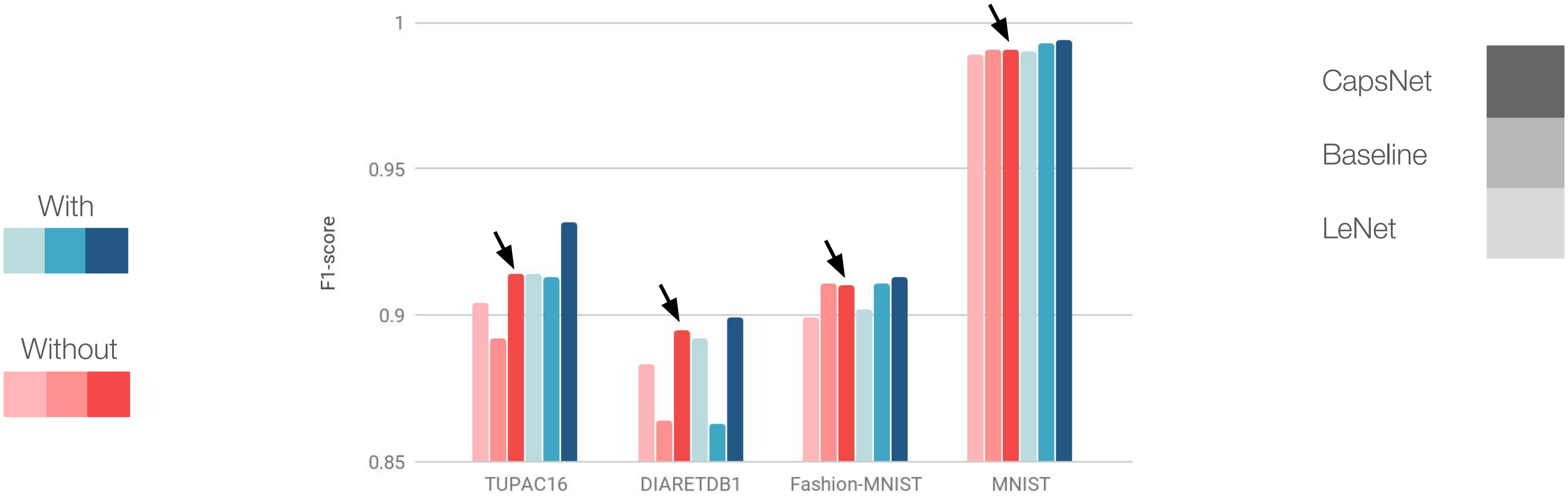
- CapsNet without data augmentation performs … than ConvNets using data augmentation.
  - similarly (TUPAC16, MNIST, Fashion)
  - better (DIARETDB1)

(3) Is there any benefit from **data augmentation** as a complementary strategy?



- CapsNet without data augmentation performs … than ConvNets using data augmentation.
  - similarly (TUPAC16, MNIST, Fashion)
  - better (DIARETDB1)
- All results were found significant.

(3) Is there any benefit from **data augmentation** as a complementary strategy?



Take home message:

- CapsNet learns **a stronger representation** with less variability of the data.

Conclusion

+ **Equivariance** modeling, requires to see fewer viewpoints of the instance of interest.
+ Allows to reduce the **number of parameters** for a comparable performance.
+ CapsNet improves CADx classification **performance** under medical data challenges.

- Routing-by-agreement is **slower** than backpropagation (≈ convergence time).
- Improvement is **limited** in more complex datasets (TUPAC16).
- **Reconstructions** are blurry for medical datasets with complex backgrounds.

Outlook

- ➥ Fully convolutional **decoder** to handle complex backgrounds.
- ➥ Explore CapsNets in a **semi-supervised** or **unsupervised** framework.
- ➥ Investigate the latent space to improve **explainability** and **interpretability**.
- ➥ Look into more suitable **medical datasets**, in which neighborhood structure plays a role for diagnosis.

ACKNOWLEDGEMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 713673.

Amelia Jiménez-Sánchez has received financial support through the ''la Caixa'' INPhINIT Fellowship Grant for Doctoral studies at Spanish Research Centres of Excellence, ''la Caixa'' Banking Foundation, Barcelona, Spain.

Thank you for your attention!

Amelia Jiménez-Sánchez
amelia.jimenez@upf.edu