Deep learning in the ultrasound evaluation of neonatal respiratory status

Michela Gravina[*], Diego Gragnaniello[*], Luisa Verdoliva[†], Giovanni Poggi[*], Iuri Corsini[§], Carlo Dani[§],Fabio Meneghin[¶], Gianluca Lista[¶], Salvatore Aversa[‖], Francesco Raimondi[‡], Fiorella Migliaro[‡] and Carlo Sansone[*]

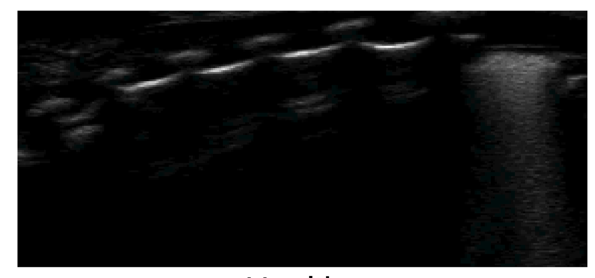
[*]Dept. of Electrical Eng. and Information Technology, University of Naples Federico II, Naples, Italy - [†]Dept. of Industrial Engineering, University of Naples Federico II, Naples, Italy

[‡]Division of Neonatology, Section of Pediatrics, Dept. of Translational Medical Sciences,University of Naples Federico II, Naples, Italy

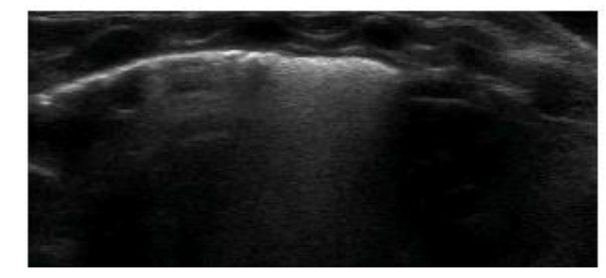
[§]Hospital Careggi of Florence, Italy - [¶]Hospital "Vittore Buzzi", Milan, Italy - [‖]Civil Hospital of Brescia, Italy

1. Lung Ultrasound Advantages

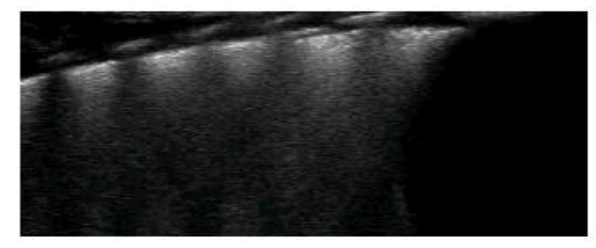
- Lung Ultrasound (US) for neonatal respiratory status evaluation:
  - ✓ Exhibits a high descriptive power
  - ✓ Is able to analyze anatomical structures, like pleural line and artifacts, exploited by experts to asses the lung composition and status
  - ✓ Provides a visual assessment of the respiratory status of newborn patients
  - ✓ Reveals the need for respiratory support
  - ✓ Is a harmless technique, compared to the standard X-rays (RX)
- **US images are prone to many investigations through either visual inspection or computer-aided analysis**
- The proposed work:
  - ➢ Exploits Deep Learning (DL) approaches to define a score directly obtained from US images
  - ➢ Test its correlation with pulse oximetric saturation ratio SpO_2/FiO_2 (SF)
  - ➢ Analyzes the performance of recent state-of-art Convolutional Neural Networks (CNNs)



Healthy



Transient Tachypnea of the Newborn (TTN)

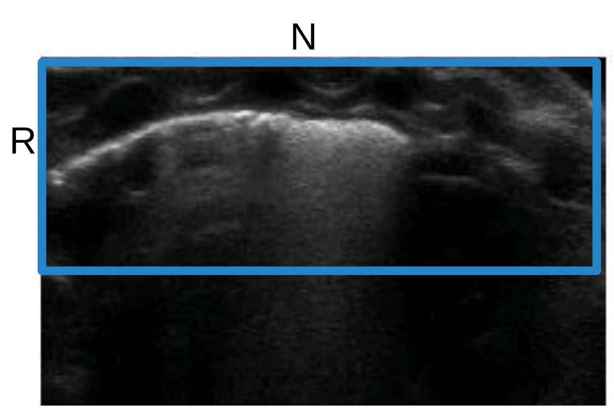


Respiratory Distress Syndrome (RDS)

2. Methodology

Pre-Processing

- Images are resized to have the same horizontal resolution (N) and the first R rows are retained.
- The obtained rectangular-shaped (RxN):
  - ✓ always includes the pleural line and part of the lung below it
  - ✓ allows the network to process the whole US from left to right without discarding relevant regions
  - ✓ does not include the bottom region of the US images, which carries no useful information



Respiratory Status Assessment

SF value prediction

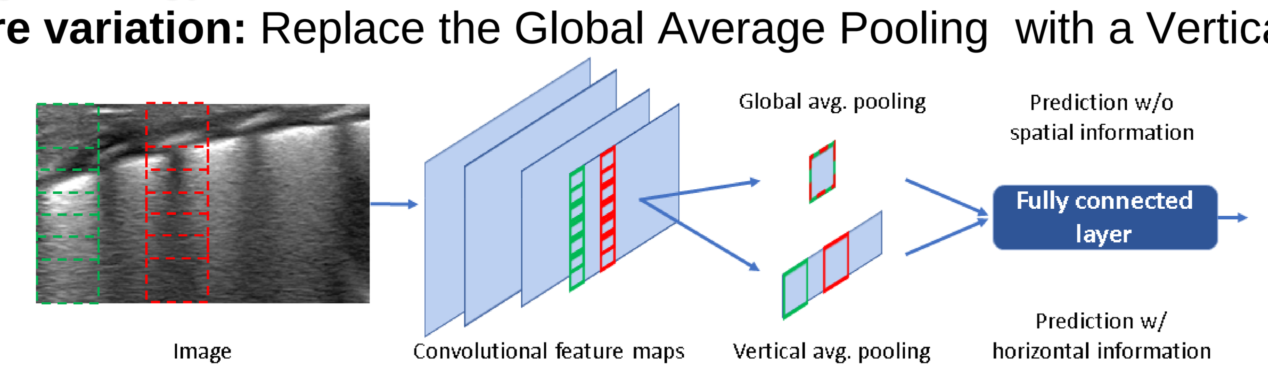
- Train the CNN to directly predict the SR value with regression loss

Healthy/Sick classification

- Train the CNN with class labels (i.e healthy or sick) that will be predicted with a certain confidence:
  - ➢ The confidence can be regarded as class probability

Advanced Training Strategy

- **CNN architecture variation:** Replace the Global Average Pooling with a Vertical Average Pooling



- **Curriculum Learning:** The training procedure consists of two phases:
  - ➢ Train the CNN with **easy** example
  - ➢ Add **hard** or **borderline** samples in a second step

3. Experimental Setup

Performance evaluation

SF value prediction

- ➢ Mean absolute percentage of Error (MAPE)
- ➢ Spearman's rank correlation between predicted SF and SF value

Healthy/Sick classification

- ➢ Accuracy (ACC)
- ➢ Spearman's rank correlation between prediction probability and SF value

Involved CNN:

- AlexNet[1]
- ResNet34[2]
- EfficientNet-B0[3]
- EfficientNet-B1[3]
- EfficientNet-B2[3]

Dataset

Disease	Number of patient	Number of videos	Patients per center Naples	Milan	Florence
None	43	328	23	10	10
RDS	30	727	11	8	11
TTN	13	211	13	-	-
Tot.	86	1266	47	18	21

4. SF Value Prediction – Results

REGRESSION RESULTS OF THE CONSIDERED BASELINES WITH DIFFERENT AUGMENTATIONS. FROM TOP TO BOTTOM WE REPORT THE RESULTS WITHOUT ANY AUGMENTATION AND BY ADDING THE FOLLOWING OPERATIONS, CONSECUTIVELY: RANDOM HORIZONTAL FLIP, RANDOM ROTATION IN THE RANGE $[-10°, +10°]$, RANDOM BRIGHTNESS AND OR CONTRAST VARIATION IN THE RANGE $[-25\%, +25\%]$.

Network	Input size	Augmentation	Correlation frame	video	session	MAPE frame	video	session
AlexNet	224×461	None	0.6678	0.6857	0.7371	0.1472	0.1430	0.1368
ResNet34	224×461	None	0.6291	0.6812	0.7398	0.1614	0.1465	0.1339
EfficientNet-B0	224×461	None	0.6383	0.6739	0.7357	0.1521	0.1399	0.1276
EfficientNet-B1	240×461	None	0.6735	0.7024	0.7487	0.1428	0.1345	**0.1250**
EfficientNet-B2	260×461	None	0.6754	0.7033	0.7630	0.1474	0.1402	0.1298
AlexNet	224×461	hor. flip	0.6587	0.6772	0.7263	0.1474	0.1454	0.1420
ResNet34	224×461	hor. flip	0.6504	0.6969	0.7556	0.1453	0.1414	0.1355
EfficientNet-B0	224×461	hor. flip	0.5857	0.6119	0.6653	0.1653	0.1609	0.1552
EfficientNet-B1	240×461	hor. flip	0.6645	**0.6998**	0.7637	**0.1377**	**0.1338**	0.1282
EfficientNet-B2	260×461	hor. flip	0.6620	0.6905	0.7576	0.1424	0.1389	0.1317
AlexNet	224×461	hor. flip, ±10° rot.	0.6644	0.6804	0.7328	0.1506	0.1492	0.1442
ResNet34	224×461	hor. flip, ±10° rot.	0.6621	0.6963	**0.7662**	0.1428	0.1389	0.1329
EfficientNet-B0	224×461	hor. flip, ±10° rot.	0.5994	0.6349	0.7070	0.1669	0.1620	0.1557
EfficientNet-B1	240×461	hor. flip, ±10° rot.	0.6642	0.6935	0.7534	0.1448	0.1419	0.1359
EfficientNet-B2	260×461	hor. flip, ±10° rot.	0.6661	0.6953	0.7592	0.1441	0.1408	0.1326
AlexNet	224×461	hor. flip, ±10° rot., bri./cont. adj	**0.6695**	0.6889	0.7442	0.1498	0.1481	0.1447
ResNet34	224×461	hor. flip, ±10° rot., bri./cont. adj	0.6649	0.6959	0.7623	0.1391	0.1349	0.1282
EfficientNet-B0	224×461	hor. flip, ±10° rot., bri./cont. adj	0.6067	0.6471	0.7176	0.1668	0.1619	0.1555
EfficientNet-B1	240×461	hor. flip, ±10° rot., bri./cont. adj	0.6563	0.6847	0.7416	0.1457	0.1429	0.1374
EfficientNet-B2	260×461	hor. flip, ±10° rot., bri./cont. adj	0.6638	0.6916	0.7577	0.1445	0.1415	0.1338

5.Healthy/Sick classification– Results

BINARY CLASSIFICATION RESULTS OF THE CONSIDERED BASELINES WITH DIFFERENT AUGMENTATIONS. FROM TOP TO BOTTOM WE REPORT THE RESULTS WITHOUT ANY AUGMENTATION AND BY ADDING THE FOLLOWING OPERATIONS, CONSECUTIVELY: RANDOM HORIZONTAL FLIP, RANDOM ROTATION IN THE RANGE $[-10°, +10°]$, RANDOM BRIGHTNESS AND OR CONTRAST VARIATION IN THE RANGE $[-25\%, +25\%]$.

Network	Input size	Augmentation	Correlation frame	video	session	Accuracy frame	video	session
AlexNet	224×461	None	0.6586	0.6749	0.7518	0.8121	0.8245	0.8461
ResNet34	224×461	None	0.6507	0.6830	0.7562	0.8129	0.8229	0.8397
EfficientNet-B0	224×461	None	0.6536	0.6794	0.7463	0.8159	0.8316	0.8461
EfficientNet-B1	240×461	None	0.6647	0.6928	0.7578	0.8214	0.8332	0.8653
EfficientNet-B2	260×461	None	0.6471	0.6687	0.7353	0.8166	0.8340	0.8654
AlexNet	224×461	hor. flip	0.6593	0.6729	0.7458	0.8054	0.8166	0.8462
ResNet34	224×461	hor. flip	0.6414	0.6701	0.7526	0.8096	0.8182	0.8397
EfficientNet-B0	224×461	hor. flip	0.6425	0.6641	0.7351	0.8141	0.8190	0.8397
EfficientNet-B1	240×461	hor. flip	0.6461	0.6752	0.7533	0.8194	0.8356	0.8718
EfficientNet-B2	260×461	hor. flip	0.6390	0.6690	0.7431	0.8152	0.8237	0.8397
AlexNet	224×461	hor. flip, ±10° rot.	0.6485	0.6658	0.7363	0.8124	0.8229	0.8526
ResNet34	224×461	hor. flip, ±10° rot.	0.6602	0.6869	0.7546	0.8248	**0.8387**	**0.8782**
EfficientNet-B0	224×461	hor. flip, ±10° rot.	0.6696	0.6883	0.7543	**0.8292**	0.8364	0.8526
EfficientNet-B1	240×461	hor. flip, ±10° rot.	0.6539	0.6794	0.7540	0.8156	0.8356	0.8462
EfficientNet-B2	260×461	hor. flip, ±10° rot.	0.6378	0.6634	0.7264	0.8123	0.8245	0.8590
AlexNet	224×461	hor. flip, ±10° rot., bri./cont. adj	0.6557	0.6690	0.7310	0.8061	0.8206	0.8590
ResNet34	224×461	hor. flip, ±10° rot., bri./cont. adj	**0.6777**	**0.7030**	**0.7782**	0.8208	**0.8387**	0.8526
EfficientNet-B0	224×461	hor. flip, ±10° rot., bri./cont. adj	0.6581	0.6819	0.7586	0.8249	0.8379	0.8590
EfficientNet-B1	240×461	hor. flip, ±10° rot., bri./cont. adj	0.6592	0.6828	0.7564	0.8138	0.8340	0.8590
EfficientNet-B2	260×461	hor. flip, ±10° rot., bri./cont. adj	0.6620	0.6846	0.7501	0.8104	0.8316	0.8718

6. Advanced Training Strategies – Results

RESULTS OF THE BEST PERFORMING NETWORKS WITH CURRICULUM LEARNING AND HORIZONTAL POSITION INFORMATION PRESERVING.

Train mode	Network	Advanced training	Correlation frame	video	session	MAPE or Accuracy frame	video	session
Regression	ResNet34	hor. flip position	0.6734	0.6778	0.7440	0.1681	**0.1427**	**0.1344**
	EfficientNet-B1	hor. flip position	0.6361	0.6754	0.7505	**0.1511**	0.1475	0.1395
Classification	ResNet34	hor. flip position	0.6766	**0.7051**	**0.7821**	0.8179	0.8324	**0.8718**
	EfficientNet-B1	hor. flip position	0.6459	0.6747	0.7608	0.8092	0.8261	0.8333
	ResNet34	Curriculum Learning	**0.6780**	0.6995	0.7604	**0.8195**	0.8348	0.8654
	EfficientNet-B1	Curriculum Learning	0.6645	0.6852	0.7600	0.8161	**0.8616**	0.8590

Correlation between **LUS score**[4] **assigned by humans expert** and SF value: **0.8259**

7. Conclusion

- Results show that ResNet34 trained for binary classification achieves the best performance in terms of correlation with the selected reference marker
- The correlation further improves by modifying the CNN architecture in order to take into account the horizontal position of the extracted convolutional networks
- It is worth observing that the proposed approach performs comparably with the human operator (+4.4%)

8. Future works

- Enlarge the dataset including data from other medical centers
- Improve the training strategy by exploiting the temporal information of the lung US videos

9. References

[1]A. Krizhevsky et al, "Imagenet classification with deep convolutional neural networks"

[2]K. He et al. "Deep residual learning for image recognition"

[3]M. Tan et al. "Efficientnet: Rethinking model scaling for convolutional neural networks"

[4]F. Raimondi et al. "Visual assessment versus computer-assisted gray scale analysis in the ultrasound evaluation of neonatal respiratory status," PLOS ONE