Oleksii Hyka
Research & Development Centre Netrix S.A
Oleksii.hyka@netrix.com.pl
ORCID: 0000-0003-2129-5138
Jan Kowalski
Lubelska Akademia WSEI
Jan.kowalski@wsei,lublin.pl
ORCID: 0000-0000-000-000
Jan Kowalski
Lubelska Akademia WSEI
Jan.kowalski@wsei,lublin.pl
ORCID: 0000-0000-000-000

System for cough monitoring
System do odsłuchu kaszlu

Summary

Our research investigates the application of machine learning and Internet of Things (IoT) technologies in healthcare, focusing on the detection and classification of coughing episodes. Leveraging deep learning architectures and a comprehensive IoT infrastructure, we developed an automated system capable of monitoring audio signals from a microphone array module to accurately detect coughs and classify their types. The study utilized the COUGHVID dataset for model training and evaluation, employing rigorous preprocessing techniques to ensure data integrity. Through comparative analysis, we identified MobileNet as the optimal model for cough detection, achieving promising results in terms of accuracy, area under the ROC curve (AUC), and F1 score. Furthermore, our emphasis on privacy safeguards and remote medical examination facilitation underscores the practical implications of our research in enhancing healthcare delivery. Overall, our study contributes to the advancement of technology-enabled healthcare solutions, offering valuable insights and solutions for improving patient care and outcomes.

Streszczenie

Nasze badania dotyczą zastosowania technologii uczenia maszynowego i Internetu rzeczy (IoT) w opiece zdrowotnej, skupiając się na wykrywaniu i klasyfikacji epizodów kaszlu. Wykorzystując architektury głębokiego uczenia i kompleksową infrastrukturę IoT, opracowaliśmy zautomatyzowany system zdolny do monitorowania sygnałów dźwiękowych z modułu mikrofonowego w celu dokładnego wykrywania kaszlu i klasyfikacji jego rodzajów. W badaniu wykorzystano zbiór danych COUGHVID do szkolenia i oceny modeli, stosując rygorystyczne techniki wstępnego przetwarzania w celu zapewnienia integralności danych. Poprzez analizę porównawczą zidentyfikowaliśmy MobileNet jako optymalny model do wykrywania kaszlu, osiągając obiecujące wyniki pod względem dokładności, obszaru pod krzywą ROC (AUC) i wskaźnika F1. Ponadto, nasze naciski na zabezpieczenia prywatności i ułatwienie zdalnych badań medycznych podkreślają praktyczne implikacje naszych badań w poprawie dostarczania opieki zdrowotnej. Ogólnie nasze badanie przyczynia się do rozwoju technologicznych rozwiązań w dziedzinie opieki zdrowotnej, oferując cenne spostrzeżenia i rozwiązania dla poprawy opieki i wyników dla pacjentów.

Keywords: *Machine Learning, MQTT, Cough Detection, Healthcare, Deep Learning, Spectrogram Analysis*

Słowa kluczowe*: Uczenie maszynowe, MQTT, Wykrywanie kaszlu, Opieka zdrowotna, Głębokie uczenie, Analiza spektrogramu, Prywatność danych*

Introduction

In recent years, the widespread integration of smart devices has reshaped daily routines, simplifying tasks ranging from household chores to environmental control through voice-activated mechanisms and other innovative features. This technological evolution now extends into the domain of healthcare, where the implementation of machine learning methodologies is revolutionizing both home-based(Shin 2020; Takahashi et al. 2004) and institutional medical practices. Of particular interest is the application of these techniques towards the detection and classification of coughing episodes (Irwin et al., 2018; Pahar et al., 2021) , given the profound impact of coughing, particularly chronic coughing, on individuals' quality of life. Thus, accurate assessment and evaluation of cough severity are paramount in clinical settings, necessitating a multifaceted approach encompassing various parameters such as frequency, duration, energy, intensity, and the resultant impact on overall well-being.

Our research endeavors focus on the development of an automated system designed to monitor audio signals captured by a microphone array module, with a primary objective of accurately detecting coughs and discerning between dry and wet cough types. A critical consideration in our system design is the stringent maintenance of privacy safeguards (Dragomir et al. 2016; Nguyen, Laurent, and Oualha 2015), achieved through the execution of all data processing and prediction tasks within the confines of the embedded system, thereby preempting the risk of data breaches. Initially conceived for institutional deployment, our solution facilitates seamless interaction between healthcare professionals and patients, enabling remote medical examinations initiated via mobile applications. Subsequently, the results of cough detection procedures are made accessible to both patients and medical practitioners through a dedicated web interface.

Despite the ubiquity of speech recognition systems, their efficacy in accurately discerning non-verbal cues such as coughing remains limited, often relegating cough signals to the realm of background noise. However, coughing serves as a vital indicator of individuals' daily health status, rendering precise cough detection instrumental in the monitoring of disease progression and severity.

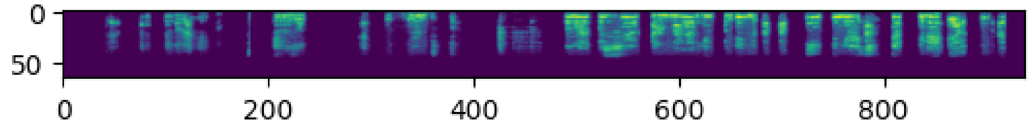
Research Methodology

The research methodology employed in this study was predicated on the utilization of the 'The COUGHVID crowdsourcing dataset: A corpus for the study of large-scale cough analysis algorithms (Bansal, Pahwa, and Kannan 2020), obtainable from (Zenodo 2021), as the foundational dataset for our investigation. Prior to model training and evaluation, the dataset underwent meticulous pre-processing procedures to ensure data integrity and uniformity. Notably, audio files shorter than 15 seconds were augmented with padding to standardize their length, while spectrograms (Bhatia 2020; Gunawan et al. 2021) with consistent dimensions were generated from the recordings. These spectrograms presented in Figure 1 served as the primary input data for training the classification models, alongside corresponding labels indicating the presence or absence of coughing events.

Subsequently, a comparative analysis of several deep learning architectures was conducted to ascertain the optimal model for classifying cough occurrences within the source recordings (Abaza et al. 2009; Balamurali et al. 2021). Specifically, the performance of Convolutional Neural Network (CNN) (Maj et al. 2022, 2023), RESNET-50, and MobileNet architectures was evaluated. The selection of these models was guided by their established efficacy in diverse computer vision tasks and their compatibility with resource-constrained devices, aligning with our objective of developing a lightweight and efficient solution for cough detection. Each model underwent rigorous testing to identify the optimal architecture and hyperparameters, thereby ensuring robust performance in the classification task.

Furthermore, to augment the dataset and enhance model stability, negative cough samples were collected from our prototype and incorporated into the training data as negative samples. This additional step served to bolster the model's resilience to false detections and refine its ability to accurately discriminate between coughing and non-coughing events, thereby enhancing the overall reliability and efficacy of cough detection.

Figure 1. Spectrogram of audio signal



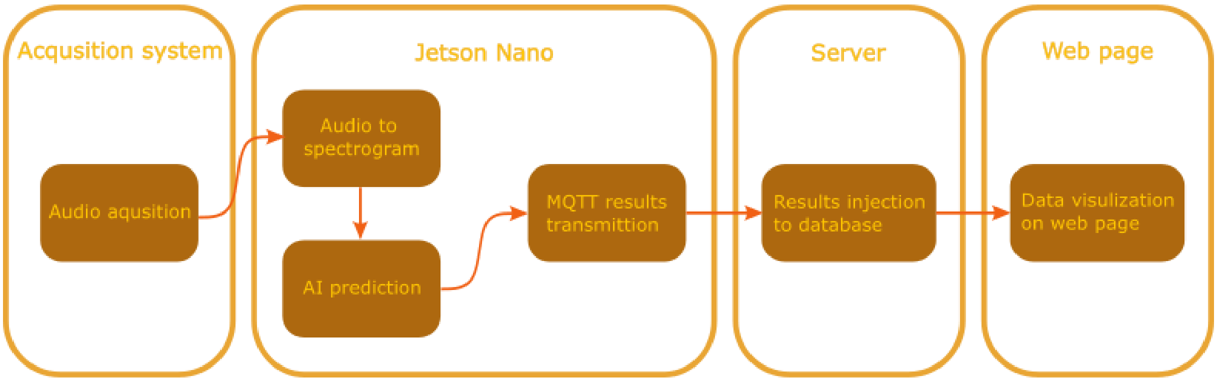
Acquisition System

The acquisition system facilitates the collection and retention of data from a microphone array. In our study, extensive deliberation was undertaken to assess various sound sampling modules, aiming to fulfill specific requirements, particularly in achieving accurate cough detection across diverse environmental conditions. Following meticulous evaluation, the ReSpeaker Mic Array v2.0 module emerged as the preferred choice. This decision was predicated on several considerations. Primarily, the module's heightened sensitivity, attributed to its implementation of four high-performance digital microphones (ST MP34DT01TR-M), endowed it with superior sound-capturing capabilities. Such enhanced sensitivity proved instrumental in facilitating precise and dependable detection of cough sounds.

Computation Unit

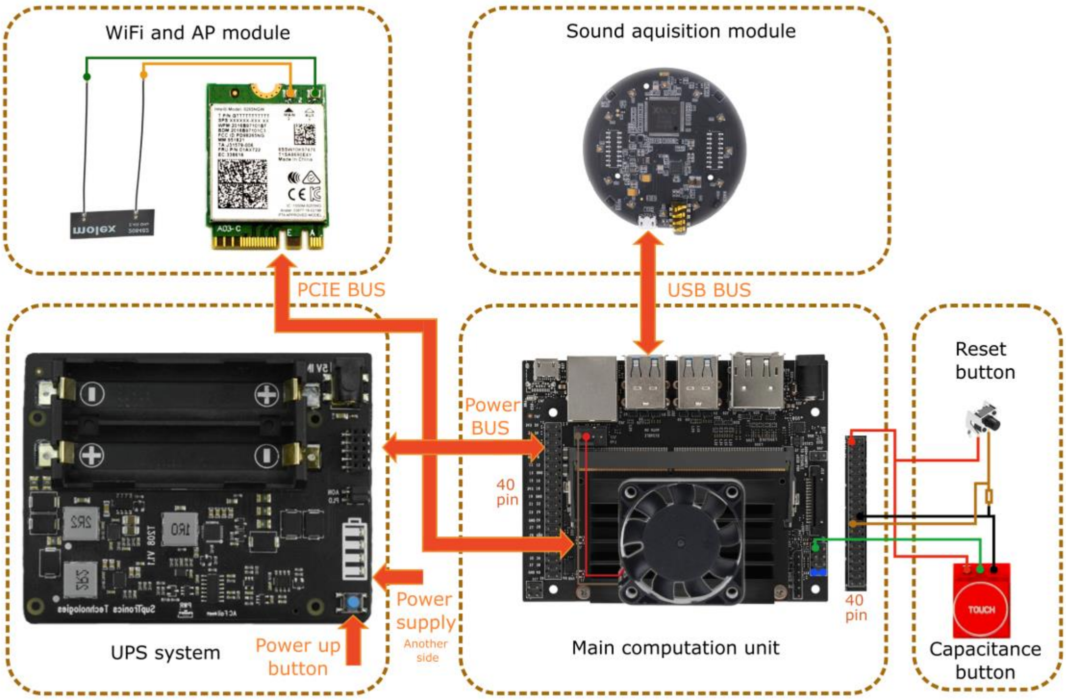
The computational framework of our system is underpinned by the Jetson Nano embedded board, distinguished by its incorporation of a dedicated GPU accelerator (Mittapalli et al. 2023). This hardware configuration serves as the cornerstone for executing intricate machine learning computations, encompassing tasks such as model training and real-time inference. Leveraging the parallel processing capabilities of the GPU, the Jetson Nano facilitates expedited execution of computationally intensive operations, thereby ensuring responsive performance even when handling voluminous datasets or complex algorithms. Furthermore, the compact form factor and energy-efficient design of the Jetson Nano (Anon n.d.; S.K, Kesanapalli, and Simmhan 2022) render it eminently suitable for deployment in embedded applications, affording scalability and operational versatility in resource-constrained environments. This integration underscores our commitment to leveraging state-of-the-art hardware solutions to achieve optimal performance and efficacy in our machine learning-driven endeavors.

Figure 2.Data Flow Overview Diagram



Presented solution allow to create access point (AP) and perform first configuration of device by connecting to the provided network. For device configuration a mobile application was developed for android and IoS systems. Then prototype automatically connect to selected wireless area network (WAN) after registration on web page and after providing device name on mobile application the prototype can be controlled and running for cough monitoring (Pahar et al. 2021; Serrurier, Neuschaefer-Rube, and Röhrig 2022; Wang et al. 2015). The collected data immediately are available for visualization on patient profile web page

Figure 3. General hardware connection diagram



Communication architecture

Efficient and secure data transmission constitutes a pivotal aspect of our system, facilitating seamless communication between devices while preserving data integrity. In our endeavor to establish a robust data transmission mechanism, we opted for the utilization of the MQTT (Message Queuing Telemetry Transport) protocol. MQTT (Singh et al. 2015) offers a lightweight, yet reliable, communication protocol suitable for the transmission of data between remote devices. Specifically, our implementation involved the establishment of a communication channel from external networks to a predefined server port, which subsequently redirected the incoming data to the Mosquitto server (Styła et al. 2021).

The adoption of the MQTT protocol was underpinned by its inherent features that align with the requirements of our system. Notably, MQTT's (Yassein et al. 2017) publish-subscribe architecture enables efficient message delivery, facilitating real-time communication between devices (Rymarczyk et al. 2020). Furthermore, its support for Quality of Service (QoS) levels ensures reliable message delivery, even in scenarios characterized by intermittent connectivity or network disruptions.

As elucidated in the seminal work by (Atmoko, Riantini, and Hasin 2017), the MQTT protocol operates over TCP/IP, offering low overhead and minimal power consumption. This data-agnostic protocol accommodates various data formats, including binary data, text, XML, or JSON, thus providing versatility in data transmission. Moreover, the widespread support for MQTT in microcontrollers, such as the STM32Fx7 series, and common market devices like Wemos and Raspberry Pi, underscores its suitability for diverse IoT applications.

Additionally, a comprehensive MQTT-based system comprises two primary software components: the MQTT Client installed on devices and the MQTT Broker responsible for handling publish and subscribe data. The former facilitates communication with the MQTT Broker, while the latter manages message delivery and ensures data integrity. Through the MQTT protocol's publish/subscribe model, data senders and receivers remain decoupled, enhancing system scalability and resilience to network fluctuations.

Results

The culmination of our study led to the development of a comprehensive prototype, as depicted in Figure 4, aimed at facilitating cough detection and visualization for both patients and healthcare providers. Evaluating the performance of the implemented models, MobileNet demonstrated promising results, achieving an accuracy level of 84%, an AUC (Area Under ROC Curve) of approximately 0.902, and an F1 score of approximately 0.846. In contrast, both MobileNet V2 and NASNet Mobile displayed moderate classification quality indicators for cough type. Notably, DenseNet121 exhibited the lowest performance in the classification task, with an accuracy level of 71%, an AUC of approximately 0.787, and an F1 score of approximately 0.757. Overall, our analysis underscores MobileNet's superiority in classifying cough types, as detailed in Table 1.

Figure 4.Prototype Overview Image

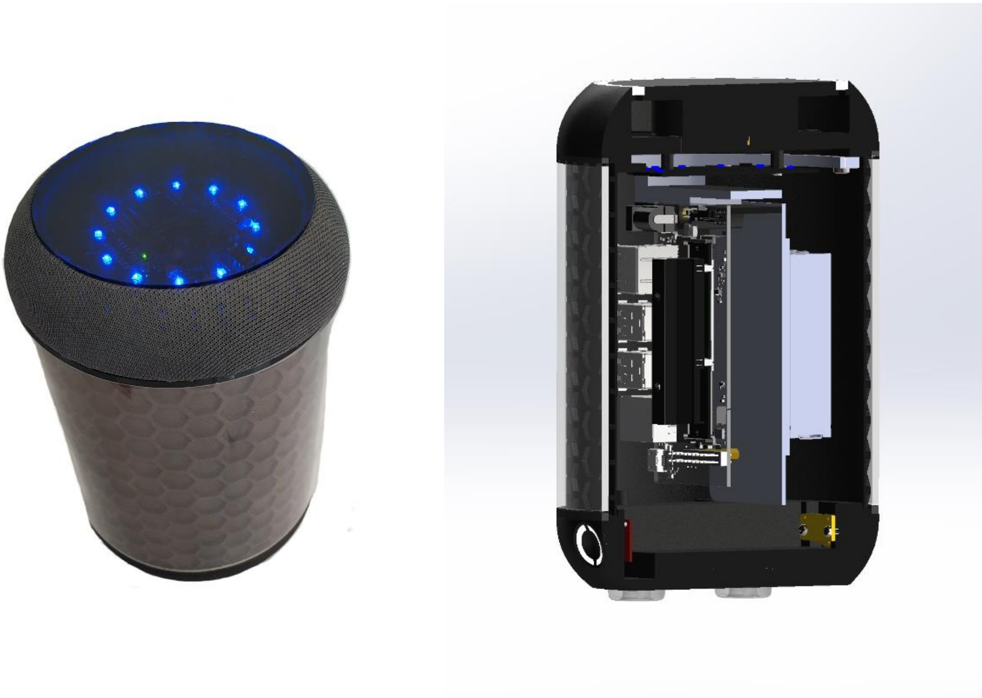


Table 1. Model Accuracy Evaluation: Comparative Table

Model	Accuracy	AUC	Precision	Recall	F1
MobileNet	84%	0.902	0.808	0.887	0.846
MobileNetV2	79%	0.856	0.747	0.883	0.809
NASNet Mobile	75%	0.851	0.695	0.887	0.779
DenseNet121	71%	0.787	0.647	0.911	0.757

Conclusions

In conclusion, our research represents a significant step forward in leveraging machine learning and IoT technologies for healthcare applications, particularly in the domain of cough detection and classification. The integration of smart devices and machine learning algorithms has the potential to revolutionize medical practices, offering remote monitoring and assessment capabilities that were previously inaccessible. By developing an automated system capable of accurately detecting coughing episodes and classifying their types, we have addressed a critical need in clinical settings.

Our study demonstrates the feasibility and effectiveness of utilizing deep learning architectures, such as MobileNet, for cough detection tasks. The promising results obtained, particularly in terms of accuracy, AUC, and F1 score, underscore the potential of these models in augmenting healthcare assessment practices. Moreover, our emphasis on privacy safeguards and remote medical examination facilitation reflects a commitment to ensuring patient confidentiality and accessibility in healthcare services.

References

Abaza, Ayman A., Jeremy B. Day, Jeffrey S. Reynolds, Ahmed M. Mahmoud, W. Travis Goldsmith, Walter G. McKinney, E. Lee Petsonk, and David G. Frazer. 2009. "Classification of Voluntary Cough Sound and Airflow Patterns for Detecting Abnormal Pulmonary Function." *Cough* 5(1):8. doi: 10.1186/1745-9974-5-8.

Anon. n.d. "Performance Evaluation of the Nvidia Jetson Nano Through a Real-Time Machine Learning Application | SpringerLink." Retrieved April 27, 2024 (https://link.springer.com/chapter/10.1007/978-3-030-68017-6_51).

Atmoko, R. A., R. Riantini, and M. K. Hasin. 2017. "IoT Real Time Data Acquisition Using MQTT Protocol." *Journal of Physics: Conference Series* 853(1):012003. doi: 10.1088/1742-6596/853/1/012003.

Balamurali, B. T., Hwan Ing Hee, Saumitra Kapoor, Oon Hoe Teoh, Sung Shin Teng, Khai Pin Lee, Dorien Herremans, and Jer Ming Chen. 2021. "Deep Neural Network-Based Respiratory Pathology Classification Using Cough Sounds." *Sensors* 21(16):5555. doi: 10.3390/s21165555.

Bansal, Vipin, Gaurav Pahwa, and Nirmal Kannan. 2020. "Cough Classification for COVID-19 Based on Audio Mfcc Features Using Convolutional Neural Networks." Pp. 604–8 in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*.

Bhatia, Nikhil. 2020. "Using Transfer Learning, Spectrogram Audio Classification, and MIT App Inventor to Facilitate Machine Learning Understanding." Thesis, Massachusetts Institute of Technology.

Gunawan, Kevin, Alam Ahmad Hidayat, Tjeng Wawan Cenggoro, and Bens Pardamean. 2021. "A Transfer Learning Strategy for Owl Sound Classification by Using Image Classification Model with Audio Spectrogram." *International Journal on Electrical Engineering and Informatics* 17:546. doi: 10.15676/ijeei.2021.13.3.3.

Maj, Michał, Tomasz Rymarczyk, Tomasz Cieplak, and Damian Pliszczuk. 2022. "Deep Learning Model Optimization for Faster Inference Using Multi-Task Learning for Embedded Systems." Pp. 892–93 in *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking, MobiCom '22*. New York, NY, USA: Association for Computing Machinery.

Maj, Michał, Tomasz Rymarczyk, Łukasz Maciura, Tomasz Cieplak, and Damian Pliszczuk. 2023. "Cross-Modal Perception for Customer Service." Pp. 1–3 in *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*. New York, NY, USA: Association for Computing Machinery.

Mittapalli, Pardha Saradhi, M. R. N. Tagore, Pulagam Ammi Reddy, Giri Babu Kande, and Y. Mallikarjuna Reddy. 2023. "Deep Learning Based Real-Time Object Detection on Jetson Nano Embedded GPU." Pp. 511–21 in *Microelectronics, Circuits and Systems*, edited by A. Biswas, A. Islam, R. Chaujar, and O. Jaksic. Singapore: Springer Nature.

Pahar, Madhurananda, Marisa Klopper, Robin Warren, and Thomas Niesler. 2021. "COVID-19 Cough Classification Using Machine Learning and Global Smartphone Recordings." *Computers in Biology and Medicine* 135:104572. doi: 10.1016/j.compbiomed.2021.104572.

Rymarczyk, Tomasz, Tomasz Cieplak, Grzegorz Kłosowski, and Edward Kozłowski. 2020. "The Use of LSTM Networks in the Detection of Outliers in IoT- Based Air Quality Monitoring Systems." *Przeglad Elektrotechniczny* 5:91–94.

Serrurier, Antoine, Christiane Neuschaefer-Rube, and Rainer Röhrig. 2022. "Past and Trends in Cough Sound Acquisition, Automatic Detection and Automatic Classification: A Comparative Review." *Sensors* 22(8):2896. doi: 10.3390/s22082896.

Shin, Woochang. 2020. "Implementation of Cough Detection System Using IoT Sensor in Respirator." *International Journal of Advanced Smart Convergence* 9(4):132–38. doi: 10.7236/IJASC.2020.9.4.132.

Singh, Meena, M. A. Rajan, V. L. Shivraj, and P. Balamuralidhar. 2015. "Secure MQTT for Internet of Things (IoT)." Pp. 746–51 in *2015 Fifth International Conference on Communication Systems and Network Technologies*.

S.K, Prashanthi, Sai Anuroop Kesanapalli, and Yogesh Simmhan. 2022. "Characterizing the Performance of Accelerated Jetson Edge Devices for Training Deep Learning Models." *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 6(3):44:1-44:26. doi: 10.1145/3570604.

Styła, M., P. Adamkiewicz, K. Niderla, and T. Rymarczyk. 2021. "Indoor Navigation System Using Radio Tomography." *Journal of Physics: Conference Series* 1782(1):012035. doi: 10.1088/1742-6596/1782/1/012035.

Takahashi, Shinya, Tsuyoshi Morimoto, Sakashi Maeda, and Naoyuki Tsuruta. 2004. "Cough Detection in Spoken Dialogue System for Home Health Care."

Wang, Hui-Hui, Jia-Ming Liu, Mingyu You, and Guo-Zheng Li. 2015. "Audio Signals Encoding for Cough Classification Using Convolutional Neural Networks: A Comparative Study." Pp. 442–45 in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*.

Yassein, Muneer Bani, Mohammed Q. Shatnawi, Shadi Aljwarneh, and Razan Al-Hatmi. 2017. "Internet of Things: Survey and Open Issues of MQTT Protocol." Pp. 1–6 in *2017 International Conference on Engineering & MIS (ICEMIS)*.

Internet sources:

The COUGHVID crowdsourcing dataset: A corpus for the study of large-scale cough analysis algorithms. (2021, February 3). Zenodo. https://zenodo.org/record/7024894#.Y7fzWS8w3C0