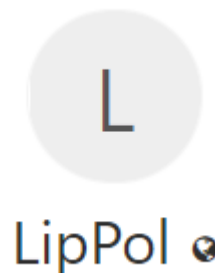




### KATEDRA ALGORYTMÓW I MODELOWANIA SYSTEMÓW

<b>Zespół projektowy:</b> 4@KAMS'2018	1. Natalia Moszczyńska – kierownik 2. Sławomir Strehlau 3. Michał Kowalewski 4. Jan Majkutewicz 5. Paweł Kraśnicki
<b>Opiekun:</b>	dr inż. Krzysztof Manuszewski
<b>Klient:</b>	dr inż. Krzysztof Manuszewski
<b>Data zakończenia:</b>	20.02.2019
<b>Słowa kluczowe:</b>	rozpoznawanie mowy, sieć neuronowa, wizemy



### TEMAT PROJEKTU:

**Opracowanie rozwiązania pozwalającego na czytanie z ruchu warg**

### CELE I ZAKRES PROJEKTU:

Celem projektu było opracowanie rozwiązania pozwalającego na czytanie z ruchu warg w języku polskim.

Zakres projektu:

1. Analiza istniejących opracowań
2. Stworzenie bazy próbek uczących w języku polskim
3. Stworzenie i nauczenie sieci neuronowej
4. Trening sieci na zebranych próbkach i eksperymenty
5. Dokumentacja

### OSIĄGNIĘTE REZULTATY:

1. Zebranie i przygotowanie danych: gotowych zbiorów i odpowiednich filmów
2. Przegląd literatury na temat rozpoznawania mowy z ruchu warg
3. Stworzenie środowiska umożliwiającego uczenie systemów do rozpoznawania mowy z ruchu warg z wykorzystaniem reprezentacji wizemowej
4. Stworzenie, nauczenie i przetestowanie sieci neuronowej rozpoznającej wizemy (z niezadowalającą dokładnością)

### CECHY CHARAKTERYSTYCZNE ROZWIĄZANIA, KIERUNKI DALSZYCH PRAC:

Cechy charakterystyczne:

- Wypróbowano nową metodę pozwalającą (w założeniu) na uczenie rozwiązań dla języków o ubogich zbiorach filmów
- Zebrano bogaty zbiór danych uczących w języku angielskim
- Przygotowano wygodne w użyciu środowisko do uczenia rozwiązań opartych na reprezentacji wizemowej

Kierunki dalszych prac:

- Udoskonalenie sieci klasyfikującej wizemy
- Dodanie odtwarzania zdania na podstawie ciągu wizemów
- Wzbogacanie zbiorów danych
- Zastosowanie nowych metod uczenia maszynowego



**Department of Algorithms  
and Systems Modelling**

<b>Project team:</b> 4@KAMS'2018	1. Natalia Moszczyńska – leader 2. Sławomir Strehlau 3. Michał Kowalewski 4. Jan Majkutewicz 5. Paweł Kraśnicki
<b>Supervisor:</b>	dr inż. Krzysztof Manuszewski
<b>Client:</b>	dr inż. Krzysztof Manuszewski
<b>Date:</b>	20.02.2019
<b>Key words:</b>	Speech recognition, neural network, visemes



**PROJECT TITLE:**

**A solution for lip-reading**

**OBJECTIVES AND SCOPE:**

The goal of the project was to develop a solution for lip-reading speech in Polish.

Scope of the project:

1. Analysis of existing solutions
2. Preparation of a dataset of videos in Polish
3. Design and implementation of a neural network
4. Training and experimentation on the neural network using the previously prepared datasets
5. Documentation

**RESULTS:**

1. Data gathering and preparation: previously made datasets and a new dataset of videos in Polish
2. An overview of literature on the topic of lip-reading
3. An environment for training lip-reading systems using a viseme representation
4. Design, training and evaluation of a neural network for recognizing visemes (the quality turned out to be unsatisfactory)

**MAIN FEATURES, FUTURE WORKS:**

Features:

- A new method was tried for creating lip-reading solutions for languages with a small number of speech samples
- Rich dataset of videos in English
- An easy to use environment for training lip-reading systems based on a viseme representation

Future research directions:

- Improvements to the viseme classifier neural network
- A separate neural network for converting sequences of visemes to sentences
- Enriching the datasets
- Utilizing new machine learning methods